Copyrighted Material

VOLUME II

Handbook of INTERNATIONAL TRADE

Economic and Legal Analyses of Trade Policy and Institutions

Edited by

E. KWAN CHOI

JAMES C. HARTIGAN

Blackwell Publishing

Copyrighted Material

Handbook of International Trade
Volume II

Blackwell Handbooks in Economics

Handbook of International Macroeconomics
Edited by Frederick van der Ploeg

Handbook of Environmental Economics
Edited by Daniel W. Bromley

Handbook of the Economics of Innovation and Technological Change
Edited by Paul Stoneman

Handbook of Applied Econometrics, Vol. I: Macroeconomics
Edited by M. Hashem Pesaran and Michael R. Wickens

Handbook of Applied Econometrics, Vol. II: Microeconomics
Edited by M. Hashem Pesaran and Peter Schmidt

Handbook of International Trade
Edited by E. Kwan Choi and James Harrigan

Handbook of International Trade, Vol. II:
Economic and Legal Analyses of Trade Policy and Institutions
Edited by E. Kwan Choi and James C. Hartigan

Handbook of International Trade

Volume II

Economic and Legal Analyses of Trade Policy and Institutions

Edited by

E. Kwan Choi and James C. Hartigan

© 2004 by Blackwell Publishing Ltd
except for editorial material and organization © 2004 by E. Kwan Choi and
James C. Hartigan

BLACKWELL PUBLISHING
350 Main Street, Malden, MA 02148-5020, USA
108 Cowley Road, Oxford OX4 1JF, UK
550 Swanston Street, Carlton, Victoria 3053, Australia

The right of E. Kwan Choi and James C. Hartigan to be identified as the Authors of the Editorial Material in this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

First published 2004 by Blackwell Publishing Ltd

Library of Congress Cataloging-in-Publication Data

Handbook of international trade / edited by E. Kwan Choi and James Harrigan.
p. cm. — (Blackwell handbooks in economics)
Includes bibliographical references and index.
ISBN 0–631–21161–6 (hardcover : alk. paper)
1. International trade. I. Choi. Eun Kwan. 1946–. II. Harrigan. James. III. Series.
HF1379.H364 2003
382—dc21 2002155271

A catalogue record for this title is available from the British Library.

Set in 10/12 Times
by Newgen Imaging Systems (P) Ltd, Chennai, India
Printed and bound in the United Kingdom
by
MPG Books, Bodmin, Cornwall

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on
Blackwell Publishing, visit our website:
http://www.blackwellpublishing.com

Contents

List of Figures vii

List of Tables ix

List of Contributors xi

Introduction 1
E. Kwan Choi and James C. Hartigan

1 What is Free Trade?: The Rorschach Test at the Heart of the Trade and Environment Debate 5
David M. Driesen

2 Rules of Power in an Age of Law: Process Opportunism and TRIPS Dispute Settlement 42
Ruth L. Okediji

3 Teaching Old Laws New Tricks: The Legal Obligation of Non-Attribution and the Need for Economic Rigor in Injury Analyses Under US Trade Law 73
James P. Durling and Matthew P. McCullough

4 Trade-Related Labor and Environment Rights Agreements? 107
Chantal Thomas

5 A Comparative Analysis of Compliance Institutions in International Trade Law and International Environmental Law 134
Brett Frischmann

6 The National Treatment Principle in International Trade Law 185
Michael J. Trebilcock and Shiva K. Giri

7 Do not Ask Too Many Questions: The Institutional Arrangements for Accommodating Regional Integration within the WTO 239
Petros Constantinos Mavroidis

8 Trade and Informal Institutions 279
James E. Anderson

9 The Economics of Preferential Trade Agreements 294
Pravin Krishna

10 Conditionality, Separation, and Open Rules in Multilateral Institutions 313
Paola Conconi and Carlo Perroni

11 Antitrust Policy in Open Economies: Price Fixing and International Cartels 333
Eric W. Bond

12 Modern Commercial Policy: Managed Trade or Retaliation? 358
Thomas J. Prusa and Susan Skeath

13 Antidumping versus Antitrust: Trade and Competition Policy 383
Ian Wooton and Maurizio Zanardi

14 Trade and the Globalization of Patent Rights 403
Rod Falvey, Feli Martinez, and Geoff Reed

15 Mixed Markets with Counterfeit Producers 427
E. Kwan Choi

16 Endogenous Injury 459
James C. Hartigan

17 International Trade in Services: More Than Meets the Eye 472
Lawrence J. White

18 The Dynamic Effects of Trade Liberalization and Environmental Policy Harmonization 499
Larry Karp and Jinhua Zhao

19 Do Bilateral Tax Treaties Promote Foreign Direct Investment? 526
Bruce A. Blonigen and Ronald B. Davies

Index 547

Figures

Figure 9.1	Trade creating tariff preferences: change in welfare = (3 + 4)	296
Figure 9.2	Trade diverting tariff preferences: change in welfare = (3 − 2)	297
Figure 9.3	Change in welfare for home = −(1 + 2 + 3 + 4); change in welfare for partner = (1 + 2 + 3); and change in welfare for union = −(4)	299
Figure 11.1	Profitable cartel outputs for $n^h = n^f = 2$	336
Figure 11.2	Efficient frontiers at $t^h = 0$ (B′C′) and $t^h > 0$ (D′E′)	337
Figure 11.3	Sustainable profit levels with $\delta = 0.7$ (DEFG) and $\delta = 0.5$ (IJK)	340
Figure 12.1	Worldwide use of antidumping, 1980–98	362
Figure 12.2	Antidumping actions consistent with alternative hypotheses	369
Figure 12.3	Traditional users (Specification B, all years)	376
Figure 12.4	Traditional users (Specification B, first three years)	377
Figure 12.5	New users (Specification B, all years)	378
Figure 12.6	New users (Specification B, first three years)	379
Figure 13.1	Countries with AD laws and membership of GATT/WTO (as of December 31, 2001)	385
Figure 13.2	The geography of AD laws	386
Figure 13.3	AD initiations from 1980, distinguishing nontraditional users	387
Figure 13.4	Average tariffs and definitive AD measures for nontraditional users	390
Figure 14.1	Patent protection under autarky	406
Figure 14.2	Welfare economics of product patents in a two-country model	409
Figure 15.1	Risk aversion and Nash equilibrium	434
Figure 15.2	Law enforcement and Nash equilibrium	438
Figure 15.3	Adverse effect of law enforcement	438
Figure 15.4	Stackelberg equilibrium	440
Figure 15.5	Entry of new counterfeiters	443
Figure 15.6	Long-run Nash equilibrium	444
Figure 15.7	Long-run survival of a Stackelberg monopolist	445
Figure 15.8	Duopoly equilibrium with counterfeiters	447

Figure 18.1 Possibilities of steady states. (a) Autarky and (b) Trade 507
Figure 18.2 Structure of the economy 510
Figure 18.3 The extraction function: $\delta_1 > \delta_2$ 512
Figure 18.4 Possibilities of going from autarky to trade 515
Figure 18.5 Resource dynamics with multiple steady states. (a) Autarky: $\hat{\eta}^a < \eta < \eta^{*a}$ and (b) Trade: $\hat{\eta} < \eta < \eta^*$ 518
Figure 18.6 Possible ranking schemes of critical η values 519

Tables

Table 3.1	Estimating the effect of price of imports on the domestic price of cold-rolled sheet	91
Table 3.2	Magnitude of different factors in explaining domestic price declines	97
Table 12.1	Antidumping activity: 1980s versus 1990s	363
Table 12.2	Probit regression: Specification A	371
Table 12.3	Probit regression: Specification B	372
Table 12.4	Probit regression: Specification C	373
Table 12.5	Probit regression: Specification D	374
Table 13.1	Antidumping initiations, 1995–9	388
Table 13.2	Targets of antidumping initiations, 1995–9	389
Table 13.3	Stages of bilateral integration	397
Table 17.1	Total World Trade, 1980–2000	480
Table 17.2	World trade by regions, 1980–2000	481
Table 17.3	Ten leading country exporters and importers of goods and services, 2000	482
Table 17.4	Major categories of trade in services and ten leading country exporters/importers, 2000	483
Table 17.5	US trade in services, 1992–2002	484
Table 17.6	Sales of services through nonbank majority-owned affiliates, 2001	485
Table 17.7	Aggregate sales of services through nonbank majority-owned affiliates, 1986–2001	486
Table 19.1	FDI data coverage by OECD country	533
Table 19.2	Descriptive statistics of variables	535
Table 19.3	New treaties by OECD countries from 1983–92	537
Table 19.4	Estimated treaty effects on OECD outbound FDI stock and flows using OLS	538
Table 19.5	Estimated treaty effects on OECD outbound FDI stock and flows using fixed effects	541

Contributors

James E. Anderson received his BA from Oberlin College and his PhD from the University of Wisconsin, and has spent his career at Boston College. He has published extensively on the theory of international trade and trade policy. In recent years, he has focused on invisible trade costs such as those associated with low-quality institutions, and in particular on the endogeneity of such costs.

Bruce A. Blonigen is the Knight Professor of Social Science in the Economics Department at the University of Oregon. He received his PhD in economics in 1995 from the University of California-Davis. While completing his dissertation, he also worked as an economist for the Research Division of the Office of Economics at the US International Trade Commission. Prof. Blonigen has research interests in empirically examining international trade issues from a microeconomic and political economy perspective, especially with respect to multinational corporations and antidumping policies. His work has been published in such journals as the *American Economic Review*, *Review of Economics and Statistics*, and the *Journal of International Economics*. Prof. Blonigen is also an Associate Editor for the *Journal of International Economics* and a Research Associate with the National Bureau of Economic Research.

Eric W. Bond is the Joe Roby Professor of Economics at Vanderbilt University. He has written on a number of topics in international trade, including the design of international trade agreements, the strategic interactions between trading blocs, and the taxation of foreign investment. He received his PhD from the University of Rochester in 1979.

E. Kwan Choi is Professor of Economics at Iowa State University. He is currently the editor of *Review of International Economics* and *Review of Development Economics*. He has published articles in *Quarterly Journal of Economics* and *Journal of Political Economy*, as well as edited *Economic Growth and International Trade* (Blackwell Publishing, 2000) and *Handbook of International Trade* (volume I, Blackwell Publishing, 2004).

Paola Conconi's research interests are in the areas of International Trade, Regional Integration, Environmental Economics, and Political Economy. She is a Professor of

Economics at the Université Libre de Bruxelles (ECARES), a research fellow of the Centre for the Study of Globalization and Regionalism (CSGR) of Warwick University, and a Research Affiliate of the Centre for Economic and Policy Research (CEPR).

Ronald B. Davies joined the University of Oregon Economics Department as an assistant professor in 1999 after receiving his PhD in Economics from the Pennsylvania State University. His research interest focuses on the determinants of foreign direct investment with a particular emphasis on strategic tax policy. His work has appeared in *American Economic Review*, *International Economic Review*, and *Journal of Public Economics*.

David M. Driesen is an Associate Professor at Syracuse University College of Law, an affiliate of the Maxwell School of Citizenship Center for Environmental Policy and Administration, and an Adjunct Associate Professor of the State University of New York College of Environmental Science and Forestry. He holds a JD from the Yale Law School.

James P. Durling is a partner in the International Trade Practice of Willkie Farr & Gallagher LLP. His practice focuses on the various US laws affecting the pricing practices of foreign companies, including antidumping law, countervailing duty law, safeguard measures, transfer pricing under Section 482, and US customs law. He is also actively involved in World Trade Organization (WTO) matters, both advice about WTO obligations and dispute settlement proceedings. Mr Durling received his law degree (JD) in 1984 from New York University School of Law, where he was an articles editor of the *NYU Law Review*. He also received a masters degree (MPA) from the Woodrow Wilson School at Princeton University in 1984, where he studied international economics. He graduated from Haverford College with a BA degree in 1980. He is a member of *Phi Beta Kappa*. In addition to various professional associations for lawyers, Mr Durling is a member of the American Economics Association.

Rod Falvey is Professor of International Economics at the University of Nottingham. He is a graduate of the University of Canterbury (New Zealand) and the University of Rochester. His research interests lie in the general area of trade theory and policy.

Brett M. Frischmann joined the faculty of Loyola University of Chicago, School of Law in 2002. Prof. Frischmann graduated Order of the Coif from the Georgetown University Law Center, where he was an Olin Research Fellow in Law and Economics and the Executive Development Editor for the *Georgetown International Environmental Law Review*. After graduating from law school he was an associate with the law firm of Wilmer, Cutler & Pickering in Washington, DC and a judicial clerk for the Honorable Fred I. Parker of the US Court of Appeals for the Second Circuit. Prof. Frischmann has written on a wide variety of topics, including the economics of science and technology policy, international emissions trading, copyright misuse, privatization of the internet infrastructure, judicial decision making in cyberlaw disputes, and the role of compliance institutions in international law.

Shiva K. Giri is a doctoral student at the University of Toronto Law School and is writing a thesis on the National Treatment principle in International Trade Law.

James C. Hartigan is a Professor of Economics at the University of Oklahoma. He formerly taught at SUNY-Buffalo and the Pennsylvania State University. He has had visiting positions at Linkoping University in Sweden, Dalian Institute of Technology in China, and the Australian National University. He has served on the Board of Editors of the *Journal of Economic Integration* and the *Pacific Economic Review*, as well as the Council of Editors of the *Review of International Economics*. He also edited a Special Issue of the *JEI* on *Antidumping Laws and Their Enforcements*. He was a founding member of the International Economics and Finance Society, and served as its secretary-treasurer. He has published in such journals as the *Review of Economics and Statistics*, the *Journal of International Economics*, and *Economica*.

Larry Karp is Professor of Agricultural and Resource Economics at the University of California, Berkeley. His research interests include resource and environmental economics, international trade, and industrial organization. Recent publications include articles on the regulation of pollution under asymmetric information, the effect of imperfect property rights under international trade, dynamic consistency and government policy, and hyperbolic discounting. These papers have appeared in the *Journal of Public Economics, Journal of Economic Theory, International Economic Review, Journal of Economic Dynamics and Control*, and *Journal of Environmental Economics and Management*.

Pravin Krishna is Professor of Economics at Brown University and a Research Fellow at the National Bureau of Economic Research. He has also held appointments at Stanford University, the University of Chicago and Princeton University. Prof. Krishna conducts research on international economics, political economy and development. He has published articles in a number of scholarly journals including the *Journal of Political Economy*, the *Quarterly Journal of Economics* and the *Journal of International Economics*. Prof. Krishna holds a bachelors degree in engineering from the Indian Institute of Technology, Bombay, and a PhD in economics from Columbia University.

Feli Martinez had degrees in Law from the universities of La Laguna and Valencia (Spain), and an MA and PhD in Economics from the University of Nottingham. At the time of her death she held a lectureship in Law at the University of Leicester. Her research interests were in intellectual property rights in open economies.

Petros C. Mavroidis is Professor of Law at Columbia Law School, New York, University of Neuchatel and a member of the Centre of Economic Policy Research (CEPR). He is the chief co-reporter of the American Law Institute (ALI) project "Principles of International Trade: The WTO."

Matthew Paul McCullough is an associate in the International Trade Practice of Willkie Farr & Gallagher LLP. He has over a decade of experience working on trade

policy matters, particularly those related to antidumping, countervailing duty and safeguards issues, steel, US–Japan and US–Korea trade relations. Mr McCullough's work on trade remedy cases has involved a number of high profile dumping, countervailing duty and safeguards investigations in the United States, appeals before the US Court of International Trade and challenges before World Trade Organization (WTO) dispute settlement panels and the WTO Appellate Body. He received his law degree from Catholic University's Columbus School of Law in Washington, DC (JD 2002, *magna cum laude*). He earned his undergraduate degree from Austin College in Sherman, Texas (BA 1991).

Ruth L. Okediji is the William L. Prosser Professor of Law at the University of Minnesota Law School. She specializes in international intellectual property law, copyright and Internet Regulation. She is the author of numerous articles on intellectual property rights in developing countries and the international intellectual property system.

Carlo Perroni's research interests include international trade, environmental regulation, and public finance. His research on international policy cooperation has focused on multilateral trade negotiations, climate treaties, and interjurisdictional tax coordination. He is Professor of Economics at the University of Warwick.

Thomas J. Prusa is Professor of Economics at Rutgers – The State Univeristy of New Jersey and is a Research Associate of the National Bureau of Economic Research. He has written and published extensively on the economics of dumping and antidumping policy. He is on the editorial boards of *Review of International Economics* and *Journal of International Economics*.

Geoff Reed is Reader in Economics at the University of Nottingham, where he gained his BSc. His research interests are in trade policy and agricultural economics.

Susan Skeath (van Mulbregt) is Professor of Economics and Chair of the Department of Economics at Wellesley College where she has been on the faculty since 1989. She is a graduate of Haverford College and holds the MA and PhD degrees from Princeton University. Prof. Skeath is the author of a number of papers on international trade theory and antidumping and is also co-author of the introductory level game theory text "Games of Strategy."

Chantal Thomas is Professor of Law at Fordham University School of Law. Recent publications include articles published in the *Harvard Journal on Legislation, Washington & Lee Law Review*, and the *Journal of International Economic Law.*

Michael J. Trebilcock is a Professor of Law and Economics at the University of Toronto, where he teaches international trade law, contract law, and law and development.

Lawrence J. White is the Arthur E. Imperatore Professor of Economics at the Stern School of Business, New York University. He has taken leave from NYU to serve

in the US government three times: during 1986–9 he was a board member on the Federal Home Loan Bank Board; during 1982–3 he was the Chief Economist of the Antitrust Division of the US Department of Justice; and in 1978–9 he was a Senior Staff Economist on the President's Council of Economic Advisers. Among his publications is *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation* (1991), and he is the co-editor (with John E. Kwoka, Jr.) of *The Antitrust Revolution: Economics, Competition and Policy* (fourth edition, 2004).

Ian Wooton is Professor of Economics at the University of Strathclyde, Glasgow, and is a Research Fellow of the Centre for Economic Policy Research. He has a PhD in Economics from Columbia University and previously held posts at the universities of Western Ontario and Glasgow. He is an established authority in international economics and has published widely in the area. He is the co-founder of the European Trade Study Group (ETSG), a major forum for international trade research.

Maurizio Zanardi holds an MA and PhD in Economics from Boston College and is currently an Assistant Professor in Economics and a member of the Center for Economic Research (CentER) at Tilburg University. He previously held a lectureship in Economics at the University of Glasgow. His research interests focus on international trade and the political economy of trade policies.

Jinhua Zhao is an Associate Professor of Economics at Iowa State University. He received a PhD from the University of California at Berkeley in 1997. His research interests include real option theory, technology adoption, dynamic welfare measurement, mechanism design, international trade and the environment, and global warming.

Introduction

E. Kwan Choi and James C. Hartigan

When we first discussed the possibility of organizing and editing a *Handbook* for Blackwell, we quickly decided to include works from both legal scholars and economists in the volume. The increasing availability of abstracts on the Internet had encouraged us to keep apprised of legal scholarship concerning international trade. As part of our growing attention to legal scholarship, we became more aware of the lack of cross-disciplinary citations between international law and international economics. In fact, the *Journal of International Economics* published five issues so far in 2002 (January, March, June, August, and October). Of the 1326 citations appearing in those issues, only 12 appear to qualify as legal sources. This is 0.90 percent of the total. The *Review of International Economics* published three issues so far in 2002 (February, May, and August). Of the 764 citations in those issues, six appear to qualify as legal documents. This is 0.79 percent of the total.

The legal profession does better. In the June 2002 issue of the *Journal of International Economic Law* (*JIEL*), 11.97 percent of the citations by the authors are to economics sources. In the June 2002 issue of the *Journal of World Trade* (*JWT*), 15.06 percent of the citations are to economics sources. However, this is not quite as encouraging as it may appear, as 84.9 percent of the citations to economics sources in the *JWT* occur in two of the eight articles it contains. For the *JIEL*, 77.8 percent of the citations to economics sources appear in one of the six articles in that issue. Nonetheless, it does seem that legal scholars are more receptive to economics scholarship than international economics scholars are to legal analysis.

The barrier to entry to economics for international legal scholars seems fairly obvious: the extent of mathematical formality endemic to the discipline. What was not apparent to us was the barrier to entry to legal scholarship. We believed that it must have something to do with the style of presentation. We assumed initially that it must be the prevalence of footnotes in legal scholarship. However, we found that in reading the contributions to this volume, the footnotes quickly ceased to be a distraction. Only after reading all of the legal contributions did we decide what the barrier was. It was that economists are not comfortable with the subtlety and ambiguity of language.

A primary objective of this book is to increase cross-disciplinary fertilization. As such, we requested that each of the contributors keep the counterpart discipline in mind

when writing their chapter. The economists were asked to use as little mathematical formality as possible, and they were very receptive to this request. Not having as much experience in law as we did in economics, we were less specific in our requests to the contributors from that discipline. However, the legal participants are sympathetic to economic analysis, and have provided chapters that we believe are interesting and accessible to economists.

In selecting topics to be represented in the volume, we were guided by two criteria: (1) the topic must be of significant interest to both disciplines, and (2) the topic must be at the forefront of current research in international trade policy. In satisfying the first criteria, we eschewed issues such as national sovereignty, which are important to legal scholars, but do not get much attention from economists. We also avoided topics, such as explanations for trade patterns, which are important to economists, but not legal scholars. To the extent possible, we included an author from each discipline for the subjects discussed in the volume.

Because antidumping (AD) duties are the only form of protection that is still increasing, and because their use is proliferating beyond the traditional utilizers (the United States, Australia, the European Union, Canada, and New Zealand) to include many less-developed countries (LDCs), we felt that it was important to include experts from both disciplines to address this issue. From the legal profession, James Durling and Matthew McCullough discuss the material injury investigation/decision of the US International Trade Commission. In particular, they focus upon the legal obligation to not attribute other causes of injury to imports in this investigation. From economics, Thomas Prusa and Susan Skeath discuss economic and strategic motives for the filing of AD complaints, finding support for the hypothesis that countries use AD petitions to deter future use or punish past use against them. Ian Wooton and Maurizio Zanardi analyze the interface of AD and competition policies, concluding that reducing reliance on AD will require increased supranatural coordination of antitrust policies. James Hartigan provides a model of cyclical reciprocal dumping with simultaneous innovation. While providing a justification for an economics-based injury decision in unfair trade investigations, he contends that its requirements are not likely to be met frequently.

A chapter blending very nicely with Wooton and Zanardi is Eric Bond's economic analysis of trade liberalization and its relationship to the behavior of international price fixing cartels. Bond also addresses the issue of international rule formation in the application of national competition policies.

One of the most significant achievements of the Uruguay Round was the extension of General Agreements on Tariffs and Trade (GATT) discipline to trade in services. The result was the General Agreement on Trade in Services (GATS). Economist Lawrence White provides a documentation of the growing importance of international trade in services, and a discussion of the achievements of the GATS.

Another issue of increasing importance in international trade is regionalism, or the formation of Preferential Trading Areas (PTAs). Petros Mavroidis provides a legal analysis of the compatibility of PTAs with the World Trade Organization (WTO) contract. As his chapter considers PTAs and dispute settlement, it also blends with Ruth Okediji's chapter discussed below. Pravin Krishna offers an economic analysis of PTAs, and provides conditions for the institutional design of a PTA to be welfare improving.

Another significant achievement of the Uruguay Round was the agreement on trade and intellectual property known as TRIPS. Ruth Okediji provides a legal perspective, with a particular emphasis upon compliance and dispute settlement. Rod Falvey, Feli Martinez, and Geoff Reed have contributed an economic analysis of TRIPS, focusing upon global patent enforcement. This chapter, in and of itself, is an example of the interdisciplinary cross-fertilization for which this volume is striving, as Feli Martinez has legal training.

Brett Frischmann's legal analysis of compliance institutions in international trade law and international environmental law integrates effectively with several other chapters in this volume. Because compliance is a close relative of dispute settlement, this chapter has an interface with those of Mavroidis and Okediji. It also complements the chapters addressing trade and the environment, to be discussed below. Further, it invokes game theoretic analysis, a cornerstone technique of economists.

Kwan Choi has contributed a chapter on economics that is in the interface of the TRIPS literature and the literature addressing counterfeit products. That is, he analyzes the market equilibria that may arise when a producer of a good entailing intellectual property must compete with firms that copy the product illegitimately. He then contrasts the results arising when imitation is tolerated with those when intellectual property is enforced.

As the WTO has been under pressure to permit members to address environmental issues through trade policy, including a softening of Most Favored Nation and National Treatment obligations, we cannot imagine not including chapters on this subject in a volume of this nature. Economists Larry Karp and Jinhua Zhao highlight the complexity of this matter in a dynamic model of international trade that is based upon different natural resource stocks, different degrees of environmental resiliency, and different regulatory policies regarding environmental exploitation. They disclose that, under various plausible assumptions, trade can be welfare enhancing or diminishing. With a focus upon trade and environment, economists Paola Conconi and Carlo Perroni discuss multilateral institutional forms for linking issues in international cooperation. That is, they consider the possibilities for cooperation when countries can negotiate binding agreements with different partners along more than a single policy dimension. Legal scholar Chantal Thomas analyzes the institutional competence of the WTO to address labor and environmental issues, and contrasts the lack of progress on these matters with the success in obtaining a TRIPS agreement in the Uruguay Round. This chapter also complements the aforementioned chapters addressing TRIPS. Legal scholar David Driesen examines the interface of international trade and regulation of the environment in the context of a much more general objective of defining free trade. Although economists typically view free trade as entailing nondiscrimination among sources of supply, Driesen discusses whether or not its definition should be extended to include freedom from attempts by one WTO member to impose its regulatory regime upon another.

In a thorough analysis and critique of panel and Appellate Body case law, Michael Trebilcock and Shiva Giri advocate criteria for determining whether or not products are "like," that is, based upon existing or potential competitive relationships between products. Their discussion of National Treatment highlights the extent to which facially neutral tax or regulatory measures can have protectionist implications.

Although the WTO has had limited success in the negotiation of Trade Related Investment Measures, it is generally recognized that foreign investment flows can have a significant impact upon trading patterns. Using OECD data, economists Bruce Blonigen and Ronald Davies examine the relationship between tax treaties and foreign investment. They suggest that recent tax treaties may be designed to reduce tax evasion rather than to promote foreign investment.

Economist James Anderson analyzes the endogenous relationship between the formation of institutions and trade. This chapter provides a deep fundamental background to the policy issues discussed elsewhere in this volume.

We hope that the reader deems this volume to be a success, and is stimulated by the research being done in this exciting area. Both disciplines can be strengthened by recognition of each other's insights and contributions.

1

What is Free Trade?: The Rorschach Test at the Heart of the Trade and Environment Debate

David M. Driesen

Chapter Outline

This chapter argues that a fundamental question, "what is free trade?," lurks behind the ongoing debate about the relationship between international trade law and competing legal regimes. Although the literature contains volumes about the reasons for free trade, it says remarkably little about free trade's definition.

This chapter explores three possible concepts of free trade, trade free from discrimination against foreign companies, trade free from coercion, and trade free from restraint, that is, laissez-faire, primarily in the context of trade and environment disputes. The misunderstanding between environmentalists and free traders reflect trade law's tendency to amalgamate the antidiscrimination, anti-coercion, and laissez-faire concepts. Free traders tend to think of trade law as primarily aimed at policing discrimination, while environmentalists tend to think of it as aimed at laissez-faire, the least legitimate concept. The trade law provides some support for both views.

1 Introduction

A large literature addresses relationships between free trade and other policy areas that trade law increasingly affects, including environmental law,[1] intellectual property,[2] labor relations,[3] human rights,[4] and competition policy.[5] These materials rarely include

a precise definition of "free trade."[6] They do not answer a crucial question, what precisely must trade be free of in order to be "free" rather than inappropriately shackled? This chapter addresses that question.

Instead of defining free trade, scholars seem to assume that "free trade" has an obvious (although unspecified) meaning.[7] Decisions interpreting the General Agreement on Tariffs and Trade (GATT)[8] and academic writing use vague phrases like "trade barriers,"[9] "trade restrictions,"[10] and "protectionism,"[11] to describe that which trade should be free of. But these phrases, absent clarification, may be broad enough to collectively embrace almost any regulation or commercial tax serving competing values, as demonstrated below.[12]

This failure to articulate a normatively attractive and clear legal concept of free trade leaves the World Trade Organization (WTO), the administrator of the GATT and related multilateral trade agreements,[13] unable to defend its legitimacy in a convincing manner.[14] Increasing tension between the WTO and other legal regimes has made the question of the WTO's legitimacy quite salient.[15] Decisions holding environmental and public health regulations contrary to GATT have contributed to paralyzing division among WTO member governments and triggered a campaign by non-governmental organizations (NGOs) to stop new trade talks.[16]

A decade that witnessed the WTO's creation and a significant expansion of international trade law has brought the WTO into conflict with international and domestic environmental law. During this decade, the WTO became increasingly concerned with "nontariff trade barriers." This creates enormous potential for conflict, because, in a globally integrated world, most regulations and commercial taxes might be described as nontariff trade barriers, since they burden commercial activity, much of which is international.

In the early 1990s, two GATT panels held the unilateral imposition of a ban on tuna imports caught in a manner that unduly endangers dolphins, contrary to GATT.[17] More recently, WTO dispute resolution panels held an import restriction aimed at protecting endangered sea turtles contrary to GATT and a European ban on the sale of beef injected with growth hormones contrary to the Agreement on Sanitary and Phytosanitary Measures (SPS),[18] another WTO-administered trade agreement.[19] A stream of articles and books addressing the proper relationship between free trade and environmental protection followed the "Tuna/Dolphin" and "Shrimp/Turtle" decisions, but rarely addressed the definition of free trade.[20]

This inattention to first principles may reflect the formal legal structure of GATT, which imposes a set of trade disciplines upon contracting parties, rather than explicitly requiring free trade. Nevertheless, free trade provides the normative justification for the WTO and the agreements it administers, and differing concepts of free trade sometimes help explain the results of cases interpreting trade agreements. Hence, an adequate legal concept of free trade would greatly enhance the debate about the WTO.

An analysis of possible definitions shows that the *ad hoc* and uncertain nature of trade law stems from a failure to choose a clear, limited, and coherent concept of free trade from among the available alternatives, rather than from theoretical necessity.[21] Current trade law amalgamates three different ideas about what trade should be free of. Article III of GATT's text reflects a concept of free trade as trade free of laws,

both taxes and regulations, which discriminate between foreign and domestically produced goods.[22] But, this chapter will argue that the Tuna/Dolphin and Shrimp/Turtle decisions implicitly rely upon an anticoercion concept of free trade, that is, trade unimpeded by efforts to enforce even nondiscriminatory environmental law (or other bodies of nontrade law) against noncomplying nations.[23] The WTO took a step toward an even broader concept of free trade, as trade free of national regulation under a broad laissez-faire conception, when it adopted the SPS agreement during the Uruguay Round of trade negotiations.[24]

Since trade law conflates three different ideas of what free trade is, trade law appears quite *ad hoc* and difficult to justify. Because these ideas are not equal in their normative attractiveness and their implications for other legal regimes, free trade becomes something of Rorschach test. Commerce advocates identify free trade with the most normatively appealing idea, that of nondiscrimination, and environmentalists tend to identify it with the least normatively appealing idea, laissez-faire government.

Section 2 of this chapter begins by identifying the roots of the ambiguity in the legal concept of free trade in the classical economics of Adam Smith and David Ricardo. It then develops three concepts of free trade based on: the principle of nondiscrimination; an international noncoercion principle; and a principle of laissez-faire government. This section describes the theoretical support for these concepts, identifies some of their sources in international trade law, and elucidates their implications for focusing efforts to expand free trade. It closes by using the concepts to help explain why the concepts of "trade barriers" and "trade restrictions" cannot adequately substitute for a definition of free trade.

Section 3 applies these concepts to show how they illuminate scholarly and judicial efforts to justify the WTO. It shows that application of these concepts yields fresh insights into the most important trade and environment cases, helps explain continued misunderstandings between free traders and environmentalists, and reframes the ongoing trade and environment debate. It concludes that a nondiscrimination concept offers the most hope for advancing acceptance of the WTO beyond the world of economists and trade specialists.

Much of the scholarly commentary about the trade and environment issue has framed the debate as one about "exceptions" to free trade. The inquiry into the meaning of free trade invites more critical thinking about the GATT and SPS trade disciplines themselves, not just the exceptions to GATT disciplines. This chapter focuses on trade and environment issues in order to make the topic manageable, but the analysis offered here will contribute to the broader "trade and . . ." debate, as well as general discussions of WTO legitimacy.[25]

2 The Law and Theory Supporting the Nondiscrimination, International Noncoercion, and Laissez-Faire Concepts

Economists have, over the years, sometimes defended free trade as an extension of the principle of laissez-faire and sometimes as a principle of avoiding a somewhat narrower set of "distortions."[26] This section will show that the ambiguity that my conceptual

framework identifies in the legal concept of free trade has its roots in the problems addressed in the writings of David Ricardo and Adam Smith.[27] It will then sketch out the three concepts and their sources.

2.1 Classical Roots of the Ambiguity

Economists write volumes about the reasons for "free trade," but often say very little about its definition.[28] Adam Smith, in "The Wealth of Nations," advanced the argument that efforts to protect a country's producers by banning or levying high tariffs upon imports would not only harm the nation making the taxed or banned goods, but also the nation imposing the restriction.[29] David Ricardo refined Smith's insights into a more nuanced theory of comparative advantage.[30] The theory holds that free trade would allow each country to make that which it is best suited to make, thereby increasing worldwide consumption.[31] This theory articulates the reasons why free trade offers benefits.[32]

Smith and Ricardo's work have less to say about what exactly free trade is. Smith's theories constitute an extended argument against the mercantilist system of his day.[33] This system levied high protective tariffs or banned imports outright as an economic strategy.[34] Smith's work showed that this strategy was economically counterproductive.[35]

The dominant view of Smith and Ricardo's work holds that it extends the arguments for a laissez-faire theory to trade among nations.[36] Their work supports this view. Smith, for example, characterizes his endorsement of the navigation acts (trade restrictions to advance national security) and compensatory taxation (taxation of imports compensating for other country's taxation of exports) as "limitations" upon the principle of free trade.[37] Similarly, Ricardo discusses a "system of perfectly free commerce" implying trade with no burdens whatsoever.[38] This suggests that the free trade principle really involved absolute license to trade without any impediments or restrictions. This would imply no commercial taxes or regulations, at least upon goods traded internationally and the processes that produce such goods.

Analysis of Smith's policy recommendations, however, shows that they fit a model of nondiscrimination in trade relations better than they fit a laissez-faire model. The import bans and high protective tariffs that Smith opposed discriminated against imports, since they applied to imports, but not to competing domestic industry. Smith endorsed compensatory taxation and general taxation for legitimate public purposes, positions at odds with strict laissez-faire, but consistent with antidiscrimination.[39]

Ricardo's work focuses more on the mechanics of comparative advantage and less on policy recommendations. Since he does not adamantly oppose taxes, Ricardo too does not really endorse laissez-faire in a strict sense either. In discussing taxes upon produce, for example, he states "the sum required by the taxes must be raised."[40] He then claims that a produce tax would not "materially interfere with foreign trade."[41] At the same time, he strikes a laissez-faire note in stating that the tax "would . . . prevent the very best distribution of the capital of the whole world . . .."[42] On balance, he treats taxation as necessary, not as a trade restraint to be abolished.

This analysis reveals a problem with the classical foundation for free trade. A principle different from the laissez-faire principle, which Smith and Ricardo are known for, best accounts for their policies. This problem matters a great deal for legal theory, because a definition of free trade must help guide institutional policy decisions to function as a useful legal concept.

Either a laissez-faire or a nondiscrimination principle justifies abandonment of mercantilist policies, which involve discriminatory government activism. The ambiguity, therefore, mattered little to the argument against mercantilism. The two concepts, however, diverge sharply in their implications for modern environmental and health regulations. And Smith and Ricardo, not surprisingly, have little to say about modern regulation that addresses health and environmental concerns.

The theory of comparative advantage does not directly provide a definition of free trade. Rather, the theory explains why international trade takes place and how it provides benefits.[43] In saying this, I do not deny that the theory's insights have proven useful in thinking about the definition of free trade, at least as an economic concept.

To see this, it will help to review the role the theory of comparative (and absolute) advantage plays in the case for free trade. International trade takes place because a foreign producer can offer some advantage to domestic consumers over domestic goods.[44] The foreign good may cost less than domestic substitutes or offer superior quality. A domestic purchaser may also purchase a foreign good because no domestic substitute exists. In order to supply a good that costs less, offers better quality, or does not exist in the purchasing country, the foreign producer must have some advantage (such as lower labor cost or superior technology) in making that good. The advantages may be either absolute or comparative. Comparative advantages arise when the pretrade ratios of prices of different goods vary. This makes it possible for a country to enjoy comparative advantage without necessarily having an absolute advantage in any one good's production. As Alan Sykes has recently explained, "the theory of comparative advantage offers the predominant explanation of why such circumstances arise."[45]

But the theory does more than simply describe why international trade takes place. It shows how international trade produces advantages for both trading partners.[46] Of course, demonstrating the advantages of international trade strengthens the case against anything that interferes with those advantages, but specifying the things that interfere with those advantages requires additional analysis. For example, suppose that trade arises because a foreign country offers a superior product, because it has a better-educated workforce. Suppose further that the home country finances improvements in its educational system that will improve the workforce so that it can produce an equally good product. If one assumes that the costs of these products are the same, should one conclude that an improvement in an educational system interferes with free trade by subsidizing elimination of an advantage? Nobody argues that improvements in an educational system interfere with free trade. But an explanation as to why not involves more than the theory of comparative advantage.

Economists typically do not rest their views about what trade should be free of directly upon a definition of free trade.[47] Rather, economists employ general equilibrium models to evaluate the costs and benefits of various policy interventions.[48] In so doing, they employ an allocative efficiency test that does not differ fundamentally from that employed to recommend "optimum" domestic policies.[49]

Applying this to health and environmental regulations, neoclassical economic principles would suggest that the benefits of such regulations should equal their costs.[50] Application of this principle, however, poses numerous practical and theoretical problems.[51] Not surprisingly, trade law does not apply a cost–benefit analysis to challenged regulations.

This summary suggests two possible definitions of free trade. One might think of free trade as trade free of burdens, a broad laissez-faire principle. One might, on the other hand, think of free trade as trade free of discrimination.

2.2 GATT Article III: Trade Free from Discrimination

Since its negotiation in 1947, the GATT has formed the basis for much of international trade law,[52] with 133 countries agreeing to abide by the GATT eventually.[53] GATT Article III supports “free trade,” defined as trade free of discrimination against foreign goods as a tool of economic policy.[54] Article III read in isolation would suggest that GATT seeks to facilitate international trade – and thereby spread prosperity – by establishing a principle of nondiscrimination against foreign goods.[55] GATT’s preamble emphasizes nondiscrimination[56] and the WTO provides a forum for lowering tariffs.[57] WTO member governments commit themselves to the principle of “national treatment” for imports, a requirement that taxes and regulations not discriminate between foreign and domestic goods without an adequate noneconomic justification.[58] Members must also provide other GATT contracting parties with the same treatment they provide the “most-favored” nation with which they trade, a limited principle of nondiscrimination between foreign trading partners.[59]

Although GATT’s text lacks a definition of discrimination, a working definition will help clarify the concept. One might define discrimination as imposition of a standard or restriction on imports that one does not impose upon one’s nationals.[60] A concept of free trade as trade free of discrimination against foreign producers implies a focus upon tariff reduction, elimination of regulations and taxes that expressly discriminate between foreign and domestic goods, termination of subsidies that apply to only domestic manufacturers of products (thereby discriminating against imports), and abolition of import quotas.[61]

2.3 The Sanitary and Phytosanitary Agreement, Article XI, and Article XX’s Evisceration: The Laissez-Faire Concept

One can define free trade more broadly than trade free of discrimination. We might mean by free trade, trade unencumbered by national laws that might increase prices, such as taxes and regulation.[62]

GATT Article XI: 1 offers the potential for a substantial move toward laissez-faire government.[63] Article XI generally prohibits “quantitative restrictions” upon exports or imports.[64] One might construe this article narrowly to embrace import quotas and

little else, rendering it consistent with a nondiscrimination principle. But the WTO has interpreted it broadly to apply to any border measure imposing any burden upon international trade.[65] This implies that any violation of the laissez-faire principle administered at the border offends GATT Article XI: 1.

While Article XI in isolation would go far toward establishing a laissez-faire concept, the Ad Note to Article III should limit Article XI's push toward laissez-faire. Trade experts agree that the Ad Note to Article III acts as a defense to claims that product regulations applied at the border are *per se* violations of Article XI.[66] It subjects such regulations to Article III's national treatment obligation in lieu of the rule of *per se* invalidity that generally applies to trade restrictions under Article XI.[67] Hence, the scope of Article III and its Ad Note determines the limits that apply to Article XI's push toward laissez-faire government.

GATT contains a set of defenses in Article XX that arguably reflects a conscious choice to leave decisions about the appropriate scope of national regulation to advance at least citizens' noncommercial welfare to national governments.[68] These defenses, assuming that they have meaning, would allow a country to otherwise impose GATT illegal trade restrictions when they meet Article XX's requirements.[69] In other words, Article XX would allow quantitative restrictions on trade and discriminatory regulation of foreign commerce under some circumstances. These exceptions apply to environmental laws.[70]

Trade panels, however, have usually construed these provisions very narrowly.[71] As a result, only one panel has ever upheld a health or environmental regulation under an Article XX defense, the panel adjudicating a challenge to French asbestos regulation.[72] And on appeal, the WTO's appellate body determined that this regulation had not offended the GATT trade disciplines in the first place.[73] While GATT does not expressly embrace a laissez-faire philosophy, the evisceration of Article XX defenses makes it quite difficult to identify meaningful limits to a WTO panel's ability to pursue a broad laissez-faire agenda indirectly.

The Shrimp/Turtle case rejected a very broad anticoercion rationale that might automatically eliminate any possibility of an Article XX defense.[74] But this decision struck down the measure before it and it is too soon to tell whether subsequent panels will regularly allow Article XX defenses to validate otherwise GATT illegal environmental measures.

Even if a government regulation complies with all relevant GATT trade disciplines or somehow manages to satisfy the WTO's interpretation of Article XX, the recent SPS agreement invites WTO panels to second guess national government's claims that the problem a regulation addresses warrants a regulatory remedy.[75] And a recent WTO panel decision did precisely that, declaring illegal a European Community restriction on beef from cattle injected with hormones, some of which had been found to cause cancer in laboratory animals.[76] A panel of trade experts with no expertise in public health concluded that the European Community had failed to show that the hormones in the banned beef posed a significant risk.[77] The WTO's Appellate Body affirmed the panel decision, while reversing some of its subsidiary rulings.[78]

The new SPS agreement, as interpreted so far by the WTO, creates hurdles for governments applying nondiscriminatory, but strict, standards to protect public health.[79] Governments wishing to enact stricter standards than existing advisory

international standards must base their standards on a risk assessment.[80] WTO panels will scrutinize national regulations that determine whether risk assessments "reasonably support" the regulatory measure at stake.[81] The Beef/Hormone Appellate Body acknowledged, in dictum, national governments' right to regulate on the basis of minority scientific views.[82] But it held that the single divergent opinion of a well-respected scientist could not justify the regulatory program before it, because the scientist did not himself carry out research directly addressing hormone residues in beef fattened with hormones.[83] It also apparently held that a government cannot regulate carcinogens without scientific studies addressing the specific application of the carcinogen it banned, at least in the face of the studies of expert opinion finding the disputed application "safe."[84] Finally, it rejected an apparently undisputed body of research identifying misapplication of growth hormones as a problem. The panel found the handful of studies on this issue "insufficient" to constitute a risk assessment of that issue.[85] This would suggest that governments cannot, under the SPS agreement, permanently regulate any problem that has not been studied extensively, even when there is little scientific controversy about it.

The Appellate Body stated, in dictum, that the SPS does not require quantification of risk.[86] But its holdings, both in Beef/Hormone and subsequent cases, cast doubt on whether any measure based on a qualitative assessment of limited information could pass muster.[87]

Judicial scrutiny of scientific justifications can cripple regulatory programs where great scientific uncertainty exists.[88] Because of ethical limitations on controlled human experimentation, precise data about the effects of contaminants at all levels on human beings usually does not exist.[89] In this context, burdens of proof can become critical.[90] Whichever party bears the burden of proof in a case with totally incomplete data has a good chance of losing.[91]

The WTO has placed the burden of proof on regulating governments. The Beef/Hormone Appellate Body reversed a panel decision that imposed the burden of proof upon regulating governments in all cases.[92] But the Appellate Body's decision may still support regular application of the burden of proof to regulators. The Appellate Body endorsed shifting the burden to the regulating party once the complaining party establishes a *prima facie* violation of the SPS Agreement.[93] While the Appellate Body did not articulate a set of principles defining a *prima facie* violation, the case may support finding a *prima facie* violation any time a bona fide scientific dispute exists about the relationship between a risk assessment and an adopted measure.[94] Because such disputes are inevitable when little direct data exists about human exposure at various levels (a very common situation), a *prima facie* case of a violation may exist frequently.[95]

Subsequent WTO panels may regularly require regulators to affirmatively prove that specific evidence directly supports their standards, rather than show some deference to government inferences from incomplete data or require complaining parties to show that regulated substances are safe.[96] If this occurs the SPS agreement could significantly impede regulation, because complete data exists about very few potentially significant public health problems.[97] In addition, the SPS agreement generally requires WTO members to use the least trade restrictive means available to protect

public health.[98] This least restrictive means test also provides a significant laissez-faire element.[99]

Defining free trade as trade free of government regulations or taxes requires efforts to expand free trade to focus on weakening regulations designed to protect the public health or advance any other values that compete with sales of goods at lower prices. Preliminary steps might involve creating burdens that governments must meet in order to impose regulations – precisely what the SPS agreement has done. While the WTO has not embraced laissez-faire government as an explicit goal, the WTO has taken a substantial step in that direction.[100]

2.4 The Tuna/Dolphin and Shrimp/Turtle Decisions: Trade Free from International Coercion

Governments frequently employ trade restrictions for reasons other than protection of domestic industry. Governments generally seek to resolve disputes about a wide range of matters, including national security, natural resource conservation, public health and safety, international trade, and human rights, through negotiation. Negotiation does not always solve the problems it addresses and countries may feel compelled to use various forms of coercion to achieve results. International institutions usually lack coercive power to enforce international legal obligations. Hence, national governments must generate solutions to international problems. Many countries employ trade restrictions instead of war when negotiation fails.[101] Trade restrictions have played an important role in the development of international law and in international policy.

Trade sanctions have encouraged governments to further develop GATT, to become GATT contracting parties, and to comply with its terms.[102] GATT reflects a specific decision to countenance unilateral trade measures to encourage nonparties to undertake GATT commitments. GATT's drafters could have specified that GATT's requirements apply to all goods that a GATT contracting party imports. This would have required GATT signatories to persuade nonparties through negotiation to sign GATT. But GATT requirements only apply to imports from GATT contracting parties, and contracting parties remain free to restrict imports from nonparties.[103] Hence, GATT's drafters decided to allow GATT contracting parties to restrict imports from nonparties, instead of drafting a provision forbidding restrictions against any country's exports.[104]

This feature of GATT played an important role in its development. The possibility of imposition of trade restrictions against nonparties coupled with the promise of escape through signature of GATT provided incentives for many countries to become parties.[105]

GATT contracting parties, however, may not always honor their GATT obligations, so the regime needs an enforcement mechanism. Prior to the recent creation of the WTO, GATT incorporated a dispute settlement process that relied upon consensus adoption of dispute resolution panel decisions.[106] As a result, the losing party generally could block implementation of a GATT panel decision by simply opposing its adoption.[107]

The United States increased its reliance on unilateral trade restrictions to leverage favorable resolution of trade disputes under section 301 of the Trade Act of 1974,[108] in part, to address failures to enforce GATT panel decisions.[109] Partially in order to escape the pressures from unilateral imposition of trade sanctions, GATT contracting parties agreed to create the WTO in 1994.[110] Hence, unilateral trade sanctions have played an important role not just in attracting new GATT contracting parties, but also in securing international agreement to strengthen GATT enforcement.

The principal procedural difference between the WTO and the prior GATT organization involves the nature of dispute settlement. The Dispute Settlement Understanding adopted as part of the Uruguay Round of GATT negotiations requires GATT's contracting parties to adopt a WTO panel decision (or an appellate decision following a panel decision), unless the parties reach consensus *against* adoption.[111] This means that WTO decisions will generally bind parties to the dispute.

The agreement creating the WTO explicitly authorizes the use of trade sanctions, called "suspension of concessions," to enforce decisions of panels established under the WTO to resolve disputes regarding alleged breaches of GATT obligations.[112] Hence, the WTO itself adopts the principle of using coercion to achieve trade goals.[113]

Competing legal regimes may, like the WTO regime, find trade measures useful in achieving their goals. As global integration proceeds, national governments experience a loss in their power to assure adequate environmental quality for their own people, unless they can influence conduct abroad that harms their environment or that of the global commons (e.g., oceans) upon which they depend.[114] This has led to the growth of international environmental law. Governments have threatened import restrictions (and occasionally export restrictions) to encourage the development of agreed upon international environmental standards and compliance with the terms of adopted standards.[115] None of these uses of trade sanctions necessarily involve protectionist economic strategy, but instead employ trade restrictions as a strategy to meet noneconomic goals.[116]

Steve Charnovitz has pointed out that unilateral trade restrictions aimed at encouraging multilateral action have preceded the adoption of most significant health and environmental treaties (just as they preceded the formation of the WTO).[117] Multilateral agreements to impose trade sanctions also encourage parties to join international environmental treaties. Parties to the Montreal Protocol on Ozone-Depleting Substances, for example, included a provision that barred imports of ozone-depleting chemicals from countries that did not agree to limit their production of ozone depleters under the Protocol.[118] This provision, in combination with other provisions providing more positive incentives for compliance, tended to discourage transfer of ozone-depleting chemical production from parties to the Protocol to rapidly developing nonparties.[119] Such a transfer might have otherwise defeated efforts to control ozone-depleting chemicals, destroyed the stratospheric ozone layer, and created a public health and environmental catastrophe.[120]

International environmental law has grown in the last few decades, but now suffers from "treaty proliferation."[121] As a result, governments have often agreed to treaties with rather broadly expressed obligations, but often have not taken the actions necessary to meet treaty objectives.[122]

Prior to the 1990s, GATT seemed to treat international efforts to limit pollution and destruction of natural resources through trade restrictions as GATT compliant. During the development of the Montreal Protocol on Substances that Deplete the Ozone Layer,[123] a "legal expert from the . . . GATT secretariat" stated that GATT did not forbid the use of sanctions to encourage nonsignatories to comply with the agreements.[124]

Some governments and organizations effectively used trade sanctions and import bans to encourage compliance with existing treaties.[125] The United States, for example, made unilateral threats of import restrictions to encourage governments to comply with international fisheries agreements.[126] Many nations, including several underdeveloped countries, responded to the pressure by improving conservation practices.[127] The European Union has also used import bans to meet environmental and public health goals.[128] The Convention on International Trade in Endangered Species of Flora and Fauna (CITES) generally forbids international trade in listed endangered species and their parts without a permit.[129] The treaty has enjoyed some success because it aims squarely at limiting trade as a strategy.[130] Agreements seeking to generally protect biodiversity that do not rely on trade sanctions have often achieved little.[131]

Countries also sometimes impose trade restrictions unilaterally to enforce environmental standards. The United States, for example, used unilateral trade restrictions to encourage several nations to abandon tuna fishing techniques that killed dolphins.[132] In the 1990s, however, first Mexico and then the Netherlands and the European Economic Community mounted challenges to a United States ban on tuna imports caught with purse seine nets in a manner that would kill many dolphins. GATT panels held the ban contrary to GATT.[133] At the time, GATT required adoption of a panel decision by consensus of member countries in order for the decision to bind member governments, and this never occurred.[134] Even though adopted panel decisions lack formal precedential value under GATT law, these decisions greatly influenced the WTO.[135]

The two Tuna/Dolphin decisions' central rationales have little to do with a concept of free trade as trade free from discrimination against imports. The decisions seem instead to uphold a principle limiting international coercion aimed at advancing competing policy goals.[136] The Tuna/Dolphin decisions both criticize the use of national coercion through trade restrictions to force foreign countries to protect dolphins.[137] Similarly, the Shrimp/Turtle appellate decision held that the United States may not seek to force other countries to adopt regulatory regimes identical to those of the United States through trade restrictions.[138] The Shrimp/Turtle appellate panel considered the "coercive effect" of the United States' turtle protection program its "most conspicuous flaw."[139] These panel decisions implicitly define free trade as including a principle of noncoercion, at least through trade measures.

The theoretical support for a noncoercion principle in international law comes not from economic theory, but from theories of international relations.[140] Scholars have debated extensively on the appropriateness of coercion in various contexts and situations.[141] Widespread agreement exists that nations should not resort to coercion without first attempting to resolve differences through negotiation. However, most scholars recognize that when negotiation cannot resolve important disputes, then some degree of coercion may be appropriate. Disagreement arises, of course, about when to employ coercion and how much coercion is appropriate in various situations. The practice of international relations seems consistent with the theory since nations typically

try to solve disputes through negotiation and resort to coercion only as a last resort. Countries vary in their willingness to employ coercion to resolve different issues when negotiations break down.

Countries can coerce each other without discrimination. For example, a country may demand compliance with an environmental standard as a condition of importation and impose an identical standard on its own domestic producers. Such a standard may coerce, but it does not necessarily discriminate.

Defining free trade as trade free of the effects of international coercion would create a different agenda for expansion of free trade. Expansion of free trade would involve decreasing reliance upon coercive measures to advance policy goals. The WTO moved in this direction in the Tuna/Dolphin and Shrimp/Turtle decisions.

2.5 Trade Restrictions and Barriers

Free traders often state that they do not oppose environmental protection; they only oppose trade barriers and restrictions as the means of protecting the environment.[142] This suggests that one can protect the environment adequately without trade barriers and restrictions.[143] The three concepts provide useful analytical tools to help understand why the concepts of trade barriers and trade restrictions do not offer an acceptable substitute for an adequate definition of free trade.

A laissez-faire definition of "free trade" as involving absolute license – trade without any burdens – might well require the elimination of almost all even-handed national regulation and taxation of business, all international coercion, and all regulation and taxation discriminating against foreign commerce. Only regulation that discriminates against domestic production for the domestic market can avoid creation of burdens upon international trade.[144] Any international coercion, any discrimination against imports, and most even-handed government tax or regulation burdens international trade.

The statement that even-handed taxation and regulation creates burdens for international trade requires some explanation. Even-handed taxation and regulation implies taxes and regulations that apply equally to all relevant businesses, including importers and exporters. Any tax or regulation that applies to all relevant products sold in the taxing or regulating jurisdiction may increase the cost of imports entering the jurisdiction. Any tax or regulation of production processes that applies to all relevant production within the taxing or regulating jurisdiction may increase the cost of goods that the jurisdiction exports. For these reasons, even-handed taxation and regulation burdens international trade.

Governments may tax or regulate only domestic producers that produce only for the domestic market without burdening international trade. But a country that taxes or regulates even-handedly (i.e., that does not systematically discriminate against companies with no involvement with international trade) will often create burdens upon international trade. This means that as international integration proceeds, even-handed regulation and taxation creates more and more burdens upon international trade.[145] A jurisdiction with no international trade could even-handedly tax and regulate everything sold in the jurisdiction with no impact upon international trade. At the other extreme, if all sectors have some involvement with international trade, then all

even-handed commercial regulation and taxation burdens international trade. More integration implies greater burdens upon international trade from routine domestic regulation and taxes.

As global integration proceeds, demands not to use trade restrictions or barriers become indistinguishable from a demand for laissez-faire government. The terms "trade restriction" and "trade barrier" plausibly apply to almost every tax and regulation directly affecting business. The vague idea that international trade law should eliminate "trade restrictions" or "trade barriers" implies rejection of all taxes and regulation, and the acceptance of all three concepts of what free trade is. It embraces laissez-faire, anticoercion, and antidiscrimination as free trade goals. Of course, a full embrace of laissez-faire makes other goals unnecessary.

Most regulations that apply equally to imported and domestically produced goods and services include a prohibition on sales and/or shipment made without obeying the regulation. To make this ban on sales of noncompliant goods effective, almost every statute providing for domestic regulation of products either explicitly or implicitly forbids importation of goods that do not comply with national laws.

An enormously wide range of federal laws rely upon these import restrictions, including criminal, intellectual property, transportation, telecommunications, national security, health and safety, and conservation laws. Since almost all federal regulation rests upon federal constitutional authority to regulate commerce,[146] these prohibitions usually couple a ban on interstate sales or shipment of noncompliant goods with a ban on importation of noncomplying goods.[147] Some statutes employ less explicit formulations (such as a ban on introduction or delivery for introduction into interstate commerce) that still use import restrictions as part of a ban on sales or shipment of noncomplying goods.[148] However, almost all regulations applicable to goods rely upon the threat of a sales ban to secure compliance. If this threat is removed, then regulation becomes virtually impossible to enforce. If a person can sell goods without complying with applicable regulations, then she probably will do so. Hence, regulations must forbid the import or export of noncomplying goods in order to be effective.

Once one realizes that virtually all "domestic" regulations rely upon coercion the distinction between a laissez-faire concept and an anticoercion concept begins to collapse. Any domestic regulation, insofar as it applies to imported goods, aims to coerce a foreign country (or its nationals) to make goods acceptable to the regulating jurisdiction, upon pain of a ban.[149] Regulations that commentators tend to regard as domestic, and usually legitimate, and those that some regard as extraterritorial, and therefore potentially illegitimate under an anticoercion definition, function identically from the standpoint of direct burdens upon international trade and both involve coercion.

The WTO may have realized in the Shrimp/Turtle case that all regulation involves coercion. The WTO Appellate Body reversed a panel ruling broadly prohibiting countries from requiring exporting countries to meet the importing countries' policies in order to obtain market access.[150] In doing so, the Appellate Body recognized that "conditioning access to" an importer's "domestic market on" the exporting country's compliance with an importing country's "unilaterally prescribed" policy "may, to some degree, be a common aspect of measures falling within the scope" of the Article XX exceptions to GATT.[151] Because Article XX exceptions to GATT cover most subjects of

national regulation, this statement suggests a broad recognition that regulation generally involves coercion. Hence, a free trade principle based on opposition to international coercion and a free trade principle based on allegiance to laissez-faire principles might have very similar results. All regulation and taxation burdens international trade and usually imposes quantitative restrictions on noncompliant shipments.

Nondiscriminatory even-handed regulation, however, burdens economic activity in general; it does not increase burdens upon international trade beyond those imposed upon similar domestic economic activity. Insofar as either national or international regulatory efforts rely upon import restrictions as an enforcement mechanism, the exporting country or company may export anyway. But in order to do that, the exporting company or country must meet the importing country's regulatory standards, hence the burden. Similarly, a domestic firm making a product for the domestic market must comply with the same regulatory standard in order to sell within its own market. Both domestic and foreign firms have the same burden. Discrimination against imports, however, constitutes a burden on international trade that does not apply to commercial activity in general. The very logic of discrimination disadvantages foreign products.

In sum, almost every tax and regulation of business constitutes a trade barrier or restriction in a global economy. A laissez-faire or anticoercion definition of free trade might imply elimination of most taxes and regulations applicable to business, since all regulations and taxes coerce and most burden trade economically. An antidiscrimination concept functions more narrowly, systematically ferreting out especially problematic treatment of international trade, as opposed to general taxation and regulation of commerce.

3 Explaining Judicial and Scholarly Efforts to Justify the WTO

Understanding the core features of the three concepts helps clarify the trade and environment debate, and these concepts help explain how trade panels seek to legitimate the WTO. Trade panels often try to justify their decisions in terms of nondiscrimination, even when other concerns seem to drive the decisions. They also regularly disavow laissez-faire goals. This disjunction between the decisions' dicta and the actual grounds helps explain why so much disagreement exists about whether the WTO poses a serious threat to competing values. The three-part model provides an alternative means of thinking about possible reconciliation between free trade goals and competing noneconomic goals.

3.1 Judicial Disavowal of Laissez-Faire and Embrace of Antidiscrimination

As previously discussed, concerns about international coercion best explain the Tuna/Dolphin and Shrimp/Turtle decisions. Similarly, a perceived need to have national governments affirmatively justify even nondiscriminatory regulation drives the

Beef/Hormone decision. But these decisions feature explicit, and not always credible, repudiation of laissez-faire goals and use of strained antidiscrimination arguments.

3.1.1 DISCRIMINATION TALK IN COERCION CASES

The first Tuna/Dolphin panel strained to use antidiscrimination rhetoric to defend its holding. The panel held that the Marine Mammal Protection Act (MMPA) did not constitute a regulation of a product triggering Article III's national treatment obligation.[152] Since Article III did not apply to the case, the panel did not need to address the question of whether the MMPA discriminated against foreign tuna in violation of Article III. Nevertheless, the panel, in dicta,[153] accused the United States of discriminating against Mexican tuna in violation of Article III.[154] It stated that the MMPA "provided treatment to tuna . . . from Mexico that was less favourable than the treatment accorded to like United States tuna . . . ."[155] In light of the fact that the United States imposed requirements on its own fleet similar to those it demanded of foreign countries, this statement seems at least questionable. The panel justified it by claiming "Article III: 4 . . . obliges the United States to accord treatment to Mexican tuna no less favorable than that accorded to United States tuna, whether or not the incidental taking of dolphins by Mexican vessels corresponds to that of United States vessels."[156] Thus, the panel converted an apparently nondiscriminatory regulation into a de jure discriminatory regulation.

At the end of its opinion, the Tuna/Dolphin I panel added "concluding remarks" returning to antidiscrimination.[157] The panel emphasized that "a contracting party is free to tax or regulate imported products and like domestic products as long as its taxes or regulations do not discriminate against imported products . . . ."[158] And it admonished the United States stating, "[A] contracting party may not restrict imports of a product merely because it originates in a country with environmental policies different from its own."[159] These statements suggest that the MMPA was GATT illegal because it discriminated against Mexican tuna, when, in fact, the antidiscriminatory rationale only appears in dicta.

Furthermore, the panel disclaimed any allegiance to laissez-faire goals. It not only emphasized parties' freedom to tax or regulate absent discrimination against foreign products, it also emphasized parties' freedom "to tax or regulate domestic production for environmental purposes."[160] But the Tuna/Dolphin principle, that only allows a party to tax or regulate domestically, prohibits some nondiscriminatory regulation and requires that any regulation addressing an international problem discriminates against domestic producers. Extending an even-handed prohibition to both domestic and foreign producers of a product is precisely what the panel prohibited, at least insofar as the regulation addresses processes. Nevertheless, the panel evidently judged it necessary to disclaim any allegiance to laissez-faire.

Similarly, the Shrimp/Turtle Appellate Body strained to find discrimination. The panel report, which the Appellate Body reviewed, followed Tuna/Dolphin and held that the ban on shrimp harvested without turtle excluder devices violated GATT Article XI.[161] Since the United States did not appeal this finding, appellate review focused on the issue of whether an Article XX exception justified a violation of Article XI's prohibition upon quantitative restrictions.[162] The Appellate Body held that

this measure constitutes "unjustifiable discrimination," which disqualifies the measure from enjoying an Article XX defense based on the language in that article's chapeau.[163]

The Appellate Body used several novel arguments to justify this holding. Because the United States required turtle excluder devices of both its own fleet and foreign fleets as a condition of market access, it appeared to engage in no discrimination at all against foreigners as a class.[164] The Appellate Body, however, held that the "cumulative effect" of a number of aspects of the turtle conservation program made this apparently even-handed approach discriminatory.[165]

It implied that a failure to discriminate in favor of foreign exporters constituted discrimination against them. The Appellate Body noted that the regime requires exporting countries to adopt "*essentially the same* policies and enforcement practices as the United States."[166] The Appellate Body recognized that "the United States also applies a uniform standard throughout its territory,"[167] but found such rigidity unacceptable in foreign relations.[168] The panel suggested that the United States must allow the sale of shrimp from abroad, caught without turtle excluder devices, even though the MMPA requires forfeiture of domestic shrimp catches made without the devices.

This kind of argument sounds in anticoercion. The opinion's call for flexibility suggests an embrace of a model of international relations based on flexible negotiations alone.

But the Shrimp/Turtle Appellate Body converted this anticoercion argument into an argument that uniform even-handed regulation is generally discriminatory. It found even-handedness inappropriate, because firm even-handed regulation does not involve considering that "different conditions . . . may occur" in the exporting country.[169] It then stated "that discrimination results not only" from differential treatment, but also from measures that do "not allow for any inquiry into the appropriateness of the regulatory program for the conditions prevailing in those exporting countries."[170]

Nothing in the Shrimp/Turtle decision identifies a single condition in any country that makes a program of installing turtle excluder devices unreasonable outside of the United States.[171] In any case, the Appellate Body converted an argument for flexibility in international relations into a strained argument about discrimination.

The Appellate Body also found the United States' failure to negotiate treaties to protect the sea turtles with some of its trading partners inappropriate, because the failure to negotiate means that the procedures and policies became unilateral, rather than multilateral.[172] This sounded like a foreign relations argument based on a conception of free trade as trade free of coercion.

The Appellate Body also converted this coercion argument into a discrimination argument. Since the United States negotiated turtle protection with some trading partners, but not the appellees, the Appellate Body stated that the "effect is plainly discriminatory . . . ."[173] Although this theory of discrimination is plausible, the Appellate Body could have simply held that the United States must negotiate with all trading partners or none of them, without some justification, thereby focusing on discriminatory treatment, rather than discriminatory effect. This focus on the "effect" of creating "unilateralism," however, allowed it to cloak an argument against unilateralism, a species of coercion, in antidiscrimination garb.

Having created a need for more individualized (and therefore less transparent) bureaucratic decisions by rejecting uniform standards, the Appellate Body argued

for a kind of due process of international relations.[174] It criticized the administrative process of certifying turtle conservation programs for the failure to provide a hearing, an opportunity for rebuttal, or a reasoned decision.[175]

The Appellate Body converted this argument for due process in international relations into a discrimination argument. It claimed that because of the lack of procedural protections, the United States discriminates against applicants who do not receive certification "vis-à-vis those Members which are granted certification."[176] But the Appellate Body did not claim that the countries that received certificates of compliance with US regulations received the procedural protections denied the unsuccessful applicants. The successful applicants may simply have had a better case on the merits, because they were willing to use turtle excluder devices. Nothing in the Appellate Body's decision directly argues to the contrary. Hence, the discrimination argument appears strained.

The Shrimp/Turtle decision, like the Tuna/Dolphin decision, closes with a disavowal of the laissez-faire definition of free trade. The Appellate Body claimed that WTO members might adopt effective measures to protect endangered species.[177] But, says the WTO, the United States has applied this ban on shrimp caught without turtle excluder devices in an unjustifiably discriminatory manner.[178]

As explained previously, the heart of these decisions has nothing to do with antidiscrimination. The Tuna/Dolphin panel concluded that import bans that any country can escape by simply adopting the environmental practices of the importing state constitute prohibited quantitative restrictions, rather than regulations, under GATT. That conclusion rested almost entirely upon an anticoercion rationale. Notwithstanding the Shrimp/Turtle decision's extended nondiscrimination argument, the Appellate Body singled out the coercive nature of the shrimp ban as the turtle protection program's most egregious feature.[179]

3.1.2 Discrimination Talk in Laissez-Faire Cases

The Beef/Hormone panel also characterized a seemingly neutral regulation as discriminatory. The European Union applied its ban on growth hormones to both European and foreign beef producers.[180] So, once again, the regulation seems quite neutral.

The panel held that the measures at issue were invalid because they were not based on a risk assessment, rendering consideration of any discrimination issue unnecessary.[181] Yet, it held that the European Community discriminated by banning beef from cattle fed with growth-enhancing hormones, while not regulating naturally occurring hormones.[182] It also held that the ban on hormones in beef discriminated, because it did not apply to a different substance used in swine production.[183]

The Appellate Body reversed, finding the failure to regulate natural hormones justified, and held that the discrimination based on differential treatment of substances used in swine production did not constitute a "disguised restriction on international trade . . . ."[184] While WTO panels frequently shoehorn their rulings into findings of discrimination, the Appellate Body largely resisted the urge to do so here. The panel decision, however, conformed to the tendency to seek out antidiscrimination rationales.

The Beef/Hormone panel decision also illustrates the impulse to disclaim any movement toward laissez-faire free trade. The panel, after holding illegal a regulation of possibly carcinogenic substances for lack of an adequate scientific basis, tried to

claim that it had made no judgment about the necessity of the measure, but couldn't. It stressed "it was not our task to examine generally the . . . necessity" of the ban on growth hormones in beef.[185] In the same paragraph, however, the panel acknowledged making a judgment about "the necessity of the import ban . . . for the protection of human life or health."[186] It then tried to suggest that its ruling somehow left governments free to regulate domestically, as if a regular participant in international markets could generally regulate without affecting international trade. It stated that the ability of countries to regulate without affecting international trade "was not at issue in the present case."[187] It then feebly pointed out that it had not addressed nonhealth-related consumer concerns, thereby implying that it had left a field of regulation untouched.[188] In this way, the panel tried to draw attention away from the large swath of health-related regulations that its decision potentially implicated.

The Appellate Body may have recognized the futility of the panel's efforts to make the implications of its holding appear insubstantial. It narrowed the panel's legal reasoning, left its ruling intact, and dispensed with concluding remarks.[189]

This combination of holdings based on noncoercion and laissez-faire-related principles with rhetoric based on discrimination may help explain why observers disagree about whether the WTO poses a significant threat to environmental protection. Environmentalists may look at the logical implications of the principles directly supporting the holdings and see a grave threat. Defenders of free trade may take the limiting dicta in these cases very seriously, partly out of faith in the judgment of trade panelists.[190]

3.2 Refocusing the Trade and Environment Debate

Scholarly debates about tensions between free trade and competing policies generally treat the concept and scope of free trade as a given.[191] The debate then focuses upon what exceptions to free trade GATT should tolerate in order to accommodate the competing policy.[192]

In the legal academy, much of this debate takes the form of arguments about the appropriate scope and interpretation of Article XX exceptions to GATT trade disciplines.[193] The Tuna/Dolphin case sparked a debate about whether national regulation of the processes of foreign production was appropriate.[194] But much of the scholarship views this process/product distinction debate as another question involving exceptions to free trade.[195]

The three-part conceptual scheme calls into question the conventional structure of the trade and environment debate. From a legal perspective, it is not clear why the debate should be about exceptions to free trade. After all, if the WTO embraces a concept of free trade as trade free from discrimination, as is sometimes claimed, then we need a debate about expansion of the free trade concept beyond those bounds. The debate would be a debate about the browning, not the greening, of the GATT.[196] Furthermore, Article XX defenses simply do not apply to the most demanding trade disciplines in the regime, those found in the SPS agreement.

Surely the SPS agreement shows that the WTO is administering agreements that cumulatively move far beyond the problem of discrimination. The definitional issue

leads us to ask what concept of free trade the WTO implements. Asking this question leads to some fresh questioning of the nature of the GATT trade disciplines. One might ask why precisely the MMPA violates the GATT disciplines. This question, which is antecedent to the question of whether a defense applies, turns out to be rather difficult to answer.[197]

The Tuna/Dolphin panels concluded that the MMPA quantitatively restricts trade[198] and GATT Article XI generally forbids quantitative restriction of trade.[199] However, Tuna/Dolphin I offers no direct support for the conclusion that the MMPA establishes a quantitative restriction on imports.[200] Tuna/Dolphin II states that the "embargoes" constituted "prohibitions or restrictions" of importation under Article XI, "since they banned the import of tuna or tuna products from any country *not meeting certain policy conditions*."[201] Of course, this implies that any country meeting these policy conditions can export freely to the United States. Tuna/Dolphin does not explain why a measure, which allows any country to choose to export unlimited quantities of tuna (by choosing to comply with conservation standards), should be considered a *quantitative* restriction on trade.

The panels held that the MMPA did not involve the kind of regulation GATT authorizes in Article III.[202] They state that the MMPA did not regulate the characteristics of tuna as a product.[203] Since Article III only addresses regulation of products, the panels concluded that the MMPA provisions before them did not constitute a regulation within the meaning of that article. Therefore, the GATT requirement of national regulatory treatment did not apply.

The Tuna/Dolphin decisions suggest that the MMPA regulates the "process" of catching tuna and distinguishes process from product regulations.[204] Scholars debating these decisions have exhaustively discussed both the wisdom and legal soundness of this product/process distinction and this chapter will not revisit that issue.[205] But the argument that this law did not regulate tuna as a product, even if correct, does not establish the measure as a quantitative restriction.

The conclusion that a law allowing imports provided a country meets policy conditions quantitatively restricts trade, while having a certain surface plausibility, does not withstand analysis. Any regulation that applies to imported goods will necessarily prohibit imports in order to enforce the regulation. If the country could export the goods without compliance with the regulation, it would not have to obey the regulation. If the targeted country can comply with a regulation and thereby secure the right to export without any limit to quantity, then clearly a qualitative regulation, not a quantitative restriction, is at issue. Hence, the fact that the MMPA did not inexorably limit the quantity of tuna that Mexico could export to the United States seems to establish that it did not impose a quantitative restriction.

GATT scholars, however, apparently consider the holding that these regulations involve a *prima facie* violation of Article XI as so obviously correct as to require no explanation.[206] While Article XI on its face might support the notion that it only limits literal import quotas (including zero quotas), GATT panels have traditionally construed Article XI much more broadly to invalidate almost any regulation applied at the border that places a burden on imports.[207]

Tuna/Dolphin's narrow construction of the Ad Note to Article III made the MMPA illegal only because of the broad construction of Article XI. This broad construction

of Article XI goes beyond the antimercantilist limit on quotas necessary to sustain the nondiscrimination principle and embraces a laissez-faire rule limited only by applicable defenses. Hence, narrow construction of the Ad Note implies greater movement toward laissez-faire trade.

A historical view of the three-part conceptual scheme suggests a theoretical reason for viewing the trade and environment debate as a discussion of proposals for exceptions to free trade principles. Historically, the theory of free trade has often been viewed as a laissez-faire theory. From the perspective of this laissez-faire theory any proposal to allow regulation of the environment looks like a proposal for an exception to free trade principles, because regulations burden international trade. This may help explain why, over time, the WTO and GATT panels have construed Article XI so broadly and the defenses so narrowly.

The existence of a background laissez-faire assumption by itself has important implications for understanding "trade and . . ." debates. Since laissez-faire assumptions have very little normative attractiveness, any sense that they inform dispute settlement panels or scholarly opinions will make those opinions less persuasive. This may explain why GATT and WTO panels feel obliged to disclaim pursuit of laissez-faire goals. Without the convincing identification of a more attractive principle, claims that the WTO pursues something more limited than the maximum politically feasible move toward laissez-faire goals will prove unconvincing. Some scholars realize that the nondiscrimination principle may prove more attractive than a laissez-faire principle. Professors Farber and Hudec write, "Facial discrimination is the most attractive basis for GATT intervention . . .."[208] And they understand that thoughtful critics of the GATT regime will be worried about the SPS agreements' restrictions, previously described as moving toward laissez-faire.[209]

The three-part conceptual scheme facilitates inquiry into which principles actually explain the decisions and why laissez-faire and noncoercion principles appear more troubling than facial antidiscrimination principles. WTO decisions relying on nondiscrimination appear more legitimate than decisions on other grounds. Nondiscrimination seems rooted in fairness concerns and usually involves policy choices that an international trade institution is well placed to make. By contrast, the WTO lacks the democratic credentials and expertise to determine the appropriate degree of laissez-faire government. Similarly, trade experts seem poorly positioned to make decisions about the appropriate use of coercion. While some international control of coercion may be legitimate, trade experts will have a tendency to view fairly trivial coercion as a grave threat, even when it effectively serves ends deemed important by the government of the consumers paying the coercion's direct cost.

While legitimacy matters, functional considerations matter as well. If a WTO panel holds a law illegal under laissez-faire principles, the law might not recover. Of course, the WTO has only moved toward laissez-faire, not all the way to it. But requiring a nation to invalidate a law because it lacks an adequate scientific justification may have equally fatal consequences. If WTO judgments about the adequacy of a scientific judgment follow national consideration of the most important available evidence, then a compliant government must adopt weaker standards or eschew regulating the matter giving rise to the WTO judgment. Of course, to the extent that a WTO ruling under the SPS agreement follows significant oversights in examining available data, the government

may find itself able to regulate after improving the risk assessment. But that may be insufficient if the WTO demands greater certainty than the available data offers on the issue a national government wishes to regulate.

A noncoercion principle would endanger efforts to protect public health and the environment. Indeed, a pure anticoercion principle is very similar to a laissez-faire policy, since even-handed regulations and taxes coerce those exporting into the taxing and regulating jurisdiction. A pure noncoercion principle would abolish a great swath of public health regulation. However, a less radical principle, embraced by Tuna/Dolphin II, objects to coercion of national governments. Tuna/Dolphin II may implicitly distinguish regulations that coerce national governments from regulations that coerce private companies.[210]

A rule against coercion of foreign governments would invalidate the use of trade restrictions to force governments to police impacts of their nationals on the high seas or on global environmental problems. Since all international environmental laws rely upon domestic enforcement, this might have serious impacts upon international environmental protection, and international environmental protection is essential to a nation's ability to protect its own people from environmental harms as well as to the protection of other species.

If the WTO extends this noncoercion principle beyond the international environmental context in which it arose, such a principle could eviscerate domestic environmental law. A country, for example, might wish to assure that a product consumed in its jurisdiction does not poison its inhabitants. In order to make a regulation, banning this poison's use, effective in a globally integrated world, it must apply to imports and domestically produced goods alike. If the importing countries cannot effectively enforce such a regulation through inspections after the product is made, it may have to regulate the foreign manufacturing process.[211] A broad noncoercion rule might prevent a country from requiring a foreign country to supply data showing adequate enforcement of standards comparable to those that the national government imposed upon its own manufacturers. Hence, this noncoercion rule would prevent regulators from effectively implementing domestic measures that required some compliance verification abroad.

Measures conditioning access to domestic markets upon foreign governmental compliance with national standards have been part of domestic public health laws since the 1920s.[212] For example, the US Import Milk Act of 1927 seeks to assure a pure milk supply by requiring a sanitary inspection of the foreign dairy farm as a condition of importation.[213] Similarly, the US Food, Drug, and Cosmetic Act of 1938 prohibits commerce in food prepared, packed, or held under unsanitary conditions, and the US Poultry Products Inspection Act bars importation of poultry unless "subject to the same inspection . . . standards" as apply in the United States.[214] Hence, a rule against coercing foreign governments would invalidate many standard domestic regulations.

A pure rule against explicit discrimination, however, would only rarely have terribly significant environmental impacts. The country would usually be able to regulate equally effectively by simply applying the same standard it applied to imports to its own products. An example of this comes from the Reformulated Gasoline case. The United States could regulate stringently after the WTO ruling against its reformulated gasoline by extending to foreign manufacturers the same opportunities for using measurement techniques that domestic manufacturers had.[215] It subsequently modified its rule to

do that.[216] In short, a facial nondiscrimination requirement usually allows the country suffering WTO reversal to continue effectively addressing the public health or environmental problem at issue. Declaring measures illegal on laissez-faire or anticoercion grounds poses considerably deeper problems for environmental regulation.

A proposal to focus exclusively upon free trade as trade free of discrimination merits consideration. Such a proposal would clearly enhance the WTO's legitimacy but entail some reduction in the scope of international trade law. The concepts of the definitions of free trade help clarify what is at stake in choosing such a narrowing for the sake of legitimacy. Such a narrowing implies a particular agenda for the WTO. The analysis above suggests that a nondiscrimination concept would involve focusing WTO's institutional energy upon tariff reduction, elimination of regulations and taxes that expressly discriminate, elimination of quotas, and curbing domestic subsidies. This would leave the WTO with an ambitious agenda. On the other hand, it implies abandoning WTO efforts to decrease reliance upon nondiscriminatory coercion to achieve international policy goals and to weaken taxes and regulations. Analysts who wish to consider such a proposal should consider not just the desirability of the various objectives (or variants thereof), but also their suitability for pursuit by the WTO. The possible advantages of a sharper sense of institutional priorities and focus should form part of the analysis.

The focus on the definition of free trade invites critical thinking about GATT and SPS trade disciplines, not just the exceptions to the disciplines.[217] In particular, careful analysis of the role of coercion in regulation raises questions about whether a process standard should constitute a quantitative restriction of trade in violation of Article XI. This calls into question the United States' decision not to challenge Tuna/Dolphin I's conclusion that process standards constitute quantitative trade restrictions in Tuna/Dolphin II or in Shrimp/Turtle.[218] It also shows the inadequacy of a definition of free trade as trade free of "trade restrictions" or "trade barriers." Whether or not the reader embraces the proposal to focus on nondiscrimination, clearly the definition of free trade is important and helps explain why so many participants in the debate about trade and environment seem to talk past each other so often.

4 Conclusion

Hopefully, the identification of the three concepts of free trade will help spark a dialogue about the problem of defining free trade. The three-part conceptual scheme explains how the departure from a discriminatory concept threatens the WTO's legitimacy. The WTO could substantially reduce this threat by focusing on a credible antidiscrimination concept.

The WTO may eventually find that accepting some clear limits to the scope of its enterprise may be a price worth paying to enhance its legitimacy. The WTO's legitimacy will depend on whether it can relate its actions to a coherent concept of free trade that a specialized international institution can credibly administer.

Focusing inquiry upon the definition of free trade provides another way of thinking about "trade and . . ." issues. It invites more critical thinking about the trade disciplines themselves, rather than just the scope of exceptions. It also shows that

misunderstandings between free traders and environmentalists have roots in the conflation of concepts of free trade.

Notes

1 See, e.g., Daniel C. Esty, *Greening the GATT: Trade, Environment and the Future* (1994) [hereinafter Esty, *Greening the GATT*]; *Trade and the Environment: Law, Economics, and Policy* (Durwood Zaelke et al. eds., 1993) [hereinafter *Trade & Environment*]; C. Ford Runge, *Freer Trade, Protected Environment: Balancing Trade Liberalization and Environmental Interests* (1994).

2 See, e.g., Frank J. Garcia, *Protection of Intellectual Property Rights in the North American Free Trade Agreement: A Successful Case of Regional Trade Regulation*, 8 Am. U. J. Int'l L. & Pol'y 817 (1993); Special issue, *Trade-Related Aspects of Intellectual Property Rights*, 1 J. Int'l Econ. L. 497 (1998).

3 See, e.g., Virginia A. Leary, *Workers' Rights and International Trade: The Social Clause (GATT, ILO, NAFTA, US Laws)*, in *2 Fair Trade and Harmonization* 177 (Jagdish Bhagwati and Robert E. Hudec eds., 1996) [hereinafter Leary, *Workers' Rights and International Trade*].

4 See, e.g., James F. Smith, *NAFTA and Human Rights: A Necessary Linkage*, 27 U.C. Davis L. Rev. 793 (1994); Patricia Stirling, *The Use of Trade Sanctions as an Enforcement Mechanism for Basic Human Rights: A Proposal for Addition to the World Trade Organization*, 11 Am. U. J. Int'l L. & Pol'y 1 (1996).

5 See, e.g., Eleanor M. Fox, *Toward World Antitrust and Market Access*, 91 Am. J. Int'l L. 1 (1997).

6 Most of the materials cited in this article contain no definition of free trade. Tom Wathen defines free trade as "the unlimited exchange of commerce between buyers and sellers across national borders." Tom Wathen, *A Guide to Trade and Environment*, in *Trade & Environment*, *supra* note 1, at 5. He then states that "free trade . . . does not necessarily require the elimination of . . . regulations," because it aims only to avoid discrimination against "foreign companies." *Id.* Similarly, Steve Charnovitz rejects a laissez-faire definition of trade and suggests that it involves the absence of tariffs and "special border restrictions," both of which are usually discriminatory. See Steve Charnovitz, *Free Trade, Fair Trade, Green Trade: Defogging the Debate*, 27 Cornell Int'l L.J. 459, 471–2 (1994) [hereinafter Charnovitz, *Defogging the Debate*]. This article addresses the tension implicit in free trade under a laissez-faire definition and a more limited definition focusing upon nondiscrimination. See *id.*

7 See, e.g., Esty, *Greening the GATT*, *supra* note 1, at 3, 23–46 (discussing but not defining free trade); Leary, *Workers' Rights and International Trade*, *supra* note 3 (same). *Cf.* Brian Alexander Langille, *General Reflections on the Relationship of Trade and Labor (Or Fair Trade is Free Trade's Destiny)*, in *2 Fair Trade and Harmonization*, *supra* note 3, at 236 (questioning the "assumption that . . . there is a natural or genuinely noncontroversial mode of economic ordering" which trade theory can police); Donald H. Regan, *The Supreme Court and State Protectionism: Making Sense of the Dormant Commerce Clause*, 84 Mich. L. Rev. 1091, 1097 (1986) (explaining that the ambiguity of the phrase "free trade" encourages conflating opposition to protectionism with a commitment to laissez-faire).

8 See General Agreement on Tariffs and Trade, October 30, 1947, 61 Stat. A-11, TIAS No. 1700, 55 UNTS 194 [hereinafter GATT].

9 See Jagdish Bhagwati, *Introduction* to *1 Fair Trade and Harmonization*, *supra* note 3, at 1 (referring to the "simple elimination of trade barriers").

10 See, e.g., Daniel A. Farber and Robert E. Hudec, *GATT Legal Restraints on Domestic Environmental Regulations*, in *2 Fair Trade and Harmonization*, *supra* note 3, at 59, 64 [hereinafter Farber & Hudec, *GATT Legal Restraints*].

11 See, e.g., *id*. at 75 (noting that it would have been easy to find US corporate average fuel economy standards "protectionist"); Ved P. Nanda, *International Environmental Law & Policy* 44 (1995) ["The controversy surrounding the use of ETM's (environmental trade measures) centers on the possibility of a state imposing an ETM as a protectionist measure."]. *Cf.* Alan O. Sykes, *Regulatory Protectionism and the Law of International Trade*, 66 U. Chi. L. Rev. 1, 3–5 (1999) [hereinafter Sykes, *Regulatory Protectionism*] (providing a definition for this otherwise vague term, "protectionism").

12 But see Regan, *supra* note 7, at 1094–5 (defining protectionism and not using it as a synonym for "trade restrictions" or "trade barriers"); Sykes, *Regulatory Protectionism*, *supra* note 11, at 3–7 (clarifying "protectionism").

13 Marakesh Agreement Establishing the World Trade Organization, April 15, 1994, art. II, para. 6 reprinted in *GATT Secretariat, The Results of the Uruguay Round of Multilateral Trade Negotiations: The Legal Texts* 6–7 (1994) [hereinafter WTO Agreement]. The relevant agreements are set out in Annexes to the agreement establishing the WTO. See *id*. Annexes 1–4 at 19–439.

14 On the concept of legitimacy see Thomas M. Franck, *The Power of Legitimacy Among Nations* (1990) and Martti Kokenniemi, Book Review, 86 Am. J. Int'l L. 175 (1992) (reviewing Franck, *supra*).

15 See Kenneth W. Abbott, *Economic Issues and Political Participation: The Evolving Boundaries of International Federalism*, 18 Cardozo L. Rev. 971, 975 (1996) (pointing out risk of trade regimes losing legitimacy and public support unless noneconomic interests have a role in WTO policy and dispute settlement); Jeffrey L. Dunoff, *Resolving Trade–Environment Conflicts: The Case for Trading Institutions*, 27 Cornell Int'l L.J. 607, 621–2 (1994) [hereinafter Dunoff, *Trading Institutions*] (arguing against allowing GATT to settle trade and environment disputes); Jeffrey L. Dunoff, *Rethinking International Trade*, 19 U. Pa. J. Int'l Econ. L. 347 (1998) [hereinafter Dunoff, *Rethinking*] (arguing that "linkage issues" such as "trade and environment" raise serious practical and theoretical challenges to the trade regime); Jeffrey L. Dunoff, *"Trade and": Recent Developments in Trade Policy and Scholarship – And Their Surprising Political Implications*, 17 Nw. J. Int'l L. & Bus. 759, 764–8 (1996) [hereinafter Dunoff, "*Trade and"*] (arguing that expansion of trade law into "new substantive areas threatens to undermine international and domestic support for the trade regime").

16 See *A Global Disaster*, Economist, December 11, 1999, at 19–20; *NGOs From 60 Countries Team Up to Halt Next WTO Round on Environmental Grounds*, 22 Int'l Env't. Rep. (BNA) 446 (May 26, 1999).

17 See *GATT Dispute Settlement Panel Report on US – Restrictions on Imports of Tuna*, July 1994, 33 I.L.M. 839, 889–90, 894, 898–9 (1994) [hereinafter Tuna/Dolphin II]; *GATT Dispute Settlement Panel Report on US – Restrictions on Imports of Tuna*, September 3, 1991, GATT B.I.S.D. (39th Supp.) at 155, 195, 200–1, 205 (1993) [hereinafter Tuna/Dolphin I].

18 Agreement on the Application of Sanitary and Phytosanitary Measures, April 15, 1994, WTO Agreement, Annex 1A, *supra* note 13 [hereinafter SPS Agreement].

19 See *WTO Appellate Body Report on US – Import Prohibition of Certain Shrimp and Shrimp Products*, October 12, 1998, 38 I.L.M. 118 (1999) [hereinafter Shrimp/Turtle Appellate]; *WTO Appellate Body Report on EC – Measures Concerning Meat and Meat Products (Hormones)*, 1998 WL 25520 (January 16, 1998) [hereinafter Beef/Hormone Appellate].

20 See Annick Emmenegger Brunner, *Conflicts Between International Trade and Multilateral Environmental Agreements*, 4 Ann. Surv. Int'l & Comp. L. 74, 78 (1997) (explaining that the issue of environmental trade measures was "raised most prominently" by the Tuna/Dolphin decisions in 1991 and 1994); Steve Charnovitz, *Green Roots, Bad Pruning: GATT Rules and Their Application to Environmental Trade Measures*, 7 Tul. Envtl. L.J. 299, 301 (1994) [hereinafter Charnovitz, *Pruning*] ("The infamous dolphin–tuna dispute of the early 1990s crystallized for the public the implications for the environment of trade agreements."); Frank J. Garcia, *The Trade Linkage Phenomenon: Pointing the Way to the Trade Law and Global Social Policy of the 21st Century*, 19 U. Pa. J. Int'l Econ. L. 201, 202 n.3 (1998) (crediting GATT Tuna/Dolphin decisions with spurring public opposition to the WTO and concern with linkage issues); Thomas J. Schoenbaum, *International Trade and Protection of the Environment: The Continuing Search for Reconciliation*, 91 Am. J. Int'l L. 268 (1997) ("The Tuna/Dolphin I decision produced an explosion of rhetoric in . . . learned journals."). See, e.g., Esty, *Greening the GATT*, *supra* note 1; *Trade & Environment*, *supra* note 1; Howard F. Chang, *An Economic Analysis of Trade Measures to Protect the Global Environment*, 83 Geo. L.J. 2131 (1995) [hereinafter Chang, *An Economic Analysis of Trade Measures*]; Charnovitz, *Defogging the Debate*, *supra* note 6; Edith Brown Weiss, *Environment and Trade as Partners in Sustainable Development: A Commentary*, 86 Am. J. Int'l L. 728 (1992); Chris Wold, *Multilateral Environmental Agreements and the GATT: Conflict and Resolution*, 26 Envtl. L. 841 (1996) [hereinafter Wold, *Multilateral Environmental Agreements*]; Mark Edward Foster, Note, *Trade and Environment: Making Room for Environmental Trade Measures Within the GATT*, 71 S. Cal. L. Rev. 393 (1998) (arguing that the WTO should consider environmental trade measures legal under GATT Article XX).

21 See *WTO Committee Report Claims Success in Furthering Trade, Environment Talks*, 21 Int'l Env't. Rep. (BNA) 1127, 1128 (November 11, 1998) [hereinafter UNEP Criticism] (citing United Nations Environment Program's charge that uncertainty in international trade law "may have hampered several environmental conventions."). But see Daniel A. Farber and Robert E. Hudec, *Free Trade and the Regulatory State: A GATT's-Eye View of the Dormant Commerce Clause*, 47 Vand. L. Rev. 1401, 1404 (1994) [hereinafter Farber & Hudec, *Free Trade and the Regulatory State*] (arguing that "lack of clarity may be a necessary characteristic" of law addressing free trade).

22 See GATT art. III, *supra* note 8; see generally Dunoff, *Rethinking International Trade*, *supra* note 15, at 370–3; Farber & Hudec, *Free Trade and the Regulatory State*, *supra* note 21, at 1406 (characterizing free trade policy as a principle of evenhandedness).

23 I use the term "coercion" with some reservations. Trade measures, when enforced, do not in fact force foreign nations to change their conduct as a military invasion or the arrest of a person might. See Belina Anderson, *Unilateral Trade Measures and Environmental Protection Policy*, 66 Temple L. Rev. 751, 755 (1993).

24 See generally John O. McGinnis, *The Decline of the Western Nation State and the Rise of the Regime of International Federalism*, 18 Cardozo L. Rev. 903, 916 (1996) (arguing that the competition for trade resulting from GATT imposes limitations on national regulatory power similar to those "imposed on the power of states in nineteenth-century America.").

25 See generally Abbott, *supra* note 15, at 975 [citing G. Richard Shell, *Trade Legalism and International Relations Theory: An Analysis of the World Trade Organization*, 44 Duke L.J. 829, 908–9 (1995)] (pointing out risk of trade regimes losing legitimacy and public support unless noneconomic interests have a role in WTO policy and dispute settlement); Dunoff, "*Trade and*," *supra* note 15, at 764 ("The expansion of trade law into new substantive areas threatens to undermine international and domestic support for the trade regime."); Daniel C. Esty, *Linkages and Governance: NGOs at the World Trade Organization*,

19 U. Pa. J. Int'l Econ. L. 709, 715 (1998) [hereinafter Esty, *Linkages and Governance*] (questioning WTO legitimacy to make regulatory judgments); McGinnis, *supra* note 24, at 918–24 (explaining that the regime of "international federalism" may be unstable because "interest groups" may oppose laissez-faire policies).

26 See W. M. Corden, *Trade Policy and Economic Welfare* 2–5 (1989).

27 See David Ricardo, *On the Principles of Political Economy and Taxation*, in *The Works and Correspondence of David Ricardo* (1962); Adam Smith, *The Wealth of Nations* (Modern Library 1994) (1776).

28 See, e.g., Douglas A. Irwin, *Against the Tide: An Intellectual History of Free Trade* (1996) (describing a history of debates about free trade's merits, rather than its definition); Paul R. Krugman, *Rethinking International Trade* (1990) (theorizing about the origins of international trade, the effects of "protectionism," and optimal trade policy and building in insights about economies of scale).

29 See Smith, *supra* note 27, at 485–6.

30 See Ricardo, *supra* note 27, at 128–49.

31 See Peter H. Lindert, *International Economics* 28 (8th edn. 1986).

32 For a brief review of the theory as it has evolved to date see Alan O. Sykes, *Comparative Advantage and the Normative Economics of International Trade Policy*, 1 J. Int'l Econ. L. 49 (1998) [hereinafter Sykes, *Comparative Advantage*].

33 See Smith, *supra* note 27, at 456–717.

34 See *id*. at 479.

35 See *id*. at 481–502.

36 See, e.g., *Economic Justice in Perspective: A Book of Readings* 69–70 (Jerry Combee and Edgar Norton eds., 1991); Daniel R. Fusfeld, *The Age of the Economist* 20–54 (5th edn. 1986); Robert L. Heilbroner, *The Worldly Philosophers: The Lives, Times, and Ideas of the Great Economic Thinkers* (4th edn. 1972).

37 See Smith, *supra* note 27, at 494 (describing compensatory taxation as "the second limitation of the freedom of trade"); see also Lindert, *supra* note 31, at 17 (noting that Smith accepted "the national defense argument for restricting trade with potential enemies.").

38 See Ricardo, *supra* note 27, at 133. *Cf. id*. at 318, 338 (referring to "universally free trade" and "free trade" respectively).

39 See Smith, *supra* note 27, at 494–6, 790, 876.

40 Ricardo, *supra* note 27, at 167.

41 *Id*. at 172.

42 *Id*.

43 See Mordechai E. Kreinin, *International Economics: A Policy Approach* 224 (5th ed. 1997) (beginning a chapter on "Why Nations Trade" with a discussion of comparative advantage); Paul R. Krugman and Maurice Obstfeld, *International Economics: Theory and Practice* 11 (2nd ed. 1991) (arguing that comparative advantages helps "us to understand how differences between countries give rise to trade . . ., and why this trade is mutually beneficial."); Lindert, *supra* note 31, at 17–24 (describing Ricardo and Smith's contribution as an explanation of how the gains of trade will accrue to both trading partners); Sykes, *Comparative Advantage*, *supra* note 32, at 49–50.

44 See Sykes, *Comparative Advantage*, *supra* note 32, at 49–50.

45 This explanation, however, yields a prediction that "nations will tend to specialize in production of goods in which they have comparative advantage, exporting them to other nations in exchange for goods in which they lack comparative advantage." See *id*. at 52–3. This remains consistent with the role of a descriptive theory that explains a phenomena, rather than justifies or defines a particular social policy.

46 See Krugman & Obstfeld, *supra* note 43, at 11 (describing comparative advantage as the "essential concept" to understanding "why trade is mutually beneficial").

47 See Sykes, *Comparative Advantage*, *supra* note 32, at 57 (explaining that the normative case for free trade rests upon economic efficiency).

48 See, e.g., Avinash Dixit and Victor Norman, *Theory of International Trade: A Dual Equilibrium Approach* 1 (1980) (explaining that debates about the impacts of tariffs require a general equilibrium model making full use of the knowledge offered by microeconomic theory); Krugman & Obstfeld, *supra* note 43, at 181–93.

49 See Sykes, *Comparative Advantage*, *supra* note 32, at 57 (offering background on welfare economics because normative case for free trade rests on efficiency grounds).

50 See E. J. Mishan, *Cost–Benefit Analysis: An Informal Introduction* 154–61 (3rd ed. 1982); David M. Driesen, *The Societal Cost of Environmental Regulation: Beyond Administrative Cost–Benefit Analysis*, 24 Ecology L.Q. 545, 577–9 (1997) [hereinafter Driesen, *The Societal Cost*].

51 See, e.g., Thomas O. McGarity, *Reinventing Rationality: The Role of Regulatory Analysis in the Federal Bureaucracy* (1991); Mark Sagoff, *The Economy of the Earth: Philosophy, Law, and the Environment* (1988) (arguing that environmental policies should be based on ethical, esthetic, cultural, and historical considerations rather than aggregate personnel preferences); David W. Pearce and R. Kerry Turner, *Economics of Natural Resources and the Environment* 122–3 (1990) (noting temptation to "downgrade" environmental benefits because they are "soft" variables); Driesen, *The Societal Cost*, *supra* note 50 (arguing that administrative cost–benefit analysis does not help evaluate most important economic questions about cost or accomplish any of the other goals proponents have outlined); Duncan Kennedy, *Cost–Benefit Analysis of Entitlement Problems: A Critique*, 33 Stan. L. Rev. 387 (1980) (arguing that outcomes of cost–benefit analysis are indeterminate in theory); Laurence H. Tribe, *Policy Science: Analysis or Ideology?*, 2 Phil. & Pub. Aff. 66, 70 (1972) (arguing that pareto-optimality analysis is both indeterminate and incomplete).

52 On its status under US domestic law see Uruguay Round Agreements Act, Pub. L. No. 103-465, 108 Stat. 4809 (1994) (codified as 19 U.S.C. §§ 3501–12) (approving GATT and the Uruguay Round agreement but providing that these agreements do not amend federal laws protecting workers, the environment, and health and safety). See also John H. Jackson, *The General Agreement on Tariffs and Trade in United States Domestic Law*, 66 Mich. L. Rev. 249 (1967); C. O'Neal Taylor, *The Limits of Economic Power: Section 301 and the World Trade Organization Dispute Settlement System*, 30 Vand. J. Transnat'l L. 209, 243 n.160 (1997).

53 See http://www.wto.org.

54 See GATT, art. III, *supra* note 8; see generally John H. Jackson, *World Trade Rules and Environmental Policies: Congruence or Conflict?*, 49 Wash. & Lee L. Rev. 1227, 1236–7 (1992) [hereinafter Jackson, *Congruence or Conflict?*].

55 See GATT, art. III, *supra* note 8; Robert E. Hudec, *GATT Legal Restraints on the Use of Trade Measures Against Foreign Environmental Practices*, in *2 Fair Trade and Harmonization*, *supra* note 3, at 108 [hereinafter Hudec, *Trade Measures Against Foreign Environmental Practices*] ("The GATT's economic goal is to promote, through liberal international trade policies, the greater effectiveness of national economies . . .. The GATT regime has made a major contribution to alleviating poverty in the postwar era.").

56 See GATT preamble, *supra* note 8 [advocating the "*elimination* of discriminatory treatment" (emphasis added), but only the "substantial reduction" of trade barriers generally].

57 See *id*. (citing substantial reduction of tariffs as a goal of the parties who signed GATT); *id*. art. XXVIII (providing for negotiation of tariff reductions). Under the auspices of the WTO, the parties to GATT periodically negotiate schedules that reduce tariffs, which constitute an integral part of the GATT agreement.

58 See *id*. arts. III, XX.

59 See *id*. art. I.

60 I posit this simple model as a means of framing discussion of free trade's meaning. This model does not capture every feature of GATT.

61 See generally *The Uruguay Round and Beyond: Essays in Honour of Arthur Dunkel* 255 (Jagdish Bhagwati and Mathias Hirsch eds., 1998) (advocating elimination of border restrictions and national subsidies for domestic coal production to liberalize further WTO member economies).

62 *Cf.* Regan, *supra* note 7, at 1096–7 (distinguishing between being against protectionism and being in favor of total economic laissez-faire).

63 See GATT art. XI, para. 1, *supra* note 8.

64 See *id*. art. XI.

65 See *WTO Dispute Settlement Panel Report on India – Quantitative Restrictions on Imports of Agricultural, Textile, and Industrial Products*, 1999 WTO DS Lexis 5, para. 5.142 (April 6, 1999) [hereinafter India Panel] (construing Article XI to include "a limitation on action, a limiting condition or regulation.").

66 See, e.g., Robert E. Hudec, *The Product–Process Doctrine in GATT/WTO Jurisprudence*, in *New Directions in International Economic Law: Essays in Honour of John H. Jackson* 187, 191–2 (Marco Bronckers and Reinhard Quick eds., 2000) [hereinafter Hudec, *GATT/WTO Jurisprudence*].

67 See *id.*

68 See GATT art. XX, *supra* note 8.

69 See Tuna/Dolphin I, *supra* note 17, para. 5.27, at 199.

70 See GATT art. XX, para. 1, (b), (g), *supra* note 8; Shrimp/Turtle Appellate, *supra* note 19, paras 127–34, at 154–7 (applying to efforts to protect endangered species); Steve Charnovitz, *Exploring the Environmental Exceptions in GATT Article XX*, J. World Trade, October 1991, at 37 (1991) [hereinafter Charnovitz, *Environmental Exceptions*].

71 GATT dispute settlement panels have repeatedly applied a least restrictive means test under Article XX. See, e.g., *GATT Dispute Settlement Panel Report on Canada – Measures Affecting Exports of Unprocessed Herring and Salmon*, March 22, 1988, GATT B.I.S.D. (35th Supp.) at 114 (1989); *GATT Dispute Settlement Panel Report on Thailand – Restrictions on Importation of and Internal Taxes on Cigarettes*, November 7, 1990, GATT B.I.S.D. (37th Supp.) at 223 (1991); Tuna/Dolphin II, *supra* note 17, para. 5.35, at 897; Tuna/Dolphin I, *supra* note 17, paras 5.28, 5.33, at 199–201; see also Charnovitz, *Environmental Exceptions*, *supra* note 70, at 49–50 (showing why this test is difficult to meet); Steve Charnovitz, *Environmental Trade Sanctions and the GATT: An Analysis of the Pelly Amendments on Foreign Environmental Practices*, 9 Am. U. J. Int'l L. & Pol'y 778–9 (1994) [hereinafter Charnovitz, *Pelly Amendments*] (contrasting a literal approach to Article XX's chapeau with the case law's approach); Farber & Hudec, *GATT Legal Restraints*, *supra* note 10, at 81 (recognizing that "it is always possible to imagine *some* less restrictive alternative," but arguing that GATT tribunals have in practice exercised "good judgment and common sense in this exercise").

72 See *WTO Dispute Settlement Panel Report on EC – Measures Affecting Asbestos and Asbestos-Containing Products*, 2000 WL 1449942 (September 18, 2000) [hereinafter Asbestos] (upholding European asbestos ban under Article XX); Shrimp/Turtle Appellate,

supra note 19, paras 161–84, 187, at 166–75; *WTO Appellate Body Report on US – Standards for Reformulated and Conventional Gasoline*, May 20, 1996, 35 I.L.M. 603, 618–23 (1996) [hereinafter *Reformulated Gasoline*] (eschewing a least restrictive means type approach, but still finding law illegal under Article XX); Charnovitz, *Defogging the Debate*, *supra* note 6, at 468–70; Robert E. Hudec, *GATT/WTO Constraints on National Regulation: Requiem for an "Aim and Effects" Test*, 32 Int'l Law. 619, 622 (1998) ("[T]he interpretation of Article XX has made its requirements exceptionally demanding.").

73 See *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R; (01-1157); AB 2000-11, ¶ 192(e) (March 12, 2001).

74 See Shrimp/Turtle Appellate, *supra* note 19, paras 120–1, at 152–3. The Appellate Body held erroneous the panel's conclusion that conditioning access to a market upon the exporter's compliance with the importer's unilaterally required policies necessarily violated Article XX. In so doing, it stated that a *per se* rule against unilateral requirements might systematically make much of Article XX "inutile." See *id*. para. 21, at 152–3.

75 See SPS Agreement, *supra* note 18. For background and a preliminary assessment of the SPS Agreement, see Donna Roberts, *Preliminary Assessment of the Effects of the WTO Agreement on Sanitary and Phytosanitary Trade Regulations*, 1 J. Int'l Econ. L. 377 (1998).

76 See *WTO Dispute Settlement Panel Report on EC – Measures Concerning Meat And Meat Products (Hormones)*, 1997 WL 569984 (August 18, 1997) [hereinafter *Beef/Hormone Panel*]. This decision established that the WTO would declare illegal "sanitary" measures that did not discriminate. The panel opinion suggests that the WTO may find a measure to be in violation of the SPS without previously finding the measure to be in violation of GATT's nondiscrimination principles. See *id*. paras 8.36–8.41.

77 See Vern R. Walker, *Keeping the WTO from Becoming the "World Trans-science Organization": Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute*, 31 Cornell Int'l L.J. 251, 301–3, 308–9 (1998) (discussing Panel fact-finding and the Appellate Body's approval of the findings).

78 See Beef/Hormone Appellate, *supra* note 19, para. 253.

79 See Steve Charnovitz, *The World Trade Organization, Meat Hormones, and Food Safety*, 14 Int'l Trade Rep. (BNA) No. 41, at 1781 (October 15, 1997) [hereinafter Charnovitz, *The World Trade Organization*]. *Cf.* Ryan David Thomas, Note, *Where's the Beef? Mad Cows and the Blight of the SPS Agreement*, 32 Vand. J. Transnat'l L. 487 (1999) (arguing that the SPS agreement does not go far enough toward harmonizing regulatory measures).

80 See Beef/Hormone Appellate, *supra* note 19, para. 186. *Cf.* David A. Wirth, *International Trade Agreements: Vehicles for Regulatory Reform*, 1997 U. Chi. Legal F. 331, 339 n.22 [hereinafter Wirth, *International Trade Agreements*] (questioning the Beef/Hormone Panel's approach).

81 See Beef/Hormone Appellate, *supra* note 19, para. 193.

82 See *id.* para. 194.

83 See *id.* para. 198.

84 See *id.* paras 196, 199–200. See Walker, *supra* note 77, at 303 (faulting the Appellate Body for requiring an assessment "so 'specifically focused' that the risk determination itself must clear a high threshold of specificity").

85 See Beef/Hormone Appellate, *supra* note 19, para. 207.

86 See *id.* para. 187.

87 See, e.g., *WTO Appellate Body Report on Australia – Measures Affecting Importation of Salmon*, 1998 WL 731009, para. 124 (October 20, 1998) (reversing Panel assumption that

a document containing only "some evaluation" of risk is a risk assessment under the SPS agreement).

88 See Thomas O. McGarity, *Substantive and Procedural Discretion in the Administrative Resolution of Science Policy Questions: Regulating Carcinogens in EPA and OSHA*, 67 Geo. L.J. 729, 780 (1979) [hereinafter McGarity, *Regulating Carcinogens in EPA and OSHA*] (arguing that no governmental regulatory program is possible if it must be based solely upon accepted "facts"); Wirth, *International Trade Agreements*, *supra* note 80, at 343 (discussing the chilling effect judicial application of science-based test may have upon regulation).

89 See McGarity, *supra* note 88, at 733–6 (explaining why direct data about the effects of widespread exposure to low doses of carcinogens is rarely available); Walker, *supra* note 77, at 258–61 (describing nature of some of the uncertainties in risk assessment); Wirth, *International Trade Agreements*, *supra* note 80, at 340 (explaining that quantitative risk assessment involves the application of policy choices to uncertain data); David A. Wirth, *The Role of Science in the Uruguay Round and NAFTA Trade Disciplines*, 27 Cornell Int'l L.J. 817, 837–40 (1994) [hereinafter Wirth, *The Role of Science*] (discussing uncertainty and the precautionary principle).

90 See Beef/Hormone Appellate, *supra* note 19, para. 97 (designating burden of proof as "of particular importance").

91 See Walker, *supra* note 77, at 313 (arguing that "the party with the burden of persuasion has a difficult burden of proof" because of scientific uncertainty).

92 See Beef/Hormone Appellate, *supra* note 19, paras 99–108.

93 See *id.* para. 98.

94 The Appellate Body stated that its finding that the burden of proof shifts upon finding a *prima facie* violation "does not deal with the quite separate issue of whether the United States and Canada actually made a *prima facie* case . . . ." See *id*. para. 109 n.71. The Appellate Body concluded that the United States and Canada established a *prima facie* case but did not explain why. See *id*. para. 197 n.180. See Walker, *supra* note 77, at 317 (stating that it is not clear what generalizable threshold for a *prima facie* case emerges from the Appellate Body's decision). The Appellate Body effectively treated a range of scientific disputes as establishing a *prima facie* case. See *id*.

95 A subsequent decision supports finding that a measure is not based on a risk assessment when the complaining nation has not made a *prima facie* case. See *WTO Appellate Body Report on Japan – Measures Affecting Agricultural Products*, 1999 WL 83966, paras 109–14, 136–7 (February 22, 1999) [hereinafter *Japan – Measures Affecting Agricultural Products*]. In that case, the Appellate Body held that Japan had violated its obligation to base its measure on a risk assessment, even though the United States had made no *prima facie* case that no scientific evidence supported the requirement. See *id.* At the same time, this case provides some more guidance on what constitutes a *prima facie* case of a violation of the SPS requirement that nations base their measures on scientific evidence in article 2.2. See *id.* para. 137. A finding of a *prima facie* case of a violation of article 5.1's risk assessment provision alone constitutes a *prima facie* case of a violation of the SPS agreement. Hence, if a *prima facie* case can be easily made for violations of article 5, then a *prima facie* case of a violation of the SPS agreement will arise frequently, regardless of how other articles are treated.

96 See Walker, *supra* note 77, at 318 ("It is possible therefore that the EC has the difficult task of persuading the WTO that its evidence was sufficiently specific and probative so as to overcome a presumption of no risk.").

97 See *id.* (pointing out the "substantial burden" the ambiguous standards for evidence imposed on the EC); Wirth, *The Role of Science*, *supra* note 89, at 833 ("[S]cience can inform the regulatory process but cannot, by itself, determine the results with particularity."). *Cf.* Roberts, *supra* note 75, at 396–7 (discussing evidence that enunciation of SPS principles has prompted unilateral reconsideration of health protection measures). Roberts, however, believes that fears that SPS disciplines would occasion an "intolerable assault on . . . food safety and environmental standards have likely been overdrawn." *Id.* at 399. To support this view, she cites the fact that most of the disputes to date involve developed countries, but she does not explain why SPS disciplines might not undermine food safety and environmental standards in developed countries. See *id.* at 398–9.

Prior to the Beef/Hormone decision a number of scholars believed that the SPS agreement did not interfere with a member's freedom to set standards. See, e.g., Schoenbaum, *supra* note 20, at 285–6. Obviously, the Beef/Hormone decision requires reevaluation of these views.

98 See SPS Agreement, *supra* note 18, art. 5.6.

99 See *infra* pp. 177–8. *Cf. Japan – Measures Affecting Agricultural Products*, *supra* note 95, paras 96–100, 126 (holding that complaining party must show that an alternative measure meets the level of stringency desired by the regulating government to invoke the last trade restrictive means test).

100 See Steve Charnovitz, *Environment and Health Under WTO Dispute Settlement*, 32 Int'l Law. 901 (1998) (reviewing the substantive impact of WTO cases).

101 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 113 ("Whenever simple diplomacy fails . . . trade restrictions seem to offer just the right blend of coercion and civility."); see, e.g., Amy Blackwell, *The Humane Society and Italian Driftnetters: Environmental Activists and Unilateral Action in International Environmental Law*, 23 N.C. J. Int'l L. & Com. Reg. 313, 314–15 (1998) (explaining that the High Seas Driftnet Fisheries Enforcement Act prohibits imports of fish from nations using driftnets because several countries have not complied with UN resolutions calling for a moratorium on driftnetting).

102 See generally Taylor, *supra* note 52.

103 Article III provides for national treatment of the "products of the territory of any *contracting party*" (emphasis added) upon import into the territory of another contracting party. See GATT art. III, *supra* note 8. Similarly, Article XI bars a contracting party from implementing quantitative restrictions on imports and exports from a contracting party. See *id.* art. XI. Hence, these disciplines apply only to those who have contracted to abide by GATT.

104 See GATT arts. III, XI, *supra* note 8.

105 See John H. Jackson, *World Trade and the Law of GATT* 258 (1969) (recognizing the existence of an incentive to enter GATT in order to receive "the advantages of GATT.").

106 See Steve Charnovitz, *The Moral Exception in Trade Policy*, 38 Va. J. Int'l L. 689, 719 (1998); Taylor, *supra* note 52, at 245–6.

107 See Robert E. Hudec, *The New WTO Dispute Settlement Procedure: An Overview of the First Three Years*, 8 Minn. J. Global Trade 1, 9 (1999) [hereinafter Hudec, *The New WTO Dispute Settlement Procedure*] (discussing defendant's veto rights under GATT dispute settlement practice).

108 See Trade Act of 1974, Pub. L. No. 93-618, §§ 301–2, 88 Stat. 1978 (1975) [codified as amended at 19 U.S.C. §§ 2411–20 (1994)].

109 See Kenneth W. Abbott, *Defensive Unfairness: The Normative Structure of Section 301*, in *2 Fair Trade and Harmonization*, *supra* note 3, at 420–2; Hudec, *The New WTO Dispute Settlement Procedure*, *supra* note 107, at 13; Taylor, *supra* note 52, at 228.

110 See Hudec, *The New WTO Dispute Settlement Procedure*, *supra* note 107, at 13.

111 See *Understanding on Rules and Procedures Governing the Settlement of Disputes*, April 15, 1994, Final Act Embodying the Results of the Uruguay Round of Multinational Trade Negotiations, Annex 2, arts. 16(4), 17(14), reprinted in *The Results of the Uruguay Round of Multilateral Trade Negotiations: The Legal Texts* 417, 419 (1994) [hereinafter *DSU Understanding*].

112 See *id.* art. 22(1).

113 See Taylor, *supra* note 52, at 259 (discussing "WTO goal of coercing the offending country into compliance with its GATT obligations"). *Cf. Montreal Protocol on Substances that Deplete the Ozone Layer*, September 16, 1987, 26 I.L.M. 1550 (entered into force January 1, 1989) [hereinafter *Montreal Protocol*]; Richard Benedick, *Ozone Diplomacy* (enlarged ed. 1998) (describing how diplomacy and information have promoted treaty development); Elizabeth P. Barratt-Brown, *Building a Monitoring and Compliance Regime Under the Montreal Protocol*, 16 Yale J. Int'l L. 519, 534–7 (1991) (discussing incentive structure and its effects).

114 See Eyal Benvenisti, *Exit and Voice in the Age of Globalization*, 98 Mich. L. Rev. 167, 168, 201 (1999).

115 See David M. Driesen, *The Congressional Role in International Environmental Law and its Implications for Statutory Interpretation*, 19 B.C. Envtl Aff. L. Rev. 287, 303–8 (1991) [hereinafter Driesen, *The Congressional Role*].

116 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 149 (arguing that environmental trade measures are usually directed at genuine environmental concerns).

117 See Charnovitz, *Defogging the Debate*, *supra* note 6, at 493; see, e.g., Driesen, *The Congressional Role*, *supra* note 115, at 303–4 (discussing example of unilateral restriction helping to create strengthened international standards for oil pollution from ships).

118 See Montreal Protocol, *supra* note 113, at art. 4; Benedick, *supra* note 113, at 91–2 (diplomatic history and nature of provisions).

119 See Barratt-Brown, *supra* note 113, at 534–7. CFC production in the developing countries has not grown at rates high enough to offset the cuts in developed countries. *Cf. Montreal Treaty Seen as Major Success in Effort to Protect Stratospheric Ozone Layer*, 28 Env't Rep. (BNA) 778 (August 29, 1997) [hereinafter *Montreal Treaty Seen as Success*]. Yet, some contraband exports continue to bedevil implementation of the protocol. See *Environmental Investigating Group Finds Widespread Trade of CFCs in Europe*, 20 Int'l Env't. Rep. (BNA) 869 (September 17, 1997).

120 See Barratt-Brown, *supra* note 113, at 534 (discussing the need to discourage "the huge potential increase" in production and consumption of ozone-depleting chemicals in developing countries); Jeffrey L. Dunoff, *Reconciling International Trade with Preservation of the Global Commons: Can We Prosper and Protect*, 49 Wash. & Lee L. Rev. 1407, 1429 (1992) ("A one percent decline in ozone level may . . . cause 200,000 or more skin cancer cases"' and cause serious damage to people's eyes). Although the ozone layer has suffered substantial depletion, it may recover by the middle of the next century. See *Montreal Treaty Seen as Success*, *supra* note 119, at 20.

121 See Edith Brown Weiss, *Understanding Compliance with International Environmental Agreements: The Baker's Dozen Myths*, 32 U. Rich. L. Rev. 1555, 1555 (1999) (pointing out that more than 1,000 international legal instruments have provisions addressing environmental protection).

122 See generally David A. Wirth, *The International Trade Regime and the Municipal Law of Federal States: How Close a Fit?*, 49 Wash. & Lee L. Rev. 1373, 1391 (1992)

(stating that the current international environmental system invites "holdouts, free riders, laggards, scofflaws, and defectors").

123 See Montreal Protocol, *supra* note 113.

124 See Benedick, *supra* note 113, at 91.

125 See, e.g., Wold, *Multilateral Environmental Agreements*, *supra* note 20, at 880–6 (analyzing GATT legality of the Basel Convention, which limits shipment of hazardous waste).

126 See Charnovitz, *Pelly Amendment*, *supra* note 71, at 757. Professor Daniel Esty refers to unilateral imposition of sanctions to advance compliance with internationally agreed upon standards as a form of "multilateral unilateralism." See Esty, *Greening the GATT*, *supra* note 1, at 140. My discussion relies implicitly on Professor Esty's very useful framework. See also Steve Charnovitz, *A Taxonomy of Environmental Trade Measures*, 6 Geo. Int'l Envtl L. Rev. 1 (1993) [hereinafter Charnovitz, *Taxonomy*] (presenting a more detailed taxonomy).

127 See Driesen, *The Congressional Role*, *supra* note 115, at 305; *cf.* Charnovitz, *Pelly Amendments*, *supra* note 71, at 756–7.

128 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 104–5 (describing various European trade restrictions).

129 See *Convention on Trade in Endangered Species of Wild Fauna and Flora (CITES)*, opened for signature March 3, 1973, art. II, para. 4, 27 UST 1087, 993 UNTS 243; see, e.g., *Defenders of Wildlife v. Endangered Species Scientific Authority*, 659 F.2d 168 (DC Cir. 1981), cert. denied sub nom., *International Assoc. of Fish & Wildlife Agencies v. Defenders of Wildlife*, 454 US 963 (1981) (adjudicating validity of bobcat export decisions under Endangered Species Act provisions implementing CITES).

130 See Wold, *Multilateral Environmental Agreements*, *supra* note 20, at 870–4.

131 See, e.g., Chris Wold, *The Futility, Utility, and Future of the Biodiversity Convention*, 9 Colo. J. Int'l Envtl. L. & Pol'y 1, 1 (1998) ("Biodiversity Convention has accomplished little of substance."). Of course, the lack or presence of trade sanctions does not, by itself, explain whether a treaty will be successful. See *id.* at 4–22 (explaining reasons for biodiversity convention's failure). Nevertheless, the conspicuous coincidence of trade sanctions and success should caution one against too quickly concluding that trade sanctions are useless.

132 See Tuna/Dolphin I, *supra* note 17, para. 2.7, at 155 (noting that Panama and Ecuador prohibited setting purse seine nets on dolphins after US imposition of trade restrictions); Laurel L. Hyde, Comment, *Dolphin Conservation in the Tuna Industry: The United States' Role in an International Problem*, 16 San Diego L. Rev. 665, 691–2 (1979) (suggesting that unilateral measures brought Senegal, Congo, Spain, and New Zealand into compliance with US conservation practices).

133 See Tuna/Dolphin II, *supra* note 17, paras 6.1–6.2, at 899; Tuna/Dolphin I, *supra* note 17, para. 7.1, at 205.

134 See Charnovitz, *The Moral Exception in Trade Policy*, *supra* note 106, at 719.

135 See, e.g., *WTO Dispute Settlement Panel Report on US – Import Prohibition of Certain Shrimp and Shrimp Products*, October 12, 1998, 37 I.L.M. 832, paras 7.11–7.17 (1998) (accepting logic of Tuna/Dolphin case in finding that prohibition on importing shrimp caught in ways endangering sea turtles violates GATT Article XI), *affirmed on other grounds*, Shrimp/Turtle Appellate, *supra* note 19.

136 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 118–19 (explaining that the Tuna/Dolphin panels held US restrictions to be of a "coercive design").

137 See Tuna/Dolphin II, *supra* note 17, paras 5.25–5.27, 5.38, 5.39, at 894, 897–8 (characterizing the tuna embargoes as efforts "to force other countries to change their policies with respect to persons or things within their own jurisdiction"); Tuna/Dolphin I, *supra* note 17, para. 5.27, at 199 (characterizing tuna embargoes as unilaterally determining the policies parties must follow in order to have rights under GATT).

138 See Shrimp/Turtle Appellate *supra* note 19, paras 63–4, at 137.

139 See *id.* para. 161, at 166.

140 See *id*. para. 164, at 167 (declaring use of an economic embargo to force a foreign nation to adopt an American regulatory program unacceptable in "international trade relations"); Richard B. Bilder, *The Role of Unilateral State Action in Preventing International Environmental Injury*, 14 Vand. J. Transnat'l L. 51 (1981).

141 See, e.g., David A. Baldwin, *Economic Statecraft* (1985); *Economic Coercion and the New International Economic Order* (Richard B. Lillich ed., 1976); Gary C. Hufbauer and Jeffrey J. Schott, *Economic Sanctions in Support of Foreign Policy Goals* (1983); Gary C. Hufbauer and Jeffrey J. Schott, *Economic Sanctions Reconsidered* (1985); *Principles of International Law Concerning Friendly Relations and Cooperation* (Milan Sahovic ed., 1972); Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations* (1977); Derek W. Bowett, *Economic Coercion and Reprisals by States*, 13 Va. J. Int'l L. 1 (1972); Lori Fisler Damrosch, *Politics Across Borders: Nonintervention and Nonforcible Influence Over Domestic Affairs*, 83 Am. J. Int'l L. 1 (1989) (discussing financial assistance to political campaigns and use of economic leverage); W. Michael Reisman, *Coercion and Self-Determination: Construing Charter Article 2(4)*, 78 Am. J. Int'l L. 642 (1984); Oscar Schachter, *In Defense of International Rules on the Use of Force*, 53 U. Chi. L. Rev. 113 (1986); Oscar Schacter, *The Right of States to Use Armed Force*, 82 Mich. L. Rev. 1620 (1984).

142 See, e.g., Bhagwati, *supra* note 9, at 1 (referring to the "simple elimination of trade barriers"); Farber & Hudec, *GATT Legal Restraints*, *supra* note 10, at 63–4 (distinguishing between "trade restrictions" disguised as policing and adequate environmental protection); David Palmeter, *International Trade Law in the Twenty-First Century*, 18 Fordham Int'l L.J. 1653, 1655 (1995) ("[S]o long as they do not discriminate, so long as they do not erect new trade barriers, governments can do whatever they wish.").

143 See Palmeter, *supra* note 142, at 1655.

144 See Anderson, *supra* note 23, at 767–78 (arguing that Tuna/Dolphin I effectively barred even-handed treatment in favor of discrimination against domestic producers).

145 See generally Charnovitz, *Defogging the Debate*, *supra* note 6, at 478–9 (as the interdependence of economies increases, more environmental measures come within GATT's purview).

146 See US Const. art. I, § 8; Gibbons v. Ogden, 22 US (9 Wheat.) 1 (1824); *Lobsters from Canada*, 3 Can. Trade & Comm. Tax Cas. (CCH) 8182 para. 7.5.1 (1990) (explaining that the restrictions on the sale of lobsters in the Magnuson Act apply to interstate or foreign commerce in order to trigger federal commerce clause jurisdiction).

147 See, e.g., 16 U.S.C. § 703 (1994) (prohibiting the sale or importation of migratory birds); 16 U.S.C. § 971e (1994) (prohibiting the sale or importation of certain fish); 18 U.S.C. § 553 (1994) (prohibiting the importation of stolen vehicles); 18 U.S.C. § 2313 (1994) (prohibiting the importation or interstate sale of stolen vehicles).

148 See, e.g., 21 U.S.C. § 350a(c) (1994) (prohibiting the introduction or delivery for introduction into interstate commerce of nonregistered new infant formula).

149 See Steve Charnovitz, *Dolphins and Tuna: An Analysis of the Second GATT Panel Report*, 24 Envtl. L. Rep. (Envtl. L. Inst.) 10,567, 10,581 (October 1994) [hereinafter Charnovitz, *Dolphins and Tuna*] ("Virtually every product standard has an element of 'force' to it").

150 See Shrimp/Turtle Appellate, *supra* note 19, paras 112, 121, at 149–50, 152–3.
151 *Id*. para. 121, at 152–3.
152 See Tuna/Dolphin I, *supra* note 17, para. 5.14.
153 See *id.* para. 5.15 ("[E]ven if the provisions of the MMPA . . . were regarded as regulating the sale of tuna as a product, the United States import prohibition *would not meet* the requirements of Article III.") (emphasis added).
154 See *id*. para. 5.16.
155 *Id*.
156 *Id*. para. 5.15.
157 *Id*. para. 6.
158 *Id*. para. 6.2.
159 *Id*.
160 *Id*.
161 See *WTO: Report of the Panel on United States – Import Prohibition of Certain Shrimp and Shrimp Products*, 37 I.L.M. 832 (1998).
162 See Shrimp/Turtle Appellate, *supra* note 19, paras 98, 188.
163 *Id*. paras 176, 184.
164 See *id*. paras 143–5 (holding that the US import prohibition was accompanied by restrictions on domestic production).
165 *Id.* paras 176, 186 (indicating that the conclusion is based on the cumulative effects of several arguments).
166 *Id*. para. 161.
167 *Id*. para. 164.
168 See *id*.
169 *Id*.
170 *Id*. para. 165.
171 The decision does discuss the financial and administrative costs of implementing a turtle conservation program in the course of an argument against different phase-in periods for different countries. See *id*. paras 173–5. But the panel does not argue that the administrative costs and financial costs vary by country or make TEDs impracticable. The argument about costs addresses the phase-in periods rather than the justification for a uniform program. At least one analyst has concluded that no other approach works as effectively as a uniform requirement to use TEDs. See Susan L. Sakmar, *Free Trade and Sea Turtles: The International and Domestic Implications of the Shrimp–Turtles Case*, 10 Colo. J. Int'l Envtl. L. & Pol'y 345, 386 (1999). Because sea turtles are endangered, this would suggest that the requirement is quite reasonable.
172 See Shrimp/Turtle Appellate, *supra* note 19, para. 172.
173 *Id*.
174 See *id*. paras 178–84. The Appellate Body actually characterized this as a matter of due process. See *id*. para. 182 ("[R]igorous compliance with the fundamental requirements of due process should be required . . ..").
175 See *id*. para. 180.
176 *Id*. para. 181.
177 See *id*. para. 185.
178 See *id*. para. 186.
179 See *id*. para. 161.
180 See Beef/Hormone Appellate, *supra* note 19, paras 3–5, 244.
181 See Beef/Hormone Panel, *supra* note 76, para. 8.165.
182 See *id*. paras 8.262–8.266.
183 See *id*. paras 8.219, 8.267–8.269.

184 Beef/Hormone Appellate, *supra* note 19, para. 246. *Cf. Australia – Measures Affecting Importation of Salmon*, 1998 WL 731009, paras 159–78 (October 20, 1998) (holding Australian limitations on salmon imports constitute a disguised trade restriction under the SPS Article 5.5).
185 Beef/Hormone Panel, *supra* note 76, para. 8.274.
186 *Id.*
187 *Id.*
188 See *id.*
189 See Beef/Hormone Appellate, *supra* note 19.
190 See, e.g., Farber & Hudec, *Free Trade and the Regulatory State*, *supra* note 21, at 1440 (noting that international tribunals have reached "defensible results" and the quality of decision makers is critical).
191 See, e.g., Schoenbaum, *supra* note 20, at 271–3 (discussing Articles I, III, and XI as background "normative structure").
192 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 118–19.
193 See, e.g., Charnovitz, *Pelly Amendments*, *supra* note 71, at 777–90; Steve Charnovitz, *The Environment vs. Trade Rules: Defogging the Debate*, 23 Envtl. L. 457, 493–501 (1993); Esty, *Greening the GATT*, *supra* note 1, at 114–15 (proposing a test to replace Article XX); Foster, *supra* note 20, at 437–43; Chang, *An Economic Analysis of Trade Measures*, *supra* note 20, at 2135 (proposing an "alternative interpretation" of Article XX); Schoenbaum, *supra* note 20, at 273–80; Padideh Ala'i, *Free Trade or Sustainable Development? An Analysis of the WTO Appellate Body's Shift to a More Balanced Approach to Trade Liberalization*, 14 Am. U. Int'l L. Rev. 1129 (1999) (analyzing Article XX). *Cf.* Weiss, *supra* note 20, at 728–35 (arguing for a conception based on sustainable development).
194 See, e.g., William J. Snape, III and Naomi B. Lefkovitz, *Searching for GATT's Environmental Miranda: Are "Process Standards" Getting "Due Process?"*, 27 Cornell Int'l L.J. 777 (1994).
195 See, e.g., Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 119 (assuming that GATT law permits nothing after discussing Tuna/Dolphin); Jackson, *Congruence or Conflict?*, *supra* note 54, at 1244 (discussing the possibility of "exceptions" to GATT rules to accommodate some process-based restrictions); Charnovitz, *Environmental Exceptions*, *supra* note 70, at 53 [focusing on issue of whether Article XX(b) or (g) can sustain a processing standard].
196 *Cf.* Esty, *Greening the GATT*, *supra* note 1 (employing the title "Greening the GATT"). Professor Esty uses this title to suggest that GATT should become greener, a point that I do not dispute. I only suggest that the browning of the GATT explains why this greening is needed.
197 See Hudec, *GATT/WTO Jurisprudence*, *supra* note 66, at 193–200 (trying to identify the textual basis for the product/process distinction found in the Tuna/Dolphin cases).
198 Tuna/Dolphin I, *supra* note 17, paras 5.8–5.18; Tuna/Dolphin II, *supra* note 17, para. 5.7. See also Deidre McGrath, *Writing Different Lyrics to the Same Old Tune: The New (and Improved) 1997 Amendments to the Marine Mammal Protection Act*, 7 Minn. J. Global Trade 431, 444–68 (1998) (discussing amendments to the MMPA following the GATT panel decision).
199 See GATT art. XI, *supra* note 8.
200 After rejecting the United States' argument that the MMPA provisions at issue created a regulation, the Tuna/Dolphin I Panel quoted the language of Article XI and then stated, "The Panel therefore found that the direct import prohibition . . . and the provisions of the

MMPA under which it is imposed were inconsistent with Article XI: 1." Tuna/Dolphin I, *supra* note 17, para. 5.18.

201 Tuna/Dolphin II, *supra* note 17, para. 5.10 (emphasis added).

202 See Tuna/Dolphin I, *supra* note 17, para. 5.14; Tuna/Dolphin II, *supra* note 17, paras 5.8–5.9.

203 See *id*.

204 See Tuna/Dolphin II, *supra* note 17, para. 5.8 (distinguishing between regulation of a product as a product and regulation of "policies or practices").

205 See, e.g., Snape & Lefkovitz, *supra* note 194, at 785; Anderson, *supra* note 23, at 765–8; Charnovitz, *Pruning*, *supra* note 20, at 311–23.

206 See, e.g., Charnovitz, *Dolphins and Tuna*, *supra* note 149, at 10,576–83 (analyzing both decisions in detail without significant analysis of Article XI); Robert Howse and Donald Regan, *The Product/Process Distinction – An Illusory Basis for Disciplining "Unilateralism" in Trade Policy*, 11 Eur. J. Int'l L. 249, 252 (2000) (arguing that process-based measures do not violate Article III, but not directly challenging the view that such measures would be *prima facie* illegal under Article XI); Hudec, *GATT/WTO Jurisprudence*, *supra* note 66, at 191–3 (analyzing Article III and Ad Note to Article III issues in detail, but simply stating that as a nontariff border restriction, the MMPA's tuna embargo would violate Article XI). *Cf.* Schoenbaum, *supra* note 20, at 273 (briefly describing the broad scope of Article XI as background to a discussion of trade and the environment feature the Tuna/Dolphin cases).

207 See, e.g., India Panel, *supra* note 65, at 569–70 (describing Article XI as broad and comprehensive); *GATT Panel Report on European Community Programme of Minimum Import Prices, Licenses and Surety Deposits for Certain Processed Fruits and Vegetables*, October 18, 1978, GATT B.I.S.D. (25th Supp.) at 68, 98–103 (1979) (invalidating Article XI minimum pricing requirements for tomato concentrate enforced through import licenses). *Cf. GATT Panel Report on United States – Measures Affecting Alcoholic and Malt Beverages*, June 19, 1992, GATT B.I.S.D. (39th Supp.) at 206, 292 (1993) (stating that Article III, rather than Article XI, applies to "internal measures").

208 Farber & Hudec, *GATT Legal Restraints*, *supra* note 10, at 69.

209 See *id*. at 62.

210 See Hudec, *Trade Measures Against Foreign Environmental Practices*, *supra* note 55, at 119.

211 See Charnovitz, *Dolphins and Tuna*, *supra* note 149, at 10,573 n.73 (noting inspectors often examine processing methods because food testing for purity at the border is impractical); *Final Rule: Regulation of Fuels and Fuel Additives: Baseline Requirements for Gasoline Produced By Foreign Refiners*, 62 Fed. Reg. 45, 533, 539–41, 550–9 (codified at 40 C.F.R. pt. 80) (requiring provision of various data and segregation of reformulated gasoline from foreign companies exporting gasoline to the United States).

212 See Charnovitz, *Taxonomy*, *supra* note 126, at 13–15.

213 See *id*. at 13.

214 *Id*. at 13–14.

215 See Reformulated Gasoline, *supra* note 72.

216 See Warren Corp., 159 F.3d at 619–20.

217 The scholarly debate has included some consideration of redefining GATT disciplines. See, e.g., Schoenbaum, *supra* note 20, at 288–90 (discussing proposals to change the definition of "like product" under Article III). But larger concepts of free trade have played little role in this debate.

218 See, e.g., Tuna/Dolphin II, *supra* note 17, para. 5.7.

2

Rules of Power in an Age of Law: Process Opportunism and TRIPS Dispute Settlement

Ruth L. Okediji

Chapter Outline

The WTO Dispute Settlement Understanding is generally considered one of the most important accomplishments of the Uruguay Round. Both developed and developing countries arguably benefit from the new "rule-oriented" system characterized by mandatory timelines, a presumptive consensus for the adoption of reports, and the prospect of sanctions. For developing countries, these attributes ostensibly diminish the role of power in effecting outcomes of disputed claims. Yet, with particular respect to the construction of obligations imposed by the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement, power remains an important consideration for whether a dispute will be resolved by voluntary compliance or involuntarily through the WTO process. This paper examines the conditions under which countries are likely to invest in voluntary compliance with TRIPS, given the information signals associated with each stage of the dispute settlement process. While voluntary compliance is the first-best outcome, such compliance may include side-payments that undermine the long-term interests of developing countries in sustainable access to knowledge goods.

1 Introduction

The conclusion of the Uruguay Round of multilateral trade negotiations ushered in an era of unprecedented formalism in the settlement of intellectual property disputes.

Of all the accomplishments of this monumental Round, none has been as celebrated as the Dispute Settlement Understanding (DSU)[1], and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement).[2] Together, these two treaties constitute an innovative framework for the multilateral regulation of intellectual property rights. The DSU established, for the first time, a means to secure enforcement of global rules for the protection of intellectual property rights. Prior to the TRIPS Agreement, unilateral efforts at enforcement, particularly under the infamous Section 301 mechanism of U.S. trade-legislation,[3] were only partially successful and evoked accusations of abuse of power, particularly by developing countries that often were at the receiving end of this formidable strategy.

Generally, scholars and commentators have considered the DSU a victory for both developed and developing countries.[4] For all the covered WTO Agreements, the DSU provides remedies for noncompliance through sanctions proportionate to the value lost from an uncured violation.[5] These options[6] – compliance or sanctions – raised expectations by developed countries that the bargain represented by the TRIPS Agreement would yield measurable results via improved global market conditions for intellectual property owners. On the other hand, the rule-oriented process of the DSU was perceived as a benefit to developing countries, because the DSU system, unlike dispute settlement under the old General Agreement on Tariffs and Trade (GATT), ostensibly imposes constraints on unilateralism (and thus presumably the use of power) in resolving trade disputes.[7] Scholarly consensus that the DSU process has been significantly improved does not, however, extend to the efficacy of the DSU in securing the obligations of the WTO Agreements and, in particular, the TRIPS Agreement.[8]

The absence of a centralized legislative, adjudicative, and enforcement institution in the international legal system has engendered intractable debates about the utility of international law in influencing the behavior of sovereign states and for establishing cooperative international relations. There is, in fact, a variety of dispute settlement systems/fora in the international society, many of which are identified with specific international regimes, including the DSU for trade disputes arising from the WTO Agreements.[9] In addition, there are a few centralized institutions, such as the International Court of Justice (ICJ),[10] that, in theory, have broad jurisdiction to resolve disputes and apply the force of law to obligations between sovereign states.[11] In reality, however, no ultimate enforcer exists in the international community, and voluntary compliance, not coercion, has been the dominant paradigm for establishing order among sovereign states.[12]

International relations scholars, more specifically those associated with neoRealism, have insisted that compliance in international law is fostered by a network of political and economic interests that provide incentives for governments to conform to an agreed rule or normative principle.[13] This network of interests creates an interdependence between countries that restrains the unfettered exercise of power, because self-interested actors are invested in stable, long-term cooperation for the advancement of mutually beneficial objectives.[14] The rationalization of "compliance" as the outcome of political and economic forces external to international law is seemingly irrelevant to the new trade regime of the WTO. The DSU's mandatory adjudicatory system appears to remedy the classic objection to international law as "law" by providing sanctions for breaches of the covered WTO Agreements.

In this chapter, I argue that the DSU has in fact heightened the importance of extra-legal constraints in effecting compliance with the TRIPS Agreement. I have argued elsewhere, recently, that the "complex interdependence"[15] that is characteristic of international society will prevent the WTO adjudicatory system from becoming a significant source of harmonized copyright law.[16] Here, I explore another facet of that argument, namely that this same interdependence will considerably influence whether states will invest in voluntary compliance with TRIPS obligations. In particular, I am interested in assessing the relative efficacy of a system of sanctions for violations of the TRIPS Agreement in light of International Relations theories that justify compliance in terms of self-interested strategies of sovereign actors pursuing their own ends. Although I focus on rationalist theories of compliance, elements of norm-based and liberal theories are discernible in my analysis. Indeed, a robust evaluation of the structure of the DSU and its effect on compliance requires a careful blend of these causal theories.

Two questions frame my inquiry about the effect of the DSU on compliance with TRIPS obligations. First, under what conditions will a mandatory system of adjudication, backed by sanctions, effect greater compliance with the TRIPS Agreement? Second, is such a system more or less susceptible to the disparities in power between the Member States of the WTO?

I start with the premise that an equilibrium is established where there is no need to use the DSU because all countries are in compliance with the TRIPS Agreement. Such equilibrium would reflect, among other things, that Members desire to secure the benefit of the WTO bargain and, consequently, compliance with TRIPS is the most rational choice. More important, for this condition of equilibrium to exist, each country must *consider* all other Member countries to be in compliance. The equilibrium accomplished by such "voluntary compliance" is the first-best outcome. A condition of costless equilibrium is admittedly improbable; thus, I use the term "voluntary compliance" also to denote a state's positive response to an allegation of TRIPS violation *prior to* exhaustion of the DSU process. In other words, voluntary compliance means that, at any point prior to a final decision by the TRIPS panel, a country chooses to rectify the alleged violation. Although questions may arise over the effectiveness of the remedial action, I omit this element from my immediate analysis and focus, instead, on the fact that a state takes reasonable measures to redress a complaint of noncompliance in a manner that averts the adjudicative process. I further consider how the procedural design of the DSU influences states to choose voluntary compliance, and how information asymmetries about the TRIPS provision at issue affect such a choice.

Section 2 provides a brief overview of the design of the DSU. I employ a market analogy, characterizing each step in the DSU process as a "signal" that conveys information to another Member. Information asymmetries are inherent in the different kinds of provisions embodied in the TRIPS Agreement. In Section 3, I discuss how these asymmetries inform Member responses to the signals. In Section 4, I consider the notion of compliance as a two-stage process. Stage one focuses on integrating the international obligation into domestic law by making the required legislative changes. Stage two compliance primarily involves judicial bodies and other domestic institutions necessary to secure the benefits of protection for intellectual property owners. I briefly discuss how the DSU might affect decisions by Members to invest in stage two compliance, and the implications for the TRIPS Agreement. I conclude, as have other scholars,

that informal, voluntary resolution of TRIPS disputes is a more effective strategy for securing compliance with the TRIPS Agreement. However, this contribution addresses the structure of the DSU and how it might influence and facilitate efforts at informal dispute resolution in light of (or despite) socioeconomic and political forces that countries must consider when implementing treaty obligations and when electing strategies to enforce compliance by other states.

2 Voluntary Compliance and the Progressive Stages of Dispute Settlement Under the DSU

The various explanations for uncoerced compliance with treaty obligations have been developed mainly through the work of international relations scholars, who are skeptical about the relevance of law in a community of sovereign states.[17] Increasingly, however, international lawyers also acknowledge the significance of extra-legal factors in understanding why and how nations behave.[18] Sovereign states obey international laws for a variety of reasons unrelated to the existence of a centralized authority to interpret and enforce such laws. Justifications for compliance by states include: considerations of a state's honor and reputation, the convergence of domestic industry priorities with the provisions of a treaty,[19] and the persuasive moral pressure stemming from the existence of the international obligation.[20] Scholars have also explained compliance with international law in terms of the benefits of cooperation facilitated by international institutions or "regimes."[21]

With its mandatory system of adjudication, and the availability of sanctions, it would seem that the trade regime does not suffer from the weakness of other public international law arrangements; an effective system of enforcement presumably obviates the need for alternative justifications for state compliance with TRIPS obligations. However, the importance of steady, cooperative relations in the international sphere still imposes voluntary compliance as the model for enforcing legal obligations among sovereign states. Indeed, despite the import of the new dispute settlement system, the provisions of the DSU reinforce the primacy of cooperative dispute resolution as the optimal strategy for promoting stable international relations and strong global markets. Article 3 of the DSU cautions members to "exercise . . . judgment as to whether action under these procedures would be fruitful. . .. A solution mutually acceptable to the parties to a dispute and consistent with the covered agreements is *clearly to be preferred*."[22]

It is useful to think of the DSU process as a kind of market where information about an actor's grievance is bought and sold.[23] Disputing Members also bargain over the terms and conditions of informal dispute settlement opportunities. The DSU alternatives (of damages or sanctions), rather than compliance, reinforce this notion of a "market" for dispute settlement. In such a market, the various stages of DSU adjudication serve as "signals" from the complaining party, who invokes the formal process. Another characteristic of this market is that settlement bargains are made more possible when there is information symmetry between the disputing Members, and less possible as information asymmetry develops.[24] A country's response to initial allegations of

a TRIPS violation should ideally reflect the degree of information asymmetry. Finally, at each stage of the DSU process, the defending party has the opportunity to respond to the signals, and the complaining party may choose to stop sending signals (effectively terminating the formal process), based on the quality of information each Member has about their respective claims/counterclaims.

Under the DSU, parties may choose to opt out of the formal process, and settle the dispute informally. It is conceivable that the possibility of opting out dilutes the strength of each signal, thus making the actions of the parties harder to predict or assess. However, the fact that opting out of the formal process is subject to *mutual agreement* alleviates this problem, because neither party can be sure what conditions will induce the other to agree to opt out. Consequently, the signals triggered by each stage of the formal process will be the most accurate source of information regarding what the other party believes about the alleged violation.[25] The parties can then evaluate the information against the risk of sanctions and/or the costs of involuntary compliance.

An important point is that the closer the relationship between the disputing parties, that is, the more they share political or economic interests, the more likely that there will be bargaining chips to facilitate an agreement to opt out of the DSU system. This is particularly true where the disputing parties are developed countries that both feel strongly about their individual positions. Rather than risk undermining the integrity and credibility of the DSU system by refusing to implement Panel/Appellate Body decisions, the disputing developed countries may choose instead to settle their dispute through alternative means. This assumes, of course, that developed countries (or particular combinations of developed countries) will enjoy greater welfare gains from ensuring compliance by developing countries, than from policing the bargain between themselves.[26]

Even where a dispute has gone through the entire DSU process, the parties may still resolve their differences informally. This happened, for example, in the dispute between the United States and the European Union over Section 110(5) of the U.S. Copyright Act.[27] Such "after the fact" agreements are forms of side-payments that heighten general welfare because: (1) the resolution was mutual, thus promoting cooperation and avoiding a loss of credibility for the WTO; and (2) the DSU prohibits resolutions that are inconsistent with its Members' obligations. Thus, such side-payments can be Pareto-efficient as between the disputing parties, even if discriminatory in its impact, because developing countries may not have similar opportunities/resources to bargain around the outcome of an adjudicated dispute.

The lesser the degree of interdependence between parties (and in the absence of convergent interests/shared alliances), the more difficult agreeing to opt out may be for parties, because they have no long-term interests at stake. The political economy of international society suggests that countries with "thin" relationships should be encouraged to invest in voluntary compliance, because they have no direct or immediate incentive to pay attention to the signals sent by the other country in the dispute. The lack of incentive to acquire information, and the resulting asymmetries may mean that disputes that can and should easily be resolved will instead go through the entire DSU process. *Thus, interdependence facilitates the optimal condition of voluntary compliance.*

It is true that the political and economic costs of dispute resolution may constitute sufficient incentives to acquire the information about the alleged dispute. However, as I discuss in Section 3, this will depend on the type of legal obligation at issue – a "rule" or a "standard," the cost of acquiring information about the legal obligation, and about the other party's legal position. In addition, there is the cost of endogenous factors within each party's domestic setting that may add to or outweigh the other costs. Such factors include divisions between domestic institutions over how to comply with the international obligation,[28] policy preferences that may tilt compliance strategies in directions that are in tension with the spirit of the treaty,[29] and the activities of interest groups.[30]

The entire formal DSU process can be divided into six progressive stages: (1) consultations;[31] (2) establishment of a panel; (3) adoption of panel reports; (4) appeals; (5) recommendations of the Appellate Body; and (6) implementation of the recommendations of the Appellate Body. At first glance, only the first two stages are within the control of the Members; as mentioned earlier, however, Members may opt out, subject to mutual agreement, at any stage of the formal process to pursue alternative means of settling the dispute. The Director-General may offer good offices, conciliation, or mediation to help in such settlement talks.[32]

At the consultation stage, the complaining party signals the other Member country that it considers a feature of its laws to be inconsistent with the TRIPS Agreement. If the allegedly violating party agrees with this allegation, ideally the parties will terminate the formal process and resolve the dispute at this stage. The strength of the signal sent by initiating consultations can be weak or strong, depending on the complaining party's information about what the TRIPS Agreement requires of Member countries. Similarly, the defending party's response to the signal will depend on its own information about how to comply with the TRIPS obligations. In reality, however, the formal initiation of consultations sends a strong signal to the alleged offending country because, generally, countries would have already engaged in diplomatic talks prior to invoking the DSU mechanism. Consultations may also facilitate more active involvement from other government agencies and political actors, who together may exert sufficient political pressure to steer the process toward informal settlement or whose influence may further harden the position in favor of DSU resolution. Under the DSU, then, consultations provide a second, and now "formal," opportunity for parties to disclose relevant information so that the complaining party can fully assess the merits of its claim. Ideally, this opportunity to exchange information and assess the strength of their respective cases will facilitate settlement negotiations.[33] What will not be exchanged or easily discernible is the extent to which internal domestic politics will affect the opportunity sets for cooperative resolution. As I will discuss more fully, this factor seriously complicates the dynamics within the formal dispute resolution process by introducing information imperfections that powerfully affect the market equilibrium for TRIPS adjudication.

The DSU's provisions attempt to eliminate incentives for parties to withhold information during the consultation stage in two ways: (1) consultations are confidential; and (2) the information shared in these consultations is shared without prejudice to the rights of any of the parties in later stages of the proceedings.[34] Indeed, the dispute panel has no knowledge of what the parties discuss in consultations; the panel's function is simply to

ascertain that the parties did in fact hold consultations prior to the request for the establishment of a panel.[35] In the absence of information asymmetries, the optimal outcome at this stage of the dispute resolution process is that parties resolve the dispute or the complaining party gains additional information that strengthens its resolve to continue with the process.[36] Of course, this exchange of information may increase the defending party's confidence in the strength of its defense, causing it to ignore the signal by refusing to make the necessary adjustments to satisfy the complaining party. The dispute then moves onto the next stage, which is the request for the establishment of a panel.[37]

This second signal indicates that the complaining party feels strongly about the alleged violation, that the alleged violation effects a loss in value to the complainant, and that the complainant is willing to undertake the costs of dispute settlement, even if the only outcome is that increased and reliable information about the provision at issue is disseminated. Although the DSU gives Members the right to seek authoritative interpretation of the provisions of a covered Agreement,[38] a complaining party may lack adequate incentives to pursue a claim simply for clarification of a TRIPS provision, *since the information cannot be kept private*. Panel and Appellate Body reports are circulated to all Members for consideration.[39] Consequently, there may be insufficient incentives for one Member to acquire the information unless there is a mechanism to appropriate the returns,[40] or the potential gain from a clarification in the complainant's favor outweighs the costs of the dispute.[41] The fact that there is no costless way to acquire authoritative information about the provisions of a covered Agreement is a weakness in the structure of the DSU with regard to establishing incentives for voluntary compliance.

The defending party may respond to the request for the establishment of a panel by agreeing to change its laws/practices or by ignoring the signal. The response will depend on the quality of information held by the defending party about the legitimacy of the claim and the sufficiency of its own defense. The response may also depend on the domestic costs of compliance.[42] If those costs are significant, such as economic instability, sectoral/industry agitations, threats to the aspirations of political actors, etc., a country will nevertheless ignore the signal, even if it is persuaded that it is in violation of a TRIPS provision.[43] Such was the case in the dispute between the European Union and the United States over Section 110(5) of the U.S. Copyright Act.[44] The European Union alleged that the exceptions to copyright protection embodied in Section 110(5) were inconsistent with Article 13 of the TRIPS Agreement. The Dispute Settlement Body (DSB) ultimately agreed with the European Union by adopting the panel report, which held that one of the exceptions was inconsistent with Article 13.[45] At the enforcement stage, the United States requested the maximum time frame for compliance,[46] citing the politically sensitive nature of Section 110(5), and arguing that it could not be easily or quickly modified. The Arbitrator rejected this justification for delaying compliance.[47] Canada offered a similar argument in a complaint by the United States alleging a violation by Canada of the TRIPS provision concerning the term of protection for patents.[48] In both cases, the Arbitrator rejected these domestic concerns as a legitimate reason for delaying compliance.[49] In the Section 110(5) decision, the United States opted to pay damages to the European Union rather than amend its legislation.[50] However, as of this writing, the two parties remain in negotiations over the amount of the compensation to be paid.

Once the dispute panel has issued its rulings and recommendations, its decision may be appealed by the losing party or it may be adopted by the DSB. If the losing party appeals the decision, the Appellate Body has 60 days to conclude its proceedings.[51] A Member's use of the appeal stage sends mixed signals. It may simply be a political strategy of the losing party, attempting to allay domestic pressure from adversely affected interest groups. An appeal may also convey that the losing party strongly believes that all or part of the panel's decision is inconsistent with the TRIPS bargain, and the losing party is seeking genuine correction/clarification. At this stage of the game, there are already sunk costs associated with the process. Thus, even a minimum incentive to appeal will likely induce this signal, particularly if the returns of a successful appeal will ameliorate the loss at the panel level.[52] The India–Patent Protection dispute is illustrative.[53] Although India lost on the substantive claim, on appeal, the Appellate Body overturned one of the panel's holdings on a matter of law. Citing GATT jurisprudence, the dispute panel held that WTO Members are required to protect legitimate expectations about the competitive conditions in the domestic economies of other Members. The panel determined that India's implementation of TRIPS Article 70.8 did not "sufficiently protect the legitimate expectations" of the United States, in light of the competitive relationship between the respective countries.[54] Thus, the panel concluded that India's implementation mechanism failed to satisfy the obligation of good faith required by the Vienna Convention.[55] On appeal, the Appellate Body rejected the "legitimate expectations" test, which, in effect, enlarged the scope of TRIPS obligations.[56] Employing a contract law framework, the Appellate Body concluded that the text of the TRIPS Agreement reflects the expectations of Member States, and reinstated "strict constructionism" as the appropriate interpretive rule for disputes.[57]

From a developing country perspective, the Appellate Body's decision is particularly important, because it limits the scope of TRIPS obligations to the "four corners" of the Agreement, thus preempting possible demands for developing countries to guarantee conditions for intellectual property owners that are similar to those in developed countries, including similar interpretations of intellectual property norms.[58] The decision has global welfare implications as well: a rule of strict construction best preserves the equilibrium point of the Agreement, as negotiated, and allows Members to enjoy the full benefits of the international bargain.[59] What constitutes a "benefit" of the TRIPS Agreement is, however, subject to the interpretation of the provisions in dispute. The scope of a panel's discretion in interpreting TRIPS provisions is, in turn, affected by whether the provision was formulated as a standard or rule.

3 Norms, Rules, and the Efficacy of Signals

Law and economics literature distinguish between two types of legal principles: rules and standards.[60] For purposes of my analysis, I adopt the following assumptions that flow from this distinction. A law is considered a "rule" if it is an *ex ante* determination between negotiating parties specifying how the parties should conduct themselves in relation to a particular subject matter. Such *ex ante* determinations are costly to formulate, but easier for adjudicators to enforce. Alternatively, a "standard" simply

provides guidance about the outer boundaries of what is expected or prohibited, and does not specify in complete detail how the parties should behave. Standards are cheaper to formulate, but more costly to enforce. They are more susceptible to interpretive flexibility that invariably imports value judgments about fairness, efficiency, and normative appropriateness into "the law." A classic example of a standard is the requirement in the domestic law of contracts, mirrored in the international law of treaties, that parties must act in good faith.[61] This good faith standard is an outer boundary that all contracting parties must operate within, but there is no definitive guide as to what constitutes good faith. Courts typically make this determination on a case-by-case basis. Further, there is no such thing as a pure rule. Rules and standards are mixed in varying degrees in every legal regulation.[62] Standards inform the application of a rule, because they embody the underlying norms that gave rise to the rule, or they reflect the overarching policy concerns that require consideration of external factors when interpreting the rule. The process of rule interpretation, particularly in common law systems, invariably relies on standards to facilitate robust and context-specific application of the law.[63] Finally, rules are likely to effect greater levels of compliance, because the cost for individuals of acquiring information about what the law requires is less than what obtains under a standard. Rules convey information that allows parties to better assess the costs of compliance, because they are more certain of the (or can easily access) expectations that stem from the legal regulation.[64] A rule is particularly efficient if the conduct to which it applies is likely to be frequent.[65]

International treaties, like domestic contracts or legislation, also consist of rules and standards. As in the domestic setting, rules in international law are more costly to negotiate than standards,[66] in part because of the multivariate factors that affect a sovereign's willingness to be identified with, or bound by, a particular rule or "hard law." The interplay between domestic politics and international negotiations adds costs to international rule-making over and above the administrative, drafting, and negotiation costs that are also common to domestic rule-making. This domestic/international interplay further introduces unique strategic costs into international negotiations that are exacerbated by the number of parties and issues involved in the negotiations.[67] For example, during the Uruguay Round negotiations, developing countries took part in issue-specific coalitions, rather than forming the traditional "South" alliance on ideological grounds. Other coalitions, also formed during the Round, included a mix of developed and developing countries during the TRIPS negotiations over issues such as compulsory licensing, geographical indications, and a rental right for sound recordings.[68] Indeed, tactics facilitated by coalitional commitments were the significant characteristic of the TRIPS negotiations.[69] On several important issues, the coalition strategies resulted in the adoption of open-ended standards instead of concrete rules.[70] However, the multiple issues within the TRIPS negotiation, and between TRIPS and other subjects of the Uruguay Round, created a complex mosaic of interests and alliances that ultimately facilitated the successful conclusion of the Uruguay Round.[71]

The unique costs of negotiating rules at the international level, particularly the costs to sovereignty, may suggest that treaties are more likely to consist of standards rather than rules.[72] This is certainly true in areas where some regulatory framework is needed, but parties are unwilling to make the political commitment to a rule.[73] Certainly, negotiation of the major intellectual property treaties initially focused on promulgating

standards and then progressively to rule-making. The Berne Convention for the Protection of Literary and Artistic Works[74] is a classic model of this evolution.[75] Over time, informal interactions, strategic alliances in other areas, and the gains of multilateral cooperation led ultimately to a willingness to establish additional rules to enable countries to maximize their joint gains from cooperation in intellectual property matters.[76] The trade system is another example of such evolution: the general standards embodied in Friendship, Commerce, and Navigation (FCN) treaties developed into the network of rules embodied in the various WTO Agreements. A review of all the different Rounds of multilateral trade talks also demonstrates the progression toward rule-making as a dominant form of law in trade agreements. Both in terms of the details of the rule, and the subject matter addressed, the multilateral trade system has increasingly been characterized primarily by rules and not standards.

The TRIPS Agreement consists of both rules and standards. Whether this analytical distinction affects compliance efforts of Members is further conditioned by the perceived enforceability of the provisions and the knowable costs of violations. The adjudicated cases so far suggest that Members do consider the TRIPS provisions as legitimate obligations, and to the surprise of many, the DSU procedure has been remarkably successful in effecting change in the behavior of Member States on a variety of issues. In other words, Members consider the DSU a legitimate and authoritative source of what the WTO obligations mean and what they require in terms of behavior by Members. Further, compliance efforts are examined and taken into account during dispute panel examinations of allegations of breach. These conditions have confirmed the DSU system as a supranational adjudicatory process with the requisite legitimacy to enforce WTO obligations.

Under conditions of legitimate enforceability, a rule is more likely to have a positive effect on compliance, because parties are better able to ascertain the costs of noncompliance.[77] In the international context, these include costs associated with creating the rule, that is, negotiation, drafting, and administration, and costs associated with noncompliance, including dispute resolution, loss of reputation, and an erosion of the gains that the bargain (at least in theory) represents. Of course, if these costs combined are less than the gains from breach, breach of the agreement may still be the optimal strategy.[78]

If the conditions for enforceability are relaxed, for example in situations where the possibility of informal settlement is high, the distinction between rules and standards will be less significant with regard to whether powerful Members will adjust their behavior to conform to the legal prescription or prohibition embodied in the TRIPS Agreement. Reputation costs may not significantly affect future bargaining opportunities.[79] More important, in the context of the TRIPS Agreement, powerful countries already had similar levels of intellectual property protection and these levels were substantially incorporated into the TRIPS Agreement. Thus, one might conclude that the distinction between rules and standards, at least for *developed* countries, will not have much import in influencing compliance, since these countries were in substantial "compliance" prior to the negotiation of the TRIPS Agreement.

However, for several reasons, this conclusion is questionable. The standards employed in the TRIPS Agreement reflect those areas where developed countries could not agree on a harmonized rule. These standards thus allow for what Jerome Reichman

has famously called "wiggle room," permitting countries to implement the provision in a manner most consistent with domestic priorities. Notwithstanding this discretion, most of the disputes thus far between developed countries have involved the construction of TRIPS standards.[80] As these disputes illustrate, to the extent a country's interpretation of a standard erodes monopoly gains of foreign competitors, there will be disputes over the "proper" interpretation of these TRIPS provisions. The problem is not limited to standards alone. The interpretive flexibility that is implicated even where rules are at issue may also generate disputes about the adequacy of compliance efforts by Members.[81] Thus, the difference between standards and rules in the TRIPS Agreement may be only of limited value in predicting how *developed* countries might effectuate change in their domestic legislation, and in *ex post* assessments of the sufficiency of efforts to enforce the substantive obligations.

Whether a dispute will go through the entire DSU process is partially a problem of asymmetric information resulting from different interpretations of TRIPS provisions. While rules *per se* cannot entirely tame the flexibility inherent in the language of the law, the interpretive scope associated with rules is less than what pertains to standards. Consequently, alleged violations of a "rule" (e.g., copyright term must last for life of the author plus 50 years) involve less information asymmetry, and it is less likely that the dispute will go through the entire DSU process.[82] Conversely, disputes involving interpretations of standards (e.g., what constitutes a legitimate exception/limitation to copyright under Article 13 of the TRIPS Agreement) are more likely to go through the entire DSU process, because such disputes are characterized by significant information asymmetries. In a sense then, for those countries with the resources for it, disputes create a secondary market for renegotiation of existing bargains, much like reinsurance carriers or underwriters. Such renegotiation takes place during the DSU process, as parties exchange information, consider competing interpretations and, ultimately, when the panel or Appellate Body issues a definitive ruling about what the standard means and/or what constitutes a violation of the standard.

If TRIPS disputes are as likely to occur under rules as they are under standards, the presumption in favor of rules over standards clearly does not hold within the DSU framework. It is plausible that the market for renegotiation permits Members to recoup some of the costs associated with the rule, assuming such renegotiation results in a mutually acceptable settlement that engenders welfare gains for at least one of the disputing Members. Such welfare gains should also exceed the costs of voluntary compliance with the TRIPS provision at issue. Thus far, there is no empirical evidence to suggest that TRIPS rules are disputed less frequently than standards. However, if the first-best outcome is voluntary compliance, then the renegotiation market internalizes the costs of negotiating "new" rules, while also encouraging compliance by reducing the costs of information acquisition. Thus, a case can be made for the relevance of the distinction between rules and standards as it affects levels of voluntary compliance with TRIPS obligations.

The value of the renegotiation market is meaningful only to the extent that the "settlement rule" arrived at between disputing parties has a network effect in that other Members do not perceive that they are adversely affected by the private bargain. Article 3 of the DSU, in part, facilitates positive returns from private settlements for all Members, by requiring that such settlements be consistent with Member obligations

under the covered Agreements. This may make it more likely that other Members will buy into the agreement reached by the disputing states. A network effect is almost guaranteed in disputes between powerful sovereigns, such as the European Union and the United States, whose interests are strongly represented in the TRIPS Agreement, and strengthened by the number of states who join the complaint as interested third parties. In other disputes involving fewer powerful countries, however, the creation of a network effect will be determined at least in part by the existence of multiple complainants plus a requisite number of interested third parties whose involvement in the dispute, no matter how marginal, may serve to ensure that the settlement is substantially consistent with their understanding of the obligation at issue. In short, the settlement value of a private bargain increases proportionally to the number of states who are willing to accept the bargain as TRIPS-consistent.

The effectiveness of signals during the dispute resolution process largely depends on the information held by the parties about the TRIPS provisions at issue. Where a TRIPS dispute involves a rule, the signals sent are typically stronger and clearer and are thus more likely to yield optimal results by way of voluntary compliance. This is, of course, one of the benefits of enacting laws as rules rather than standards.[83] For example, in a complaint filed by the United States against Greece and the European Union over the lack of copyright infringement remedies in Greece, in violation of TRIPS Articles 41 and 61, Greece agreed to enact legislation to provide an enforcement remedy for copyright owners whose works had been infringed by television stations in Greece.[84] Similarly, other complaints filed by the United States against Denmark and Sweden over the failure of these countries to enact provisional measures as required by TRIPS Articles 50, 63, and 65 were resolved by mutual agreement. Denmark and Sweden amended their laws to comply with TRIPS requirements.[85] Other examples of voluntary compliance include complaints by the United States against Portugal over Portugal's failure to implement TRIPS provisions with regard to patent terms;[86] a complaint by the European Union[87] and the United States[88] against Japan over Japan's treatment of sound recordings in violation of Articles 14.6 and 70.2 of the TRIPS Agreement; a complaint by the United States against Pakistan over Pakistan's failure to comply with TRIPS Articles 27, 65, and 70, regarding patents for pharmaceuticals and agricultural chemicals, a patent application filing system for such products, and exclusive marketing rights to such products.[89] The parties settled each of these disputes prior to the exhaustion of the DSU process.

Despite the "rules" involved in these disputes, however, no dispute was settled prior to invoking the formal process of the DSU. This suggests that while mere threats of DSU action may in fact effect compliance in some cases,[90] using the DSU process conveys information that allows parties to "buy" and "sell" opportunities to comply voluntarily with the TRIPS Agreement, that is, the market for renegotiation. To reinforce my earlier point about welfare gains from such renegotiation, voluntary compliance in such cases suggests that both parties negotiated an equilibrium point that left them better off than if the dispute had: (1) never taken place; or (2) if the dispute had gone through the entire DSU process. Even for the defending party, waiting until the first stage of the DSU to settle the dispute may serve some political ends in the domestic context.

Of course, there may be other ways to explain the cases mentioned above. In particular, the dispute between the United States and Greece may be distinguished

as a case involving the difference between "plausible" compliance and "minimum" compliance.[91] The dispute was not as much over *how* Greece had implemented the TRIPS provision at issue (i.e., plausible compliance) as it was that Greece had not implemented the provision at all (i.e., minimum compliance). The latter being the case, this dispute is arguably not about the relative efficacy of rules and standards, but about using the DSU to pressure countries to send *some* signal (i.e., to do *something*) in the first instance.

These cases may also be about the cost of errors. When there is significant commitment to the underlying substantive principle, the distinction between rules and standards is useful for evaluating the conditions under which a developed country will utilize the DSU process and its outcome as a way to justify unpopular changes to domestic legislation. Or, a country could choose not to comply and employ the DSU process to ascertain the economic costs of noncompliance. Scholars have described these strategies or tactics as "process opportunism," where Members exploit information asymmetries to affect the choices that other Members might make.[92] If there is sufficient concern about domestic repercussions for "strict" voluntary compliance with the TRIPS Agreement, political actors may engage in opportunism by doing nothing until the formal DSU process is invoked, thus providing some insulation from domestic criticism and pressure from groups adversely affected by the changes required by the TRIPS Agreement.[93] The distinction between rules and standards also has some bearing here. Standards, more than rules, may be more expedient for political actors, because they allow Members to manipulate meaning in a manner that is not easily characterized as illegitimate by domestic interest groups, who may favor an interpretation different from what the government has adopted.

Where the alleged violation occurs over a norm (or standard), the signals are fuzzy, and it is not clear that only one rational response exists. The resulting quandary presents a sort of prisoner's dilemma. The defending party's best strategy will depend on factors (or incentives) such as the domestic political costs, particularly where there are domestic constituencies/interest groups vying for one outcome over another; or where there is a domestic policy that has welfare-enhancing properties within the domestic context. For example, a complaint filed by the European Union against Canada[94] over exemptions/limitations in Canadian patent law required the interpretation of, among other things, the standard articulated in TRIPS Article 30. Canada argued that the legislation in question was TRIPS consistent based on Canada's interpretation of TRIPS. Indeed, Canada had engaged in a deliberate study of the welfare effects of its legislation, including discourse with members of the public at large, before enacting the legislation at issue. Canada had also considered the TRIPS-consistency of the legislation. The combined costs of drafting the domestic legislation, what appeared to be good faith efforts to comply with TRIPS, and the domestic welfare concerns that motivated the legislation suggest that Canada's optimal strategy was to risk adjudication and acquire information about the precise scope of Article 30. The WTO panel ruled that Canada's stockpiling provision was inconsistent with Canada's obligations under the TRIPS Agreement. However, the panel found that the other challenged provision was narrow enough to be consistent with TRIPS Article 30.[95]

The dynamic relationship between domestic interests and compliance with international obligations suggests that countries cannot always easily ascertain the costs

of compliance at the point of treaty negotiation. This dynamic relationship is another factor that makes it difficult to assess when it is more efficient to invest in a rule, rather than a standard, for the purposes of affecting the behavior of parties.[96] Another important point is that countries who wish to exploit the flexibility inherent in a standard by enacting domestic legislation that furthers a domestic policy may instead choose more conservative policy options to avoid the risks of WTO adjudication. As I note later with respect to developing countries, a conservative strategy has some negative welfare implications.

When the purpose of a rule is not solely to effect compliance, but also to set the equilibrium for parties so that they can maximize joint gains, then rules seem to be more efficient than standards, *because they facilitate certainty about the returns of mutual cooperation*. Thus, as distinct from the domestic context, where the primary purpose of the rules versus standards debate is to determine how each type of legal regulation might affect a party's behavior, in international contexts the debate adds to the information available to each state in determining whether the benefits under the treaty outweigh the costs of membership and/or if the benefits are worth the costs of negotiation. Under this model, a rule is a socially optimal alternative to a standard if, and only if, the gains from mutual cooperation outweigh the possibilities of unilateral action. Indeed, as the gains of cooperation increase, the more valuable it is to have a rule. In addition, the larger the number of members of the treaty, the more efficient it is to negotiate a rule, otherwise the potential gains from mutual cooperation might be diluted through the different domestic policy preferences for specific interpretations. These two observations are corollaries to the argument in the domestic setting that when a law will apply frequently, it is best to promulgate it as a rule and not a standard.[97]

4 Compliance Investment and the Role of the DSU

I have discussed TRIPS compliance primarily in terms of Members enacting domestic legislation consistent with the TRIPS provisions. Indeed, the majority of TRIPS disputes thus far address the failure of Member States to amend domestic legislation to conform to TRIPS requirements. As several commentators have already noted,[98] the more significant challenge for TRIPS compliance, particularly in developing countries, is building domestic institutional capacity to enforce the rights in the domestic setting.[99] The TRIPS Agreement does not require countries to establish new judicial and administrative structures for intellectual property rights, nor does it require countries to devote additional resources to intellectual property enforcement.[100] However, the good faith obligation imposed by the Vienna Convention will require attention to capacity-building issues with regard to efforts of owners to enforce rights in domestic courts. Compliance requires both a formal adherence to the treaty language through legislative change, as well as substantive conformity to the expectations and objectives of the treaty, as intellectual property rights owners seek to enforce their interests in Member States. In this regard, compliance is a dynamic, long-term process involving the coordinated efforts of domestic institutions such as customs officials, courts, and other domestic agencies involved in law enforcement.[101]

In a manner strikingly different from the DSU's detailed procedural rules, the enforcement provisions of the TRIPS Agreement consist almost entirely of standards, leaving the judicial authorities in Member States to determine on a case-by-case basis how enforcement should take place and what remedies are appropriate. Specifically, the TRIPS Agreement requires Members to confer wide discretion on "judicial authorities" to order a range of remedies for parties whose intellectual property rights have been infringed. It is well understood that the capacity of domestic institutions to interpret, apply, and enforce the TRIPS provisions will vary from country to country, and will be conditioned by the socio-legal and economic forces in a particular Member State. This is the case both for developed and developing countries. The pertinent issue, then, is what role the DSU might play in encouraging fidelity to TRIPS provisions by judicial and enforcement institutions of a Member State.

The DSU process is more removed from the domestic judicial system than the other branches of government, and is less likely to be directly influenced by the political conditions that facilitated Membership in the WTO in the first place. Thus, the DSU signals may not easily "penetrate" or convey meaning to domestic courts. In constitutional democracies, judicial bodies are also generally immune from incursions by the legislative and executive branches. It might seem, then, that after integrating the TRIPS provisions into domestic legislation, Members may argue that they have satisfied their international obligations. Conforming legislation is indeed one of the first steps toward compliance. Although domestic courts are the ultimate enforcers of these obligations, there remain significant opportunities for Members to invest in TRIPS compliance. This may be done by strengthening extra-judicial agencies such as the police and customs officials. It may also be possible to utilize administrative agencies to address particular aspects of domestic enforcement procedures, such as border measures under Section 4 of the TRIPS Agreement. Certainly, Members are not obligated to invest in creative approaches to enforcement.[102] However, such endeavors could be considered by a panel as evidence of good faith efforts to comply with international obligations, as required by the Vienna Convention.

There is another dimension to the rules versus standards debate that bears more directly on the issue of compliance by domestic institutions. "Standards" negotiated at the international level do not necessarily have to be implemented as "standards" in the domestic legislation. Meaningful investment in TRIPS compliance suggests that Members implement TRIPS obligations in a manner that provides clear guidelines and objectives for domestic institutions. Specifically, domestic implementation of TRIPS obligations could consist more of rules and not ambiguous provisions. The danger of this approach to implementation is that such domestic legislation is easily accessible by other Member States[103], and can thus be challenged more readily before a WTO panel. The implementation of TRIPS standards as "standards" in domestic legislation is, in effect, a better strategy for countries wishing to reserve to judicial authorities the opportunity to apply TRIPS obligations in a manner consistent with domestic political and economic realities. The very effectiveness of the DSU process thus creates perverse incentives for countries to comply with the TRIPS Agreement using open-ended standards rather than determinate rules.

An uncritical application of the rules versus standards debate to the international context can generate some questionable results. The domestic legal context is structured

on the basic premise that the law is legitimate and enforceable. Thus, under certain conditions, legislatures can avoid the costs of enacting rules without a corresponding loss in compliance levels.[104] As noted earlier, the enforceability of international law is both controversial and controvertible: while international laws are often obeyed,[105] such compliance has been justified as the outcome of several different forces.[106] Of these, domestic interests play a particularly important role, given the incentives of political actors who enter into these international agreements, principal of which is the incentive to act in ways that will sustain, if not improve, their domestic political support.

In most developing and least-developed countries, domestic pressures are unlikely to play a significant role in securing compliance with the TRIPS Agreement. The strength of domestic governance plays an important role in whether or not process opportunism is a viable (or necessary) strategy for a country.[107] The weaker the political system (in democratic terms), the more likely it is that compliance with TRIPS will be less affected by the domestic political process. Indeed, TRIPS compliance may be at very high levels in a dictatorship;[108] a dictator may comply more readily with rules and over-comply with a standard, than a political leader of a strong democracy, who must account for the reactions of domestic constituencies.[109] Further, autocratic regimes can circumvent domestic judicial processes or overrule decisions that are considered noncompliant. The unitary model of government in many developing countries may indicate strong prospects for compliance where domestic institutions are concerned. Whether intellectual property rights owners will be willing to exploit these undemocratic conditions for their own ends is an entirely different issue.

Other factors are important for evaluating the prospects for voluntary compliance by developing countries. First, it is widely acknowledged that the real "bargain" over intellectual property protection during the Uruguay Round negotiations occurred as an exchange for freer trade in agriculture and textiles. Consequently, it may be argued that as far as developing countries are concerned, the gains from TRIPS compliance lie in the increased access to developed country markets in relation to agricultural goods and textiles. Incentives to invest in compliance arise, then, from the significant value that developing countries place on market access in those sectors over intellectual property protection. The possibility that retaliation for violations of the TRIPS Agreement may aim at these crucial areas will certainly deter some developing countries from active breach of the TRIPS obligations.

Further, few developing countries have the resources to pay for or bargain over sanctions in the event of an adjudication of breach.[110] In this regard, the analytical distinction between rules and standards is also unlikely to significantly affect compliance levels in developing countries. Like developed countries, developing countries may obtain more information from a rule. However, the cost of obtaining information under a standard may be so high for developing countries that, for all practical purposes, the developing country will simply treat the standard as a rule and modifies its behavior to conform to the strictest possible interpretation of the standard – thereby, de facto, converting the standard into a rule. Insulation from domestic interest groups, coupled with the costs of obtaining information regarding a TRIPS standard, suggest that voluntary compliance levels in some developing and (eventually) least-developed countries may be higher than in developed countries. In terms of amending domestic legislation and circumscribing judicial discretion in interpreting the TRIPS Agreement, both formal

and substantive compliance with TRIPS may be greater in developing countries than previously anticipated.

A final observation should be made about voluntary compliance with TRIPS in developing countries. Interest groups representing intellectual property owners have been extremely active in monitoring and aiding developing countries in implementing the provisions of the TRIPS Agreement. There is some concern that these activities will likely contribute to maximalist interpretations of TRIPS standards *and* rules. For example, some developing countries have eschewed the minimal copyright duration requirement of the TRIPS Agreement (life of the author plus 50 years) and adopted the maximalist copyright term of life of the author plus 70 years,[111] which is the favored duration term in developed countries. Other developing countries have simply imported language from the TRIPS Agreement directly into their domestic legislation.[112] In the context of dispute resolution, this direct appropriation entails the negotiating context of the particular provision that may not be consistent with the country's priorities or interests. Because the "grey areas" in TRIPS, in theory, give some room for countries to consider domestic priorities, this overinvestment in compliance is, in effect, the least desirable outcome for developing countries.

The combined effect of a lack of resources to pay in the event of a determination of breach, the unwillingness to incur the costs of dispute settlement, and the possibility that conflict with a developed country may impair other interests (such as foreign economic aid and foreign direct investment) may be sufficiently heavy to deter many less powerful countries from breach. Indeed, if the level of risk aversion is so high that developing countries are unwilling to go through the DSU process, even when they have a legitimate claim or defense, there will be negative welfare returns, both domestically and internationally, for intellectual property regulation. Developing countries may comply by preemptively implementing higher standards of intellectual property protection than the TRIPS Agreement requires. Thus, the more serious problem, where developing countries are concerned, may not be compliance, but the need for incentives for these countries to use the DSU to defend legitimate implementation of TRIPS provisions.

If developing countries are likely to effect maximalist interpretations of the TRIPS provisions, then the role of the DSU in *improving* compliance at the level of domestic legislation will be negligible. Instead, the DSU will have the most effect on compliance efforts in *developed* countries. As I have argued elsewhere, the DSU process provides an opportunity for developed countries to accomplish through international adjudication what they could not accomplish through negotiation, namely greater levels of intellectual property harmonization.[113] In essence, the DSU forces disputing parties back to the bargaining table whenever an allegation of breach is made concerning a standard of the TRIPS Agreement. This renegotiation is an optimal outcome, so long as the parties are able to reach a solution that costs less than the sanctions imposed through the formal adjudication process. In intellectual property terms, renegotiation is optimal when the decision of the WTO panel or Appellate Body would result in a higher level of protection than that which the parties would ultimately settle for. This is particularly the case when parties submit standards to dispute resolution for interpretation. Of course, interpretation through the adjudicatory process "thickens" the Agreement by filling in gaps that existed from the negotiation. Interpretation also potentially elevates the cost

of TRIPS membership in terms of what such interpretations may require regarding changes to domestic legislation. In this event, the equilibrium established by the terms of the TRIPS Agreement may evolve into a series of multiple equilibria.

There are clearly costs (over and above administrative ones) associated with a mandatory system of adjudication for disputes between sovereign states. These costs, typically domestic political considerations, specifically influence the willingness of governments to comply with their international obligations.[114] Of course, domestic politics may actually serve to enhance compliance.[115] This is certainly the case with the TRIPS Agreement in developed countries, particularly the United States. The effect of domestic politics on international affairs is particularly acute in the case of strong democracies, and this effect is not predictably unidimensional. Changes in the domestic political landscape often influence how current leaders feel and act with respect to international agreements that are already in force. If dispute settlement outcomes are more affected by domestic conditions and international relations than by the existence of the DSU system, did WTO Member countries overinvest in negotiating and designing such a radical international dispute settlement system? Or, in other words, does the DSU affect levels of *voluntary* compliance?

Almost every breach of the WTO Agreements can be classified as a resort to protectionism – usually in response to domestic pressures that, if unheeded, might erode domestic political support.[116] In light of this protectionism, the DSU offers opportunities for process opportunism. That is, the DSU may serve to facilitate strategies that mediate between the internal (domestic) pressures that political actors face, and the external (international) obligations to which they are committed.[117] In this sense, the DSU process may serve as a channeling mechanism, diverting domestic pressure to the international setting and thus insulating political leaders from some of the political costs of compliance.[118] Voluntary compliance, as defined in this chapter, is enhanced by invoking the DSU process where domestic interests adversely affect a Member's incentive to comply. The formal process exerts strong moral pressure on Members, heightens the reputational costs of violation by publicizing the allegation of breach to the entire trading community, and opens the possibility of cross-sectoral retaliation in the event that a panel determines that a violation does exist.

5 The Age of Law

Almost immediately after the TRIPS Agreement came into effect, the United States successfully invoked the DSU process against India.[119] Commentators hailed the outcome as evidence of the effectiveness of the DSU system and the actualized promise of multilateral dispute settlement untainted by the exercise of power in effecting a particular outcome. It was an age of law in which the DSU rules would equally apply to all nations under a structured institutional system, with defined procedures for dispute settlement administered by impartial arbiters. The rules and procedures of the DSU, however, remain susceptible to the disparities of power that characterized dispute settlement under the old GATT system. The DSU process provides parameters that facilitate new bargains between disputing countries. However, this is an opportunity limited to

those countries with the resources to do so. Power in this sense is not the ability to enforce one's will on a weaker Member, or to refuse to comply with a panel's decision. Power, in this new era of rules, is manifested in the opportunities to buy out of the system entirely. Power is further manifested in the invulnerability of certain Members to the costs of involuntary adjudication, because reputation costs are minimal and there exist opportunities to bargain around adjudicated outcomes.

The distinction between rules and standards facilitates efficient and rational decision-making, because rules convey (more) information that helps parties determine whether to change their behavior. If, however, the costs of obtaining information under a rule and a standard are equal, the rule/standard distinction has no direct bearing on compliance.[120] As stated earlier, most of the TRIPS Agreement "standards" are the result of disagreement during the negotiations about what the content of the rule should be. As a matter of international law, then, states are not under an obligation to implement these standards in any predetermined way. Scholars, in particular Jerome Reichman, have urged developing countries to assume a "pro-competitive" model of intellectual property protection by exploiting the "grey" areas of the TRIPS Agreement to craft domestic legislation that is calibrated to provide a balance between incentives to create and welfare-generating limitations on the proprietary rights of owners.[121] However, where domestic legislation implements TRIPS standards in a manner that explicitly exploits the ambiguities, countries must be prepared to defend challenges to the legitimacy of their domestic rules. Developing countries should be given incentives and means to do so. In addition, Kal Raustiala's commendation of nonbinding instruments and systems implementation review (SIRs), as a mechanism for enhancing cooperative compliance and ascertaining the requirements of an obligation *prior* to a dispute arising, should be seriously considered by Members.[122] For example, a move to give the TRIPS Council power to issue authoritative interpretations of ambiguous provisions (somewhat akin to declaratory judgment actions in common law systems) would help solve the problem of costly acquisition of information, which I identified in Section 2, and thus may forestall a number of disputes by providing an environment in which common understanding of a TRIPS provision may be developed.[123]

6 Conclusion

This chapter offers a consideration of the different ways the DSU has altered the fabric of international cooperation, particularly in the area of international intellectual property regulation. My analysis here is by no means exhaustive; my primary goal is to consider, in a preliminary fashion, the effect of the DSU on the conventional justifications for compliance in international law. Mandatory adjudication, with built-in arbitration options, adds a unique aspect to the study of this dynamic and multifaceted issue. The highly political nature of compliance with international law, with its attendant costs, suggests that the equilibrium for TRIPS dispute settlement should be characterized by infrequent use. Put differently, efficient use of the entire DSU process should be characterized by the possibility of net global welfare gains that outweigh the political costs of mandatory adjudication. My initial conclusions echo the concerns of those who anticipate that an overly aggressive use of confrontational tactics will fail to

produce results that are consistent with the objectives of the WTO Agreements and global welfare in general.[124] This is because developing countries are not legally obligated (but may be coerced) to invest more than is necessary or efficient in TRIPS enforcement at the domestic level, and because developed countries are more likely to bargain around adjudicated disputes than comply with WTO panel outcomes. Neither strategy is beneficial for the vitality of the TRIPS Agreement or the credibility of the DSU.

The signals communicated at each stage of the DSU process may aid countries in making optimal decisions about compliance, and rational responses to these signals will help secure the welfare gains. Such responses may also create opportunities for interdependence. Indeed, an idea that deserves thoughtful consideration is the possibility that the DSU process is structured to facilitate greater interdependence between countries by establishing a framework where cooperation can take place through information exchanges. As I conclude in the second section, greater levels of interdependence heighten the prospects for voluntary compliance.

Where involuntary compliance results in the payment of sanctions rather than modified behavior, it is likely that over time the equilibrium point for global intellectual property protection will allow deviations from TRIPS provisions that can be justified by domestic policies/politics. This may have the ironic result of enhancing net global welfare, particularly if the policy involves limitations on proprietary rights. It may also, however, raise the costs of determining exactly what constitutes compliance with the TRIPS Agreement and, thus, effect lower levels of voluntary compliance with TRIPS standards, particularly in *developed* countries where sectoral interests affect the construction of exceptions and limitations to intellectual property rights.

Many variables affect how countries are likely to interact with the signals of the DSU process. None of these variables is fixed, or easily measurable, thus making it impractical to accurately predict the influence of the DSU on voluntary compliance. What seems predictable, however, is that in the event of a determination of breach, the rich will comply or negotiate around sanctions for the alleged violations. The poor will simply comply and, possibly, overinvest in such compliance.

Acknowledgments

I am grateful to Graeme Dinwoodie and Uche Ewelukwa for reading and making helpful comments on earlier drafts. Jeb Boatman, Liz Mazzarella, and Julie Short provided valuable research assistance. As always, Tade Okediji read and commented on several drafts, and listened patiently as I worked my way through many of the ideas discussed in this chapter.

Notes

1 Agreement Establishing the World Trade Organization, Understanding on Rules and Procedures Governing the Settlement of Disputes, April 15, 1994, 33 I.L.M. 1226 (1994) [hereinafter DSU].

2 Agreement Establishing the World Trade Organization, Agreement on Trade-Related Aspects of Intellectual Property Rights, April 15, 1994, 33 I.L.M. 1197 (1994) [hereinafter

TRIPS Agreement]. See Ruth Okediji, *TRIPS Dispute Settlement and the Sources of (International) Copyright Law*, 49 J. Copyright Soc'y 585, 593 n.21 (2001) (discussing the ICJ as a forum for adjudicating intellectual property disputes).

3 19 U.S.C. §§ 2242, 2411–20 (2002).

4 See, e.g., Judith H. Bello and Alan F. Holmer, *Dispute Resolution in the New World Trade Organization: Concerns and Net Benefits*, 28 Int'l Law 1095 (1994); Rochelle Cooper Dreyfuss and Andreas F. Lowenfeld, *Two Achievements of the Uruguay Round: Putting TRIPS and Dispute Resolution Together*, 37 Va. J. Int'l L. 275 (1997); John H. Jackson, *The World Trading System: Law and Policy of International Economic Relations* 124–7, 133–7 (2nd ed. 1997).

5 Professor John Jackson has contended that the DSU does not allow a choice between compliance and damages. But see Judith Hippler Bello, *The WTO Dispute Settlement Understanding: Less is More*, 90 Am. J. Int'l L. 416 (1996) (proposing that compensation in lieu of compliance with a WTO dispute panel report is consistent with the DSU). Professor Jackson argues that "the DSU clearly establishes a preference for an obligation to perform," notwithstanding the weakness/lack of international enforcement mechanisms. See John H. Jackson, *The WTO Dispute Settlement Understanding – Misunderstandings on the Nature of Legal Obligation*, 91 Am. J. Int'l L. 60, 63 (1997) (stating that a panel decision is a binding obligation under international law). See also, Jackson, *supra* note 4, at 126–7. For a view consistent with Bello's, see Warren F. Schwartz and Alan O. Sykes, *The Economic Structure of Renegotiation and Dispute Resolution in the WTO/GATT System*, John M. Olin Working Paper No. 183 (2nd series) Law & Economics, 15–24 (2002) (unpublished manuscript, on file with the University of Chicago Law Review) (arguing that the DSU rules permit a Member State to deviate from its obligations in order to facilitate efficient breach. They analyze the DSU structure as a liability-rule regime and argue that the sanctions possible under the DSU suggest a "careful design" to allow states to adjust to circumstances where breach of an obligation maximizes the joint gains of all the actors).

6 I use the term "options" here to denote the possible outcomes of a dispute, and not to suggest that countries may legitimately deviate from the obligations in the covered Agreements.

7 As I discuss in Section 3, the number of WTO Members and the multiple commitments among members on different subjects outside of the trade arena engender the formation of coalitions. Coalition dynamics, in turn, facilitate tactics rooted in power that invisibly undermine the formal equality of states and the integrity of the rule-oriented system. See, e.g., Richard H. Steinberg, *In the Shadow of Law or Power? Consensus-Based Bargaining and Outcomes in the GATT/WTO*, 56 Int'l Org. 339 (Spring 2002) (examining how, despite the consensus-based rules of the GATT/WTO, countries may still use power to affect outcomes).

8 See in particular J. H. Reichman, *Enforcing the Enforcement Procedures of the TRIPS Agreement*, 37 Va. J. Int'l L. 335 (1997); J. H. Reichman and David Lange, *Bargaining Around the TRIPS Agreement: The Case for Ongoing Public–Private Initiatives to Facilitate Worldwide Intellectual Property Transactions*, 9 Duke J. Comp. & Int'l L. 11 (1998); J. H. Reichman, *The TRIPS Agreement Comes of Age: Conflict or Cooperation with the Developing Countries?* 32 Case W. Res. J. Int'l L. 441, 457–63 (2000); Martine de Konig, *Why the Coercion-Based GATT Approach is Not the Only Answer to Piracy in the Asia-Pacific Region*, 19 Eur. Intell. Prop. Rev. 59 (1997).

9 Another example is the United Nations Convention on the Law of the Sea (LOS Convention) opened for signature, December 10, 1982, 1833 U.N.T.S. 397, and the attendant International Law of the Sea Tribunal, which is one of four fora to which parties to the Convention may submit disputes.

10 See Statute of the International Court of Justice, June 26, 1949, 59 Stat. 1055, T.S. No. 993, 3 Bevans 1179.

11 Article 94 of the UN Charter provides that the Security Council may make recommendations or take measures necessary to give effect to a judgment of the ICJ. See UN Charter, June 26, 1945, 59 Stat. 1031, T.S. No. 993, 3 Bevans 1153.

12 Professor Henkin has famously stated "almost all nations observe almost all principles of international law and almost all of their obligations almost all of the time." Louis Henkin, *How Nations Behave* 47 (2nd edn. 1979). See also Abram Chayes and Antonia Handler Chayes, *The New Sovereignty: Compliance with International Regulatory Agreements* 3 (1995) (noting the assumption that states have a propensity to obey international law); Lori F. Damrosch, *Enforcing International Law Through Non-Forcible Measures* 269 Rec. Des Cours 9, 19–24 (1997) (discussing various ways that international law is enforced and stating that "there is much more voluntary compliance with international law than critics would acknowledge").

13 Kenneth N. Waltz, *Theory of International Politics* (1979); Robert O. Keohane, *After Hegemony* (1984). See also, Kal Raustiala, *Compliance and Effectiveness in International Regulatory Cooperation*, 32 Case W. Res. J. Int'l L. 387, 399–411 (2000) (providing a summary of the major theories of treaty compliance developed by political scientists).

14 Stephen D. Krasner, *Structural Causes and Regime Consequences: Regimes as Intervening Variables*, 36 Int'l Org. 185, 187 (1982); Robert Keohane, *The Demand for International Regimes*, 36 Int'l Org. 325, 330–1 (1982).

15 Robert O. Keohane and Joseph S. Nye, *Power and Interdependence: World Politics in Transition* 29–30 (1977).

16 Okediji, *supra* note 2.

17 Waltz, *supra* note 13; Keohane, *supra* note 13.

18 See, e.g., Beth A. Simmons, *Money and the Law: Why Comply with the Public International Law of Money?*, 25 Yale J. Int'l L. 323, 326 (2000); Am. Soc'y of Int'l Law, *Proceedings of the 91st Annual Meeting: Implementation, Compliance and Effectiveness* (1997).

19 This convergence was particularly evident during TRIPS negotiations as well as the two WIPO Internet treaties. See, e.g., Pam Samuelson, *The US Digital Agenda at WIPO*, 37 Va. J. Int'l L. 369 (1997) (describing negotiations for the WIPO Copyright Treaty and the WIPO Performances and Phonograms Treaty).

20 See Detlev F. Vagts, *The United States and Its Treaties: Observance and Breach*, 95 Am. J. Int'l Law 313, 324–9 (2001) [discussing five reasons for US compliance with treaties: (1) honor; (2) contract theory; (3) natural law; (4) promise, will, and consent theories; and (5) national interest, which encompasses concerns about reputation].

21 Keohane, *supra* note 14; Krasner, *supra* note 14. See also Oran R. Young, *Regime Dynamics: The Rise and Fall of International Regimes*, 36 Int'l Org. 277 (1982).

22 DSU, *supra* note 1, art. 3(7) (emphasis added). See also art. 3(10) (providing that "use of the dispute settlement procedures should not be "intended or considered as contentious acts and that, if disputes arise, all Members will engage in these procedures in good faith in an effort to resolve the dispute"). In a recent decision involving a dispute between the European Union and the United States over the Foreign Sales Corporation Tax Law, an arbitrator held that the EU is authorized to engage in countermeasures in the amount of $4,043 million against the United States. See *United States – Tax Treatment for "Foreign Sales Corporations*," WTO Doc. WT/DS108/ARB, August 30, 2002. Director-General Mike Moore issued a press release stressing the need for both parties to resolve the dispute

cooperatively:

The arbitration ruling marks a further stage of WTO involvement in this long-running and difficult dispute. I have been following the FSC case closely and I commend both parties for working in a constructive manner throughout the duration of this dispute. I urge both parties to continue to cooperate and work toward resolving this dispute and the others between them in an amicable and constructive fashion. The European Union and the United States are among the most important members of this organization and both hold a special responsibility to ensure the continued health and soundness of WTO and global trading system.

See Press Release 311, available at http://www.wto.org/english/news_e/pres02_e/pr311_e.htm (August 30, 2002).

23 Joseph E. Stiglitz, *Information and the Change in the Paradigm in Economics*, 92 Am. Econ. Rev. 460, 461 (2002) ("Information affects decision-making in every context – not just inside firms and households").

24 *Id.* at 472 (noting Akerlof's point that thin or absent markets are the result of information asymmetries).

25 Of course, the parties may design these signals and their responses to hide or distort information. Using actions to convey information generates costs that lead to market imperfections. Stiglitz, *supra* note 23, at 473.

26 Okediji, *supra* note 2, at 612 (discussing the European Union claim against the United States based on Section 110(5) of the Copyright Act and explaining the decision by both countries to opt out of the process as one that reflects, in part, their mutual cooperative relationship in the area of intellectual property, particularly with respect to influencing compliance with TRIPS by developing countries).

27 United States – Section 110(5) of the U.S. Copyright Act, WTO Doc. WT/DS160/R, June 15, 2000.

28 David Vogel and Timothy Kessler, *How Compliance Happens and Doesn't Happen Domestically*, in *Engaging Countries: Strengthening Compliance with International Environmental Accords 19* (Edith Brown Weiss and Harold K. Jacobson eds., 1998) (stating that the "division of labor between institutions that make environmental policy and those that are responsible for implementing and enforcing it" has led to a "persistent gap ... between the standards established by national regulations and the behavior of governmental institutions").

29 Helen V. Milner, *Interests, Institutions and Information: Domestic Politics and International Relations* 9–17, 34–46 (1997) (exploring how the structure of domestic politics, i.e., the policy preferences of political actors affect international cooperation). See also, Ronald J. Herring and Erach Bharucha, *Embedded Capacities: India's Compliance with International Environmental Accords*, in *Engaging Countries: Strengthening Compliance with International Environmental Accords 395* (Edith Brown Weiss and Harold K. Jacobson eds., 1998) (concluding that India will more effectively comply with environmental treaties that are consistent with its national policies).

30 Milner, *supra* note 29.

31 It is typical that prior to the formal consultations stage, the parties have unsuccessfully employed diplomatic efforts to resolve the dispute.

32 DSU, *supra* note 1, art. 5(6).

33 See India – Patent Protection for Pharmaceutical and Agricultural Chemical Products, AB-1997–5, WT/DS50/AB/R, ¶ 94 (December 19, 1997, adopted January 16, 1998).

34 DSU, *supra* note 1, art. 4(6).

35 See David Palmeter and Petros C. Mavroidis, *Dispute Settlement in the World Trade Organization, Practice and Procedure* 64 (1999).

36 The number of third parties who join the complainant may further strengthen the signal sent by the request for consultations. See DSU, *supra* note 1, art. 10 (governing third parties). There is always a risk that third parties may become multiple complainants, thus requiring the violating party to remedy the violation or agree to sanctions/withdrawal of concessions from more than one Member. See DSU, *supra* note 1, art. 9 (procedures for multiple complainants).

37 See DSU, *supra* note 1, art. 6

38 See DSU, *supra* note 1, art. 3(9).

39 See DSU, *supra* note 1, art. 16, art. 17(4).

40 Stiglitz, *supra* note 23, at 463.

41 This may explain, in part, why the European Union was willing to challenge Section 301 of the U.S. Trade Act before a WTO Panel. This was a rule of general application – everyone suffered under this formidable U.S. mechanism. But only the European Union was willing to incur the costs of acquiring information about its legitimacy under the new WTO system. See United States – Sections 301–10 of the Trade Act of 1974, WTO Doc. DS/152/R, November 30, 1998.

42 See Robert D. Putnam, *Diplomacy and Domestic Politics: The Logic of Two-Level Games*, 42 Int'l Org. 427 (1988) (using game theory to explain the role of domestic politics in the making of international agreements); Milner, *supra* note 29, 42–3 (noting the "standard assumptions" about how concerns about preserving domestic political office influence the choices of political actors regarding international cooperation); Herring & Bharucha, *supra* note 29 ("the real test of even soft international law comes when domestic costs become a matter of political process; obstacles to compliance that were invisible under the rarified negotiation process merge as it becomes clear whose ox is to be gored. Capacity to comply is embedded in the state's relation to the society it claims to represent and in positional imperatives of the international system").

43 See Kimberly A. Czub, *Argentina's Emerging Standard of Intellectual Property Protection: A Case Study of the Underlying Conflicts Between Developing Countries, TRIPS Standards, and the United States*, 33 Case W. Res. J. Int'l L. 191, 212–14 (2001) (discussing the influence of the domestic pharmaceutical industry on the Argentine legislature).

44 17 U.S.C. § 110(5) (2002).

45 See DSU, *supra* note 1, art. 16.

46 Award of the Arbitrator, United States – Section 110(5) of the U.S. Copyright Act, WTO Doc. WT/DS160/12, ¶ 36, January 15, 2001.

47 *Id.* at ¶¶ 41 and 42.

48 See Canada – Term of Patent Protection, WTO Doc. WT/DS170/R, May 5, 2000.

49 See Award of the Arbitrator, Canada – Term of Patent Protection, WTO Doc. WT/DS170/10, ¶ 60, February 28, 2001; Award of the Arbitrator, United States – Section 110(5) of the US Copyright Act, *supra* note 46, ¶ 42. See also, Award of the Arbitrator, Canada – Patent Protection of Pharmaceutical Patents, WTO Doc.WT/DS114/13, ¶ 60, August 18, 2000.

50 There has been some extensive review and analysis of the Section 110(5) dispute. See Laurence R. Helfer, *World Music on a U.S. Stage: A Berne/TRIPS and Economic Analysis of the Fairness in Music Licensing Act*, 80 B.U. L. Rev. 93 (2000). Ruth Okediji, *Toward an International Fair Use Standard*, 39(1) Colum. J. Transnat'l L. 75 (2000); Jane C. Ginsburg, *Toward Supranational Copyright Law? The WTO Panel Decision and the*

"Three-Step Test" for Copyright Exceptions, 187 Revue Internationale du Droit d'Auteur 3 (2001); Graeme Dinwoodie, *The Development and Incorporation of International Norms in the Formation of Copyright Law*, 62 Ohio St. L. J. 733 (2001).

51 DSU, *supra* note 1, art.17(5).

52 Of course, an appeal may also facilitate renegotiation. See Section 4, *infra*, discussing side-payments and renegotiation.

53 Report of the Panel, India – Patent Protection for Pharmaceutical and Agricultural Chemical Products, WT/DS50/R (1997). Another illustrative case is the dispute between the European Union and the United States in United States – Section 211 Omnibus Appropriations Act of 1998, WTO Doc. WT/DS176/R August 6, 2001.

54 *Id.* at ¶¶ 7.29–7.41 (1997).

55 See Vienna Convention on the Law of Treaties, May 23, 1969, Art. 31(1), 1155 U.N.T.S. 331 [hereinafter Vienna Convention].

56 Appellate Body Report, India – Patent Protection for Pharmaceutical and Agricultural Chemical Products, WT/DS50/AB/R (1997).

57 Jerome H. Reichman, *Securing Compliance with the TRIPS Agreement After US* v. *India*, 1 J. Int'l Econ. L. 585, 594–7 (1998) [hereinafter *Securing Compliance*].

58 *Id.*

59 Reichman, *Securing Compliance*, *supra* note 57, at 596.

60 See, e.g., Louis Kaplow, *Rules versus Standards: An Economic Analysis*, 42 Duke L. J. 557 (1992); Cass R. Sunstein, *Problems with Rules*, 83 Cal L. Rev. 953 (1995).

61 This good faith standard is embodied in the principle "*pacta sunct servanda*," meaning treaties must be observed. See Vienna Convention, Art. 26 ("Every treaty in force is binding upon the parties to it and must be performed by them in good faith.") and Art. 31 (imposing an obligation on states to interpret treaties in good faith); See also U.C.C. § 1–203 (1998) ("Every contract or duty within this Act imposes an obligation of good faith in its performance or enforcement."); Restatement (Second) of Contracts § 205 (1981) ("Every contract imposes upon each party a duty of good faith and fair dealing in its performance and its enforcement").

62 Kaplow, *supra* note 60.

63 For example, a rule in a contract that the plaintiff will buy "chicken" may require interpretation as to what constitutes a "chicken." Is it a fryer, a broiler, or a hen? A court considers several things to determine exactly what the parties meant by the term "chicken," including preliminary negotiations, trade usage, and course of performance. See *Frigaliment Importing Co.* v. *B.N.S. Int'l Sales Corp.*, 190 F. Supp. 116 (S.D.N.Y. 1960). See also *Nanakuli Paving & Rock Co.* v. *Shell Oil Co.*, 664 F.2d 772 (9th Cir. 1981) (determining that trade usage could override an explicit contractual provision).

64 Kaplow, *supra* note 60, 559–96.

65 *Id.*

66 Joel Trachtman, *The Domain of WTO Dispute Resolution*, 40 Harv. Int'l L.J. 333, 355–76 (1999) (applying rules versus standards analysis to the international trade context).

67 Gunnar Sjostedt, *Negotiating the Uruguay Round of the General Agreement on Tariffs and Trade*, in *International Multilateral Negotiation*, 44, 62–3, 68–9 (I. William Zartman ed., 1994).

68 Daniel Gervais, *The TRIPS Agreement: Drafting History and Analysis* 10–25 (1998) (providing a brief review of the events leading to the final text of the TRIPS Agreement). For a comprehensive account of the TRIPS negotiations see Gail E. Evans, *Intellectual Property as a Trade Issue – The Making of the Agreement on Trade-Related Aspects of Intellectual Property Rights*, World Competition L. & Econ. Rev. 137 (1994). See

also Ruth Okediji, *A Cartography of WTO TRIPS Dispute Settlement and the Future of Intellectual Property Policy*, Working Paper (2001) (describing the coalitional dynamics during the TRIPS negotiations); Christophe Dupont, *Coalition Theory: Using Power to Build Cooperation*, in *International Multilateral Negotiation* 148, 164 (I. William Zartman ed., 1994) (table of coalitions formed on various issues during the Uruguay Round).

69 Okediji, *supra* note 68. See also Putnam, *supra* note 42 (describing similar coalitional action with the 1978 Bonn Summit).

70 See e.g., Gervais, *supra* note 68, at 82 [commenting on an ambiguity with respect to Art. 10(2)], 84–5 (commenting on the failure to agree on a general right of distribution and importation and the resulting standard embodied in Art. 11).

71 Dupont, *supra* note 68; Okediji, *supra* note 68; See also, Michael P. Ryan, *The Function-Specific and Linkage-Bargain Diplomacy of International Intellectual Property Lawmaking*, 19 U. Pa. J. Int'l Econ. L. 535 (1998).

72 Trachtman, *supra* note 66, at 376–7 [concluding that "it is impossible to determine without specific analysis the utility of a rule versus a standard in particular (international) cases"].

73 Specific examples would include human rights treaties and environmental treaties. These treaties and associated practices or expectations that evoke compliance issues are often described as "soft law." For a collection of essays examining the use and effectiveness of soft law, see *Commitment and Compliance: The Role of Non-Binding Norms in the International Legal System* (D. Shelton ed., 2000).

74 Berne Convention for the Protection of Literary and Artistic Works, last revised at Paris, July 24, 1971, S. Treaty Doc. No. 99-27, 99th Cong. 2nd Sess. 37 (1986), 828 U.N.T.S. 221 [hereinafter Berne Convention].

75 For the leading account of the progressive development of the Berne Convention, see Sam Ricketson, *The Berne Convention for the Protection of Literary and Artistic Works*, 1886–986 (1987).

76 See, in particular, the patent provisions of the TRIPS Agreement, which reflect an evolution in rule-making as compared to the Paris Convention for the Protection of Literary and Artistic Works. See TRIPS Agreement, *supra* note 2, section 5.

77 A power-based view of the WTO suggests that compliance is simply a choice by those who can afford (or are largely unaffected by) the consequences of a violation.

78 Schwartz and Sykes, *supra* note 5.

79 See Jack L. Goldsmith, *Sovereignty, International Relations Theory and International Law*, 52 Stan. L. Rev. 959, 985 (2000) (skeptical that loss of reputation has any significant effect on the behavior of states); Robert O. Keohane, *International Relations and International Law: Two Optics*, 38 Harv. Int'l L.J. 487 (1997) (same). But see Schwartz and Sykes, *supra* note 5, at 20–3 (arguing that reputation plays an important, if imperfect, role in compliance with trade rules).

80 See, e.g., United States – Section 110(5) of the Copyright Act, *supra* note 27; Canada – Patent Protection of Pharmaceutical Products, *infra* note 94.

81 Czub, *supra* note 43, at 223–5 (discussing confusion in Argentina over application of the TRIPS rule for length of patent protection). See also, Canada – Term of Patent Protection, *supra* note 48 (dispute over the meaning of the TRIPS "rule" that patent term last for 20 years from the date of filing). Professor Raustiala has suggested that ambiguities over what TRIPS obligations require should be resolved informally through mechanisms such as systems implementation review (SIRs) which is typified by the work of the TRIPS Council. See Raustiala, *supra* note 13, at 432–8.

82 But see, Canada – Term of Patent Protection, *supra* note 48.

83 Kaplow, *supra* note 60, at 563–4, 575–7.

84 European Communities – Enforcement of Intellectual Property Rights for Motion Pictures and TV Programmes, WTO Doc. WT/DS125, May 7, 1998; Greece – Enforcement of Intellectual Property Rights for Motion Pictures and TV Programmes, WTO Doc. WT/DS124, May 7, 1998.
85 Denmark – Measures Affecting the Grant of Copyright and Neighbouring Rights, WTO Doc. WT/DS83, May 21, 1997; Sweden – Measures Affecting the Enforcement of Intellectual Property Rights, WTO Doc. WT/DS86, June 2, 1997.
86 Portugal – Patent Protection Under the Industrial Property Act, WTO Doc. WT/DS37, May 6, 1996.
87 Japan – Measures Concerning Sound Recordings, WTO Doc. WT/DS42, June 4, 1996.
88 Japan – Measures Concerning the Protection of Sound Recordings, WTO Doc. WT/DS28, February 14, 1996.
89 Pakistan – Patent Protection for Pharmaceutical and Agricultural Chemical Products, WTO Doc. WT/DS36, May 6, 1996.
90 Rufus H. Yerxa and Demetrios J. Marantis, *Assessing the New WTO Dispute System: A US Perspective*, 32 Int'l L. 795, 808–9 (1998) (citing a statement by the U.S. Trade Representative that threats of invoking the DSU process has effected compliance in some developing countries).
91 I am grateful to Professor Dinwoodie for pointing out this distinction.
92 David A. Lax and James K. Sebenius, *Thinking Coalitionally: Party Arithmetic, Process Opportunism, and Strategic Sequencing*, in *Negotiation Analysis* 153, 166 (H. Peyton Young ed., 1991).
93 I discuss this particular strategy at length, using the two-stage game model to analyze the various ways that process opportunism plays out. See Okediji, *supra* note 68.
94 Report of the Panel, Canada – Patent Protection of Pharmaceutical Products, WTO Doc. WT/DS114/R (2000).
95 See Okediji, *supra* note 68, for discussion of this dispute.
96 Professor Trachtman also reaches this conclusion. See Trachtman, *supra* note 66, at 376–7.
97 Kaplow, *supra* note 60, 571–7, 599–601.
98 Reichman, *supra* note 8; Czub, *supra* note 43.
99 *Id.*
100 TRIPS Agreement, *supra* note 2, Art. 41(5).
101 See TRIPS Agreement, *supra* note 2, Art. 41, which provides, *inter alia*, "Members shall ensure that enforcement procedures as specified . . . are available under their law so as to permit effective action against any act of infringement of intellectual property rights . . .."
102 See TRIPS Agreement, *supra* note 2, Art. 41(5).
103 Member States are required to notify all changes to intellectual property laws to the TRIPS Council. See TRIPS Agreement, *supra* note 2, Art. 63.
104 Kaplow, *supra* note 60.
105 See *supra* note 12. Recently, there has been a move to integrate international law and international relations in an effort to engage in more robust analysis about the role and efficacy of rules and institutions in international society. See, e.g., Kenneth W. Abbott, *Modern International Relations Theory: A Prospectus for International Lawyers*, 4 Yale J. Int'l L. 335 (1989); Anne-Marie Slaughter Burley, *International Relations Theory: A Dual Agenda*, 87 Am. J. Int'l L. 205 (1993).
106 See Chayes and Chayes, *supra* note 12; Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 Yale L.J. 2599 (1997) (book review); Raustiala, *supra* note 13.
107 Investment in TRIPS compliance in developed countries, and in the more advanced developing countries such as India and Brazil, will be considerably affected by domestic considerations that the structure of the DSU is ill-suited to change.

108 See Putnam, *supra* note 42, at 427–9 (1988) (giving some practical examples of domestic/international linkages); Lax and Sebenius, *supra* note 92.

109 Putnam, *supra* note 42, at 448–50 (discussing the role of political institutions in bargaining strengths of political actors).

110 Note, however, that an Advisory Centre for WTO Law was established in 2001. This Centre is intended to assist developing countries in dealing with WTO related disputes. See www.wto.org. With respect to the economic costs of dispute resolution, it is not yet apparent how this Centre will affect the willingness and ability of developing countries to go through the DSU process.

111 See, e.g., Copyright Laws of Argentina, Brazil, Hungary, and Poland. India has adopted a term of life plus 60 years. See LexisNexis.com, International Copyright Law and Practice (2001). Draft laws in Nigeria and Botswana also propose a copyright term of life plus 70.

112 I have observed this trend particularly in the draft laws of countries in sub-Saharan Africa.

113 Okediji, *supra* note 68.

114 Helen V. Milner, *Resisting Protectionism: Global Industries and the Politics of International Trade* 290–1 (1998); Andrew Moravscik, *Taking Preferences Seriously: A Liberal Theory of International Politics*, 51 Int'l Org. 513, 537 (1997).

115 Schwartz and Sykes, *supra* note 5, at 18–20; Moravscik, *supra.*

116 The recent decision by President George W. Bush to impose tariffs on steel is such an example. See Dick Polman, *Bush's Steel Tariff is an Early Bit of Politicking*, *Philadelphia Inquirer*, March 13, 2002, at A17 (stating that President Bush made the decision to impose steel tariffs to gain critical votes in key midwestern states); Mike Allen and Steven Pearlstein, *Bush Settles on Tariff for Steel Imports*, available at Washingtonpost.com, March 5, 2002 (making the same connection between domestic political gain and the decision to violate trade rules). See generally Kal Raustiala, *Domestic Institutions and International Regulatory Cooperation: Comparative Responses to the Convention on Biological Diversity*, 49 World Pol. 482, 482–3 (1997).

117 See Moravscik, *supra* note 114, at 516–24 (1997); Lax and Sebenius, *supra* note 92 at 166–73.

118 See Ruth Okediji, *Toward an International Fair Use Standard*, 39 Colum. J. of Transnat'l Law 75, 85–7 (2000) (discussing these strategies with regard to the fair use standard and its consistency with the TRIPS Agreement).

119 India – Patent Protection for Pharmaceutical and Agricultural Chemical Products, Appellate Body Report, WTO Doc. WT/D550/AB/R, December 19, 1997.

120 Kaplow, *supra* note 60, 572–4 (where the cost of acquiring information is the same for rules and standards, the choice of which one to enact should depend on other factors such as costs of promulgation, frequency of behavior that is the subject of regulation, etc.).

121 Reichman, *Securing Compliance, supra* note 57, at 589–92. See also J. H. Reichman, *From Free Riders to Fair Followers: Global Competition Under the TRIPS Agreement*, 29 NYU J. Int'l L. & Pol. 11 (1996/97). Professor Reichman suggests that a pro-competitive strategy of compliance with TRIPS includes tilting domestic intellectual property laws to favor second-comers and resisting elevation of intellectual property beyond what the TRIPS Agreement requires. *Id.* See also World Bank, Global Economic Prospects 2002: Making Trade Work for the World's Poor 129, 147 (2001) (recommending lower and more flexible intellectual property standards as a policy option for developing countries).

122 Raustiala, *supra* note 13, at 432–8.

123 *Id.*

124 *Id.* at 438–9; Reichman, *Securing Compliance*, *supra* note 57, at 456–63; Reichman and Lange, *supra* note 8; Martine de Konig, *supra* note 8.

References

Abbott, Kenneth W. 1989: Modern international relations theory: a prospectus for international lawyers. *Yale J. Int'l L.*, 14, 335–411.

Agreement establishing the World Trade Organization, Agreement on Trade-Related Aspects of Intellectual Property Rights, April 15, 1994, 33 I.L.M. 1997.

Agreement establishing the World Trade Organization, Understanding on rules and procedures governing the settlement of disputes, April 15, 1994, 33 I.L.M. 1226.

Allen, Mike and Steven Pearlstein 2002: Bush settles on tariff for steel imports, http://www.Washingtonpost.com.

Am. Soc'y of Int'l Law 1997: *Proceedings of the 91st Annual Meeting: Implementation, Compliance and Effectiveness*.

Bello, Judith Hippler 1996: The WTO dispute settlement understanding: Less is more. *Am. J. Int'l L.*, 90, 416–18.

Bello, Judith H. and Alan F. Holmer 1994: Dispute resolution in the new World Trade Organization: concerns and net benefits. *Int'l Law.*, 28, 1095–103.

Berne Convention for the Protection of Literary and Artistic Works, last revised at Paris, July 24, 1971, S. Treaty Doc. No. 99-27, 99th Cong. 2nd Sess. 37 (1986), 828 U.N.T.S. 221.

Burley, Anne-Marie Slaughter 1993: International relations theory: a dual agenda. *Am. J. Int'l L.*, 87, 205–39.

Chayes, Abram and Antonia Handler Chayes 1995: *The New Sovereignty: Compliance with International Regulatory Agreements*. Cambridge, Mass.: Harvard University Press.

Czub, Kimberly A. 2001: Argentina's emerging standard of intellectual property protection: a case study of the underlying conflicts between developing countries, TRIPS standards, and the United States. *Case W. Res. J. Int'l L.*, 33, 191–231.

Damrosch, Lori F. 1997: Enforcing international law through non-forcible measures. *Rec. Des Cours*, 269, 9.

de Konig, Martine 1997: Why the coercion-based GATT approach is not the only answer to piracy in the Asia-Pacific region. *Eur. Intell. Prop. Rev.*, 19, 59.

Dinwoodie, Graeme 2001: The development and incorporation of international norms in the formation of copyright law. *Ohio St. L. J.*, 62, 733–82.

Dreyfuss, Rochelle Cooper and Andreas F. Lowenfeld 1997: Two achievements of the Uruguay Round: putting TRIPS and dispute resolution together. *Va. J. Int'l L.*, 37, 275–333.

Evans, Gail E. 1994: Intellectual property as a trade issue – the making of the agreement on trade-related aspects of intellectual property rights. *World Competition L. & Econ. Rev.*, 18(2), 137–80.

Gervais, Daniel 1998: *TRIPS Agreement: Drafting History and Analysis*. London: Sweet & Maxwell.

Ginsburg, Jane C. 2001: Toward supranational copyright law? The WTO panel decision and the "Three-Step Test" for copyright exceptions. *Revue Internationale du Droit d'Auteur*, 187, 3.

Goldsmith, Jack L. 2000: Sovereignty, international relations theory and international law. *Stan. L. Rev.*, 52, 959–86.

Helfer, Laurence R. 2000: World music on a US stage: a Berne/TRIPS and economic analysis of the fairness in Music Licensing Act. *B.U. L. Rev.*, 80, 93–204.

Henkin, Louis 1979: *How Nations Behave: Law and Foreign Policy*. New York: published for the Council on Foreign Relations by Columbia University Press, 2nd edn.

Hongju Koh, Harold 1997: Why do nations obey international law? *Yale L.J.*, 106, 2599–659.

Jackson, John H. 1997a: The WTO dispute settlement understanding – misunderstandings on the nature of legal obligation. *Am. J. Int'l L.*, 91, 60–4.

Jackson, John H. 1997b: *The World Trading System: Law and Policy of International Economic Relations*, Cambridge MA: MIT Press, 2nd edn.
Kaplow, Louis 1992: Rules versus standards: an economic analysis. *Duke L.J.*, 42, 557–629.
Keohane, Robert 1982: The demand for international regimes. *Int'l Org.*, 36, 325–55.
Keohane, Robert O. 1984: *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, N.J.: Princeton University Press.
Keohane, Robert O. 1997: International relations and international law: two optics. *Harv. Int'l L.J.*, 38, 487–502.
Keohane, Robert O. and Joseph S. Nye 1977: *Power and Interdependence: World Politics in Transition*.
Krasner, Stephen D. 1982: Structural causes and regime consequences: regimes as intervening variables. *Int'l Org.*, 36, 185–205.
Milner, Helen V. 1997: *Interests, Institutions and Information: Domestic Politics and International Relations*. Princeton, N.J.: Princeton University Press.
Milner, Helen V. 1988: *Resisting Protectionism: Global Industries and the Politics of International Trade*. Princeton, N.J.: Princeton University Press.
Moravscik, Andrew 1997: Taking preferences seriously: a liberal theory of international politics. *Int'l Org.*, 51, 513.
Okediji, Ruth 2000: Toward an international fair use standard. *Colum. J. Transnat'l Law*, 39, 75–175.
Okediji, Ruth 2001: TRIPS dispute settlement and the sources of (international) copyright law. *J. Copyright Soc'y*, 49, 585.
Okediji, Ruth 2001: *A Cartography of WTO TRIPS Dispute Settlement and the Future of Intellectual Property Policy*, Working Paper.
Palmeter, David and Petros C. Mavroidis 1999: *Dispute Settlement in the World Trade Organization, Practice and Procedure*, The Hague, Boston: Kluwer Law International.
Polman, Dick 2002: *Bush's Steel Tariff is an Early Bit of Politicking*, Philadelphia Inquirer, at A17.
Press Release 311 2002: http://www.wto.org/english/news_e/pres02_e/pr311_e.htm
Putnam, Robert D. 1988: Diplomacy and domestic politics: the logic of two-level games. *Int'l Org.*, 42, 427–60.
Raustiala, Kal 1997: Domestic institutions and international regulatory cooperation: comparative responses to the convention on biological diversity. *World Pol.*, 49, 482–509.
Reichman, J. H. 1996/97: From free riders to fair followers: global competition under the TRIPS agreement. *NYU J. Int'l L. & Pol.*, 29, 11.
Reichman, J. H. 1997: Enforcing the enforcement procedures of the TRIPS agreement. *Va. J. Int'l L.*, 37, 335–56.
Reichman, J. H. 1998: Securing compliance with the TRIPS agreement after US v. India. *J. Int'l Econ. L.*, 1, 585.
Reichman, J. H. 2000: The TRIPS agreement comes of age: conflict or cooperation with the developing countries? *Case W. Res. J. Int'l L.*, 32, 441–70.
Reichman, J. H. and David Lange 1998: Bargaining around the TRIPS agreement: the case for ongoing public–private initiatives to facilitate worldwide intellectual property transactions. *Duke J. Comp. & Int'l L.*, 9, 11–68.
Ricketson, Sam 1987: *The Berne Convention for the Protection of Literary and Artistic Works*. London: Centre for Commercial Law Studies, Queen Mary College: Kluwer.
Ryan, Michael P. 1998: The function-specific and linkage-bargain diplomacy of international intellectual property lawmaking *U. Pa. J. Int'l Econ. L.*, 19, 535–86.
Samuelson, Pam 1997: The US digital agenda at WIPO. *Va. J. Int'l L.*, 37, 369–439.

Schwartz, Warren F. and Alan O. Sykes 2002: The economic structure of renegotiation and dispute resolution in the WTO/GATT system, John M. Olin Working Paper No. 183, (2nd series) Law and Economics, 15–24.

Shelton, D. (ed.) 2000: *Commitment and Compliance: The Role of Non-Binding Norms in the International Legal System*. New York: Oxford University Press.

Simmons, Beth A. 2000: Money and the law: why comply with the public international law of money? *Yale J. Int'l L.*, 25, 323–62.

Steinberg, Richard H. 2002: In the shadow of law or power? Consensus-based bargaining and outcomes in the GATT/WTO. *Int'l Org.*, 56, 339.

Stiglitz, Joseph E. 2002: Information and the change in the paradigm in economics. *Am. Econ. Rev.*, 92, 460–501.

Sunstein, Cass R. 1995: Problems with rules. *Cal L. Rev.*, 83, 953–1023.

Trachtman, Joel 1999: The domain of WTO dispute resolution. *Harv. Int'l L.J.*, 40, 333–77.

UN Charter, June 26, 1945, 59 Stat. 1031, T.S. No. 993, 3 Bevans 1153.

Vagts, Detlev F. 2001: The United States and its treaties: observance and breach. *Am. J. Int'l Law*, 95, 313–34.

Waltz, Kenneth N. 1979: *Theory of International Politics*. Reading, Mass.: Addison-Wesley Pub. Co.

Weiss, Edith Brown and Harold K. Jacobson (eds.) 1998: *Engaging Countries: Strengthening Compliance with International Environmental Accords*.

World Bank 2002: *Global Economic Prospects 2002: Making Trade Work for the World's Poor*. Washington, D.C.: The International Bank for Reconstruction and Development.

Yerxa, Rufus H. and Demetrios J. Marantis 1998: Assessing the new WTO dispute system: a US perspective. *Int'l Law*, 32, 795–810.

Young, Peyton H. (ed.) 1991: *Negotiation Analysis*. Ann Arbor: University of Michigan Press.

Young, Oran R. 1982: Regime dynamics: the rise and fall of international regimes. *Int'l Org.*, 36, 277–97.

Zartman, I William (ed.) 1994: *International Multilateral Negotiation: Approaches to the Management of Complexity*. San Francisco: Jossey-Bass.

3

Teaching Old Laws New Tricks: The Legal Obligation of Non-Attribution and the Need for Economic Rigor in Injury Analyses Under US Trade Law

James P. Durling and Matthew P. McCullough

Chapter Outline

This chapter explores the issue of "non-attribution" in trade remedy cases from both a legal and an economic perspective. We first describe the legal obligation not to attribute other causes to imports when deciding whether to impose trade remedies such as antidumping (AD) duties or safeguard measures. We then present a recent case study involving cold-rolled carbon steel, showing how economic theory and econometric analysis can meet this legal obligation, and contrast this approach with that traditionally used by the US International Trade Commission. We discuss what happened in the aftermath of this novel use of more formal economics, describing the severe backlash to the decision not to impose AD duties in that case. We then offer some concluding comments about the future role of economics in trade remedy cases.

1 Introduction

Economic analysis often can help sort out conflicting claims by interested parties before a regulatory agency. In many areas of regulatory policy, sophisticated economic analyses have become a widely accepted tool for decision making. Perhaps most dramatically, economic analysis of effects on consumer welfare now largely drives

regulatory and judicial decisions about competition policy. Economic analysis does not dictate a particular policy outcome; rather, it provides an analytically disciplined way to think about whether a certain policy action is consistent in fact with the stated objectives of the relevant statute or regulation. But such disciplined thinking often inconveniences decision makers who prefer the flexibility of "case-by-case" decisions.

Yet, if competition policy has largely embraced economic analysis, trade policy represents the opposite end of the spectrum. Many areas of trade policy seem remarkably resistant to any serious economic analysis. Although competition policy and trade policy are analytically similar, there are important political asymmetries. In competition policy, the usual case has two domestic parties fighting with each other; since both parties are domestic, the politics often are more easily balanced and neutralized, leaving the process free to develop sound analytic techniques over time. In trade policy, in contrast, the usual case involves domestic parties fighting with foreign parties; since the domestic interested parties seeking protectionist outcomes have more political clout than the foreign interested parties seeking free-trade outcomes, politics trumps economics. Moreover, although Congress has mandated consideration of, indeed a primary focus on, consumer interests in competition law cases, Congress insists that trade law cases focus more narrowly on producer interests. Since the narrow producer interests always have an intense interest in obtaining protection, and the diffuse consumer groups usually have only a weak interest in blocking the protection, the politics naturally follow the degree of effort exerted by each group.

Sometimes this political dynamic leads to specific statutory provisions that narrowly constrain agency decision making. For example, under the antidumping (AD) law, the authorities must focus on narrow producer interests and may not consider consumer interests. Although this limitation leads to economically inefficient outcomes, the statute requires this focus. Yet even when the statute poses a more neutral question, the politics often trump the economics. Problems that readily lend themselves to economic analysis too often are left to more subjective decision making that can be influenced by politics.

No topic better illustrates this tension than the debate over "non-attribution," and the type of analysis required to prove "causation" in trade cases. This issue arises both in the context of "material injury" investigations conducted under Title VII of the Tariff Act of 1930 (covering antidumping and countervailing duty (AD/CVD) actions), as well as "serious injury" investigations under section 201 of the Trade Act of 1974 (so-called "safeguard" actions). In both instances, protection for the domestic industry requires a finding that imports are a "cause" of a specified level of injury to an industry petitioning for relief. The law requires authorities to distinguish the injurious effects of imports from the effects of other factors, and not to blame imports for these others factors. The legal issue is clear: under both US law and international law, national authorities have an obligation to distinguish the role of imports from the role of other causes. Yet the practical realities of fulfilling this obligation continue to pose serious challenges for national authorities. Although Title VII and section 201 injury investigations logically might rely on economic analyses to understand the effects of different causal factors, and rely on econometric analyses to measure the relative role of different factors, the International Trade Commission (ITC) seldom does so to any great degree.

This chapter will explore the issue of non-attribution in trade remedy cases from both a legal and an economic perspective. Section 2 surveys the law, both to describe

the obligation of non-attribution and to frame the task for decision makers. Section 3 then describes the history and current status of non-attribution at the ITC, the agency in the United States charged with making this determination in trade remedy cases. Section 4 presents a recent case study of how economics can shed light on the issue of non-attribution. The successful use of economic and econometric analysis in a recent AD case on cold-rolled steel highlights the potential of such analysis to remove politics from the decision to impose trade remedies. Section 5 then continues the story, describing the subsequent developments and severe backlash to the ITC negative determination – and refusal to impose AD duties – in the cold-rolled steel case. We then offer some concluding comments in Section 6.

The recent case on cold-rolled steel serves to highlight both the opportunities and problems of using economics to address the issue of non-attribution. The fit is almost too perfect, with the tools of economics and econometrics providing a powerful way to address the issue. If the authorities make the effort to collect the necessary data, they can undertake quite sophisticated analysis. Yet authorities fear loss of their decision making autonomy and, thus, embrace economics slowly and reluctantly, and only in a very tentative way. Having used economics in one recent case, the ITC even more recently considered an analogous situation and backed away from formal economics. For every step forward, there seems to be a step backward.

Nevertheless, the national authority's legal obligation not to blame imports for other factors is not going away. Imports cannot be blamed in place of other factors that may be injuring the domestic industry. If the authorities reject economics, they will need to find some other serious and disciplined approach to meet the non-attribution obligation. Perhaps over time the ITC and other national authorities will become more comfortable with the usefulness of economic and econometric techniques, and will adjust their investigative procedures to collect the necessary data on a more regular basis.

2 Defining the Problem: Statutory Language and International Obligations Requiring Non-Attribution

Under two important US trade remedy statutes, non-attribution is a key legal issue. Both AD duties and safeguard actions cannot move forward unless the ITC determines that imports are in some sense "causing" the problems faced by the competing US industry. This basic obligation to find that imports cause the problems, and that other factors are not responsible, serves as a major barrier – indeed, in many cases, the only barrier – to domestic industries successfully obtaining protectionist trade remedies.[1] Because of its importance, this issue is often quite contentious in trade remedy disputes.

2.1 Injury Investigations Under Title VII

2.1.1 US Law

Title VII of the Tariff Act of 1930 provides that AD duties will be imposed if two conditions are met. First, the US Department of Commerce (DOC) must determine

that merchandise subject to investigation is sold in the United States at less than "fair value" – in other words, is being "dumped." Second, the ITC must determine that an industry in the United States is injured "by reason of" imports of that merchandise [19 U.S.C. § 1673(1)–(2) (2000)].[2] To meet this second condition, the statute directs the ITC to analyze the domestic industry and the factors affecting its performance to determine if imports are a "cause" of one of three types of injury: (1) "material injury," which is defined as "harm which is not inconsequential, immaterial, or unimportant"; (2) "threat of material injury," which must be real and "imminent"; or (3) "material retardation," which requires a finding that an industry would have developed but for unfair imports [19 U.S.C. § 1677(7) (2000)].[3] Thus, the principal issue before the Commission is "material injury," whether current or prospective.

The ITC examines the role of imports in causing material injury by evaluating various factors specified by statute. This evaluation includes, but is not limited to, considering the volume and price effects of subject imports, as well as their impact on domestic producers [19 U.S.C. § 1677(7)(B)(i) (2000)].[4] The statute elaborates on each of the enumerated factors, directing the Commission to consider whether the volume and/or increase in imports is significant, whether there has been significant price underselling and price suppressing or depressing effects, and a range of domestic industry performance indicia. The statute further directs the ITC to consider such factors "within the context of the business cycle and conditions of competition that are distinctive to the affected industry" [19 U.S.C. § 1677(7)(C) (2000)].

Although it is easy to list factors to be considered, it is harder to specify what the decision maker should do with those factors. The statute itself is not explicit. The legislative history suggests that the ITC is not to weigh causes; imports need only be "a" cause of injury, and need not be "the" cause of injury.[5] The extent to which this bit of legislative history controls the ITC decision making, however, is subject to argument. Over the years, different Commissioners have interpreted the statutory mandate differently, and have embraced different analytic approaches.

But implicit in the statutory language is the idea that the ITC must distinguish between injury caused by imports and injury caused by other factors. This distinction may seem obvious, since the ITC cannot know what the injury from imports is, and consequently whether it amounts to "material injury," unless it isolates the effects of imports. To the extent this requirement may be implicit in the statutory language, the legislative history to the statute makes this requirement explicit, stating that the ITC "must examine other factors to ensure that it is not attributing injury from other sources to the subject imports" (H.R. Doc. 103-316, 1974, p. 851).[6] Non-attribution effectively demands that the causal effects of imports be segregated from other causes of injury.

2.1.2 International obligations

The US AD statute also must be read in light of US international obligations under the World Trade Organization (WTO) agreements, including a specific agreement on AD practices. Any ambiguity contained in the US AD statute with respect to causation and non-attribution is quickly drawn into focus by the WTO "Antidumping Agreement."

Article 3.5 unequivocally states:

> The demonstration of a causal relationship between the dumped imports and the injury to the domestic industry shall be based on an examination of *all* relevant evidence before the authorities. The authorities shall also examine any known factors other than the dumped imports which at the same time are injuring the domestic industry, and the *injuries caused by these other factors must not be attributed to the dumped imports* (WTO, 1994) (emphasis added).

Based on the above treaty language, there is little doubt that WTO members have accepted an obligation of non-attribution and must not impute injury to imports caused by other factors. As we discuss later, this language has now been interpreted definitively by the WTO Appellate Body in a recent case.

2.2 Injury Investigations Under Section 201

2.2.1 US LAW

Section 201 of the Trade Act of 1974 provides that if an article "is being imported in such increased quantities as to be a substantial cause of serious injury, or threat thereof, to the domestic industry producing an article like or directly competitive with the imported article," the domestic industry may be granted import relief to effect a positive adjustment to import competition [19 U.S.C. § 2251(a) (2000)]. For purposes of the statute, "substantial cause" means a cause "which is important and not less than any other cause" [19 U.S.C. § 2252(b)(1)(B) (2000)]. "Serious injury" is defined as "a significant overall impairment in the position of a domestic industry," whether current or clearly imminent [19 U.S.C. § 2252(c)(6)(C) (2000)].

Under section 201, the ITC must examine the relative role of imports. Unlike the Title VII standard, the effect of imports has to be "important" and not merely more than negligible. The statute specifies a host of nonexclusive factors related to industry performance and import volume for the Commission's consideration [19 U.S.C. § 2252(c)(1) (2000)]. There is an added requirement, however, to weigh the effect of imports against the effects of other causes of injury to determine if imports are at least as significant a cause of injury as every other individual cause. To this end, the statute explicitly directs the Commission to "examine factors other than imports which may be a cause of serious injury, or threat of serious injury, to the domestic industry [19 U.S.C. § 2253(c)(2)(B) (2000)]. It seems obvious, therefore, that the effects of these other causes must be separately identified at some level, or no comparison can be made.

2.2.2 INTERNATIONAL OBLIGATIONS

The WTO "Safeguards Agreement," which applies to section 201, supplements the statutory language and effectively confirms that non-attribution must be a requirement under section 201. Article 4.2(b) of the Agreement states that no import relief may be granted unless the injury investigation:

> ... demonstrates, on the basis of objective evidence, the existence of the causal link between increased imports of the product concerned and serious injury or threat thereof. *When factors*

other than increased imports are causing injury to the domestic industry at the same time, such injury shall not be attributed to increased imports (WTO, 1994) (emphasis added).

Under Article 4.2(b), the authority must first demonstrate "the existence of the causal link between increased imports . . . and serious injury or threat thereof." Moreover, this demonstration must be "on the basis of objective evidence." Under the plain meaning of these provisions, the authority must therefore demonstrate an explicit "causal link" between the increase in imports and any serious injury suffered by the domestic industry.

Although the first sentence of Article 4.2(b) imposes an explicit obligation to demonstrate the "causal link," the second sentence of the article goes even further. It requires, as part of the causation analysis, that "{w}hen factors other than increased imports are causing injury to the domestic industry at the same time, such injury shall not be attributed to increased imports." Under the plain meaning of this language, authorities may not demonstrate a causal nexus between imports and serious injury simply by correlating increased imports with serious injury. Mere correlation is a necessary, but insufficient, condition. Because other factors may be causing the decline in domestic industry performance, the authorities must specifically investigate other possible causes, and the injury from those alternative causes "shall not be attributed" to imports.

Thus, like the WTO Antidumping Agreement, the WTO Safeguards Agreement explicitly requires that administering authorities ensure that injury caused by other factors is not improperly imputed to imports. As we discuss later, this international obligation has also now been interpreted and clarified by the WTO Appellate Body.

3 Searching for a Solution: ITC Practice in Injury Investigations and Recent Developments

Although causation and non-attribution are often the central issues in any trade remedy case, the ITC has developed only very primitive analytic tools to address these crucial issues. Rather than develop any systematic or principled approach, as the courts and agencies have done for competition policy, the ITC has continued to use a case-by-case approach that often varies by individual Commissioner. A few Commissioners have embraced economic analysis, but they stand out as the exceptions to the more general practice of relying on trends and testimony by members of the affected industry.[7]

Not surprisingly, when this simplistic approach has been challenged in the WTO, the ITC has a poor record of defending its decisions as consistent with international obligations. In three recent cases, the WTO has found that the ITC approach to causation and non-attribution does not meet US obligations under the WTO Safeguards or Antidumping Agreements. The ITC needs to find a new, and more rigorous approach to this crucial issue.

3.1 Formal Economic Analysis at the ITC

The ITC has been around for a long time. Established in 1916 as the Tariff Commission, this independent quasi-judicial agency has been offering analysis of tariffs and other

trade policy instruments for many years. In 1974, the agency's mandate broadened and the name changed. For much of this history, the ITC was considered a somewhat sleepy backwater of Federal agencies. But the 1980s saw a period of intense intellectual activity at the ITC concerning the possible role of economics in trade remedy cases. Two factors converged: a change in the law to trigger many more trade remedy cases, and the arrival of several Commissioners very interested in exploring the possibility of an expanded role for economics.

The world of trade policy and the incentive to bring trade remedy cases changed dramatically with the Trade Agreements Act of 1979. Prior to 1979, the Department of Treasury administered the US antidumping and countervailing duty (AD/CVD) law, enforced the law cautiously, with a preference for finding less trade restrictive solutions when possible. Domestic industries rarely bothered to bring cases, and those cases were usually settled without imposing duties. Frustrated with this failure to protect domestic industries, Congress took the authority away from Treasury and assigned it to the DOC, an agency perceived as more sympathetic to the self-serving wishes of domestic industries.[8] The coincidence of this shift in administering responsibility, changes in the statute to make trade remedies easier to obtain and more automatic, and the recession in the early 1980s all triggered a surge in AD/CVD cases.[9]

With these new responsibilities also came new personalities. The appointment of Susan Liebeler in 1984, Anne Brunsdale in 1986, Ronald Cass in 1988, and Carol Crawford in 1991 brought to the Commission individuals interested in exploring the use of economics for helping to implement the statutory mandates. Several staffers at the Commission also contributed to this debate over the potential role of economics. With the advent of more powerful personal computers, and spreadsheet and econometric software packages to take advantage of ever-increasing computing power, professional economists could much more easily develop and implement formal models to assess the economic consequences of trade policy actions. Much of this intellectual activity would later appear in various working papers and published papers by these individuals.[10]

Indeed, other economists also joined in many of these debates on how to use economics in such trade remedy cases. In a section 201 case involving copper (Pindyck and Rotemberg, 1987)[11] and another involving carbon steel (Grossman, 1986),[12] outside academic experts weighed in and subsequently wrote scholarly papers about their efforts. In addition, during this period the Department of Justice and Federal Trade Commission (FTC) sometimes submitted amicus briefs to the ITC arguing far more serious attention to consumer interests. This work also sometimes generated scholarly papers.

The major result of this intellectual activity was the "Comparative Analysis of the Domestic Industry's Condition," or CADIC, model. This simple model sought to assess the effects of "dumping" (i.e., lower priced imports) on the price and volume of the competing domestic products. The model posits a simple counterfactual scenario: if the imports had not been dumped – if imports had higher prices – by how much would domestic prices increase and by how much would domestic shipment quantities increase? This initial model was subsequently refined as the "Commercial Policy Analysis System," or COMPAS, model.[13]

Both CADIC and COMPAS are Armington partial equilibrium models, which treat imports and domestic products as imperfect substitutes. Both models are simultaneous

equation, demand–supply, simulation models that can be presented in spreadsheet form. These types of models are thus easy to use in the real world. CADIC is a log–linear model. COMPAS is a nonlinear and therefore somewhat more general model.

COMPAS (the model currently in use) utilizes three key inputs and six elasticities to calculate the price and volume effects of dumping. The principal data elements are:

- the dumping margin percentage,
- the domestic producers' market share,
- the subject imports' market share,
- the aggregate elasticity of demand for the product under investigation,
- the domestic producer's supply elasticity,
- the fairly traded import supply elasticity,
- domestic product/subject imports elasticity of substitution,
- domestic product/fairly traded imports elasticity of substitution, and
- subject imports/fairly traded imports elasticity of substitution.

In an ITC context, these elasticities are actually "guestimates." They are rarely, themselves, the product of empirical calculation.[14] Rather, ITC staff set these parameters by interpreting the facts about consumers and producers developed during the investigation. Using these inputs, COMPAS calculates the effect of subject imports on price and volume of US producers' shipments, given the assumptions of the model.

COMPAS has merits as an analytical tool, given the practical constraints. The ITC must frequently deal with narrowly specified products being investigated for which there are little or no published economic data. COMPAS provides a way to apply informed economic judgment in a disciplined manner to limited market data collected directly from concerned importers, exporters, domestic producers, and purchasers. For many investigations COMPAS therefore provides a vehicle for working with limited information.

But COMPAS has severe limitations as a tool for considering the issue of non-attribution. Most fundamentally, COMPAS as a model does not have the flexibility to allow different factors to affect the supply and demand curves underlying the model. COMPAS makes assumptions so that the model can be implemented with limited data.[15] These assumptions about constant elasticities mean that the myriad of factors other than import competition that might cause the supply or demand curves to shift cannot be taken into account. For example, what if subject import market share increased due to the fact that there was a long strike or perhaps a fire at a mill? In both cases domestic profits would be falling and imports increasing. In both cases COMPAS would calculate a dollar impact of the imports. Yet, in fact, imports were pulled into the domestic market and, therefore, cannot be said to have caused any damage to the domestic industry. COMPAS has no ability to distinguish such distinct effects. In a sense, COMPAS is an acceptable tool for measuring the *magnitude* of impact that dumped imports *may* be having on the domestic market – all else being equal – but it cannot be used to *distinguish* the relative impact of different factors. It, therefore, fails to perform the necessary analysis required to ensure that the effects of other factors are not attributed to imports.

COMPAS has had a mixed history. Ultimately, CADIC/COMPAS never replaced traditional trends analysis at the Commission,[16] but remains an exercise in most AD cases, with a COMPAS "run" generally attached to the back of the staff reports that

accompany each ITC decision. The actual decision occasionally cites to the COMPAS runs as supporting a particular conclusion, but the output and decision model are never explicitly the basis for a decision.

Even this limited use, however, may be in jeopardy. In a recent opinion handed down by the Court of International Trade, the Court remanded a case back to the ITC in which the Commission had found no injury. The Court was persuaded, in part, by the fact that the COMPAS run at the back of the ITC's opinion in the case suggested the domestic industry may have in fact been injured by imports, yet the ITC had not even discussed these results (*Altx, Inc.* v. *United States*, 2001). This ruling might discourage the ITC from attaching COMPAS runs to their opinions in future cases.

3.2 Traditional ITC Practice on Causation

The ITC's traditional analysis of causation has been driven by anecdotal evidence of domestic industry developments and simple correlations among import levels, domestic production and shipments, and pricing. The analytical procedures involved have been largely *ad hoc*. At their crudest level, the approach consists of a combination of trends analysis (i.e., a qualitative and uncontrolled comparison of industry performance and import penetration); underselling analysis (i.e., a comparison of average prices of domestic goods and imports); and lost sales analysis (i.e., based on the claims by domestic producers).[17] It is hard to discern anything more sophisticated in most ITC reports.

Under this approach, if market penetration of imports has increased when domestic industry operating profits have fallen, this fact is considered to be "evidence" that imports have been causing the problems. Similarly, if domestic prices and output are also down, these facts are considered to be further evidence that increasing imports are to be blamed. Unless there is compelling qualitative evidence to force a disconnect in this simple analysis,[18] import volumes and pricing are assumed to be the cause of the industry's condition, rather than a consequence of other market developments. ITC practice, therefore, is not particularly careful in distinguishing causes to ensure non-attribution.

With binding obligations instituted by the WTO in 1995, it was only a matter of time before these shortcomings were identified by WTO dispute settlement. Four recent cases – three involving a US safeguard action on wheat gluten, lamb meat, and line pipe imports, and another involving a US AD action on hot-rolled steel imports, exposed the weaknesses in US practice. Each case went from a panel to the Appellate Body, producing definitive legal interpretations of US international obligations.

3.3 Causation for Safeguards: Wheat Gluten, Lamb Meat, and Line Pipe

3.3.1 Wheat Gluten

On October 1, 1997, the ITC initiated a safeguard investigation regarding imports of wheat gluten. It determined that the US wheat gluten industry was seriously injured

by reason of increased imports of wheat gluten (*Wheat Gluten*, 1998), leading to the imposition of quantitative restrictions on imports of wheat gluten (Proclamation 7103, 1998). The European Communities, representing the largest exporters of wheat gluten to the United States, challenged the action at the WTO. The dispute carried to the Appellate Body, which produced significant new precedent on the issue of non-attribution and the meaning of Article 4.2(b) of the WTO Safeguards Agreement (*AB Decision on Wheat Gluten*, 2000).[19]

The Appellate Body affirmed that other causes of injury must not be attributed to imports and that all relevant injury factors be examined (para. 69). But while acknowledging that the WTO Safeguards Agreement required authorities to distinguish between effects caused by increased imports and effects caused by other factors for purposes of non-attribution, the Appellate Body also concluded that the effects of *all* injury factors count toward a determination of serious injury (para. 72). This muddied the non-attribution waters by mixing in distinct issues of causation. There was no articulation of the relationship between the non-attribution requirement and the extent to which imports must contribute to serious injury. The Appellate Body's ruling in *Wheat Gluten* left unclear whether the non-attribution requirement was merely intended to determine whether increased imports are causing *any* injury to the domestic industry, or if it was intended as a means of assessing relative causes of injury, or somewhere in-between. The Appellate Body offered no clean answers, only possible interpretations.[20] According to the Appellate Body, the burden for an administering authority, after distinguishing the effects of all relevant injury factors, was simply to show whether "a causal link exists between increased imports and serious injury, and whether this causal link involves a genuine and substantial relationship of cause and effect between these two elements" (para. 69). What this really meant for non-attribution was effectively left to the Appellate Body's holding in relation to a US safeguards measure on imports of lamb meat just five months later (*AB Decision on Lamb Meat*, 2001).

3.3.2 Lamb Meat

On October 7, 1998, the USITC initiated a safeguard investigation regarding imports of lamb meat. It determined that the US lamb meat industry was threatened with serious injury by reason of lamb meat imports (*Lamb Meat*, 1999), leading to the imposition of a safeguard measure, in the form of a tariff-rate quota, on imports of lamb meat, effective as of July 22, 1999 (Proclamation 7208, 1999). Australia and New Zealand, the leading lamb meat exporters to the United States, challenged the action at the WTO. Among the claims raised by the two countries, and validated by a WTO dispute settlement panel, was the argument that the United States acted inconsistently with Article 4.2(b) of the WTO Safeguards Agreement because the ITC's determination did not ensure that threat of serious injury caused by "other factors" was not attributed to increased imports (*AB Decision on Lamb Meat*, 2001, p. 3). On appeal from a WTO dispute settlement panel decision, the Appellate Body agreed, applying the precedent announced in its decision in Wheat Gluten, thereby cementing the non-attribution requirement.

The Appellate Body focused on the record of the ITC investigation, noting that the report issued by the ITC stated that the "worsen[ing]" financial situation of the domestic

industry, as defined by the ITC, had occurred as "a result of the increase in imports" (p. 64).[21] At the same time, the ITC had identified six factors other than increased imports alleged to be contributing to the situation of the domestic industry.[22] Applying the US statutory standard, the ITC considered whether, individually, each of these six factors was a "more important cause" of the threat of serious injury than the increased imports. It concluded that each of these factors was not a more important cause than the increased imports (*Lamb Meat*, 1999, pp. I-24–6). The ITC then concluded that "the increased imports are an important cause, and a cause no less important than any other cause, of the threat of serious injury"(p. I-26), clearing the way for a tariff rate quota to be imposed by the President.

The Appellate Body noted that by examining the relative causal importance of the different causal factors, "the USITC clearly engaged in some kind of process to separate out, and identify, the effects of the different factors, including increased imports" (*AB Decision on Lamb Meat*, 2001, p. 65). However, the Appellate Body also found that an examination of the relative causal importance of the different causal factors, though seemingly compliant with US law, did not satisfy the requirements of the WTO Safeguards Agreement. In particular, the Appellate Body stated that Article 4.2(b) of the Agreement may only be satisfied "in the light of the explanation given by the USITC for its conclusions on the relative causal importance of the increased imports, as distinguished from the injurious effects of the other causal factors" (p. 65). Yet, according to the Appellate Body, the ITC only concluded that each of four of the six "other factors" was, relatively, a less important cause of injury than increased imports (p. 65, n. 142).[23] This was an implicit acknowledgment that these factors were actually causing injury to the domestic industry at the same time. The ITC failed to be certain, however, that the injury caused by these other factors was not attributed to increased imports by assessing the injurious effects of these other factors. "[A]lthough the USITC acknowledged that these other factors were having *some* injurious effects, it did not explain what these effects were, nor how those injurious effects were separated from the threat of serious injury caused by increased imports (p. 66)."

The Appellate Body framed the problem, and why the ITC had failed to meet its obligations under Article 4.2(b) of the WTO Safeguards Agreement, as follows:

> In a situation where *several factors* are causing injury "at the same time," a final determination about the injurious effects caused by *increased imports* can only be made if the injurious effects caused by all the different causal factors are distinguished and separated. Otherwise, any conclusion based exclusively on an assessment of only one of the causal factors – increased imports – rests on an uncertain foundation, because it *assumes* that the other causal factors are *not* causing the injury which has been ascribed to increased imports. The non-attribution language in Article 4.2(b) precludes such an assumption and, instead, requires that the competent authorities assess appropriately the injurious effects of the other factors, so that those effects may be disentangled from the injurious effects of the increased imports. In this way, the final determination rests, properly, on the genuine and substantial relationship of cause and effect between increased imports and serious injury (p. 63).

For the reasons stated above, the US safeguard action on lamb meat was declared invalid since it did not adequately explain how it ensured that injury caused to the domestic industry by factors other than increased imports was not attributed to increased imports.

3.3.3 LINE PIPE

On August 4, 1999, the ITC initiated a safeguard investigation regarding imports of line pipe, a type of steel product. After finding serious injury (*Circular Welded Carbon Quality Line Pipe*, 1999, *Line Pipe* hereafter),[24] the ITC split between tariffs and tariff-rate quotas. In the end, the President adopted tariff-rate quotas. Korea challenged this measure in the WTO, on a variety of grounds, including the failure to respect the obligation of non-attribution. The dispute settlement panel agreed, and the United States then appealed.

Building on its findings in the wheat gluten and lamb meat cases, the Appellate Body clarified that authorities must identify the nature and extent of the injurious effects of the known factors other than increased imports, as well as explain satisfactorily the nature and extent of the injurious effects of those other factors as distinguished from the injurious effects of the increased imports. To summarize the obligation of Article 4.2(b), the Appellate Body explained:

> Thus, to fulfill the requirement of Article 4.2(b), last sentence, the competent authorities must establish explicitly, through a reasoned and adequate explanation, that injury caused by factors other than increased imports is not attributed to increased imports. This explanation must be clear and unambiguous. It must not merely imply or suggest an explanation. It must be a straightforward explanation in express terms (*AB Decision on Line Pipe*, 2001, para. 217).

These obligations explicitly to separate and distinguish are now reflected in a long line of decisions, including wheat gluten, lamb meat, and line pipe in the safeguards context, and hot-rolled steel in the AD context. The plain language of the text and repeated interpretation by the Appellate Body have provided an unmistakable framework within which to analyze actions by the authorities.

3.4 Causation for Antidumping: Hot-Rolled Steel

On October 22, 1998, the DOC initiated an AD investigation into imports of hot-rolled steel from, among others, Japan, publishing its final affirmative dumping determination on May 6, 1999. On June 23, 1999, the ITC published its final affirmative determination of injury to the United States' hot-rolled steel industry (*Certain Hot-Rolled Steel Products from Japan*).[25] A final dumping order was issued on June 29, 1999. Japan challenged numerous aspects of the US action, including the ITC's injury determination. Among other arguments, Japan claimed that the ITC acted inconsistently with the causation requirements in Article 3.5 of the Antidumping Agreement, first, because it did not adequately examine factors, other than dumped imports, which were also causing injuries to the domestic industry and, second, because it failed to ensure that injuries caused by these other factors were not attributed to the dumped imports (World Trade Organization, July 2001, p. 72, *AB Decision on Hot-Rolled Steel* hereafter). Japan focused on four other factors raised during the course of the AD investigation it believed had been recognized by the ITC but not adequately considered (pp. 72–3).[26] Once again, the matter ended up before the WTO Appellate Body.

The WTO Appellate Body applied the same logic it had in its decision in the lamb meat case, this time in interpreting Article 3.5 of the WTO Antidumping Agreement. It concluded that in order for investigating authorities, applying Article 3.5, to ensure that the injurious effects of the other known factors are not attributed to dumped imports, they must appropriately assess the injurious effects of those other factors:

> Logically, such an assessment must involve separating and distinguishing the injurious effects of the other factors from the injurious effects of the dumped imports. If the injurious effects of the dumped imports are not appropriately separated and distinguished from the injurious effects of the other factors, the authorities will be unable to conclude that the injury they ascribe to dumped imports is actually caused by those imports, rather than by the other factors. Thus, in the absence of such separation and distinction of the different injurious effects, the investigating authorities would have no rational basis to conclude that the dumped imports are indeed causing the injury which, under the Antidumping Agreement, justifies the imposition of antidumping duties (pp. 74–5).

The Appellate Body was unmoved by arguments that causal factors may be difficult to disentangle:

> [D]ifferent causal factors operating on a domestic industry may interact, and their effects may well be inter-related, such that they produce a *combined* effect on the domestic industry. We recognize, therefore, that it may not be easy, as a practical matter, to separate and distinguish the injurious effects of different causal factors. However, although this process may not be easy, this is precisely what is envisaged by the non-attribution language. If the injurious effects of the dumped imports and the other known factors remain lumped together and indistinguishable, there is simply no means of knowing whether injury ascribed to dumped imports was, in reality, caused by other factors. Article 3.5, therefore, requires investigating authorities to undertake the process of assessing appropriately, and separating and distinguishing, the injurious effects of dumped imports from those of other known causal factors (p. 76).

Thus, like the Appellate Body's rulings concerning US practice regarding injury analysis under the WTO Safeguards Agreement, this decision cast into serious doubt the ITC's existing practice with respect to injury under the WTO Antidumping Agreement.[27]

4 Embracing a Solution: Cold-Rolled Steel – A Possible Approach to Non-Attribution

In light of the WTO Appellate Body rulings in the wheat gluten, lamb meat, line pipe, and hot-rolled steel cases, the question remains what approach should be used by national authorities to disentangle and analyze causal factors so as to ensure non-attribution. The rulings themselves do not provide the answer; they only note that neither the WTO Safeguards Agreement nor the WTO Antidumping Agreement prescribes any specific method (*AB Report on Hot-Rolled Steel*, 2001, p. 75; *AB Report on Lamb Meat*, 2001, p. 64). Since the US statute and the WTO Safeguards and Antidumping Agreements focus primarily on economic criteria, it seems logical that more rigorous economic analyses might be employed to meet the non-attribution requirement. At the

same time, such analyses might also help answer whether injury caused by imports is more than "negligible" in the AD context, and "important" and not less than any other cause in the safeguards context.

In early 2000, the ITC was presented with a concrete example of how more sophisticated economic analyses could help disentangle causal factors and meet the non-attribution obligations under US law and at the WTO. In 1999, the domestic steel industry filed AD petitions against imports of cold-rolled steel from several foreign countries. As part of this investigation, the ITC had to conduct an injury investigation, and had to assess the relative roles of imports and other factors (*Certain Cold-Rolled Steel Products from Argentina*, et al., 2000, *Cold-Rolled Steel* hereafter).

In the world of ITC litigation, the domestic petitioners always frame the economic dynamics as unfair imports battering the domestic industry, but the reality is usually more complex. Consider, for example, a situation with both falling import prices and falling domestic prices. Considered in isolation, one might infer that import prices are affecting domestic prices. But both trends might be in reaction to some other factor, such as falling input costs or increasing intra-industry competition within the domestic industry. In many ITC cases, the foreign respondents point to these other factors qualitatively, but with limited effect. If the other factors are dramatic, the ITC will sometimes agree with these qualitative arguments. The problem, however, occurs when several factors are at play at once, and no one factor alone is so obvious. When the other factors are less obvious, the ITC tendency has been to blame imports.

The AD case involving cold-rolled steel had the potential to become such a case. The conventional wisdom during that case was that the domestic industry would surely win. The case appeared to have all the typical elements needed for the domestic industry to win: falling import prices, import prices underselling domestic prices, and falling domestic prices that lead to financial losses. In the jargon of Washington, DC trade lawyers, this was a "price case" and the correlation of falling import price and falling domestic prices, and the import underselling of domestic price levels, meant that the domestic industry would win.

The challenge in that case was to force a more serious look at what was really happening in the cold-rolled market. Since conventional approaches were unlikely to prevail, the foreign respondents in that case decided to pursue a less conventional approach. Although there had been little success in persuading the ITC to embrace formal economic and econometric models in other AD cases, the foreign respondents realized they had to try something to show dramatically that imports were not driving down domestic prices. The defense team therefore took advantage of the availability of extensive and detailed public data about the steel industry, commissioning a more sophisticated economic study of the factors affecting the domestic prices.

4.1 The Economics

The basic approach was to identify those factors affecting supply and demand, and then analyze and quantify each of those effects separately. Rather than simply correlate declining import prices and declining domestic prices, and concluding that import prices drove down domestic prices, the study instead looked at the underlying factors

affecting supply and demand. Although this approach may seem incredibly basic and obvious to economists, it is sometimes amazing how the most basic economic insights can be the most surprising and counterintuitive to noneconomists.

The study (Prusa, 2000)[28] considered several basic factors that one would expect to affect the supply of and demand for "cold-rolled steel," a particular type of steel. Although a myriad of other factors were also at play, these factors captured key determinants of demand and supply. More importantly, once identified and specified, these other factors could be measured. Such an approach thus would allow measuring the role of imports, while controlling for these other factors.

The specific model used in the cold-rolled steel case considered various demand and supply factors.[29] The demand factors included: (1) domestic cold-rolled steel shipments; (2) the price of "subject" cold-rolled steel imports;[30] (3) the price of "nonsubject" cold-rolled steel imports; (4) US auto assemblies; (5) industrial machinery; (6) the domestic price of galvanized steel; and (7) galvanized steel production capacity. The model's supply factors included: (1) the domestic price of hot-rolled sheet; (2) the import price of hot-rolled sheet; (3) coal prices; (4) steel scrap prices; (5) iron ore prices; (6) steel slab prices; (7) cold-rolled steel capacity; and (8) domestic cold-rolled shipments. We now review the economic logic behind each of these factors.[31]

4.1.1 Factors Influencing the Demand for Cold-Rolled Steel

Automobiles and other uses. Cold-rolled steel is primarily used as an intermediate product to make other products. For instance, a large amount of cold-rolled steel is ultimately used by the automobile industry. The study therefore explicitly included automobile production and industrial machinery production as two different measures of these sources of downstream demand.

Hot-dipped galvanized market. Cold-rolled steel is actually a middle stage in integrated steel production. The production flow is: slab, hot-rolled steel, cold-rolled steel, and finally galvanized steel. The "coating" market is thus actually the largest buyer of cold-rolled steel sheet. About three quarters of all coated steel is galvanized (i.e., coated with zinc), and hence changes in the galvanized market will likely have a significant effect on the cold-rolled steel market.

In the 1990s, large investments in new capacity have been the driving force behind the changes in the galvanized market. The new capacity had a significant impact on the cold-rolled market in at least three ways. First, the new galvanized capacity increased the demand for cold-rolled steel and put upward pressure on cold-rolled prices. Second, the new galvanized capacity caused substantial downward price on galvanized prices which in turn has put downward pressure on the price of cold-rolled steel. Third, and perhaps most importantly, the galvanized steel market is where the highest profit margins are located. Sensibly, profit-maximizing producers that produce both cold-rolled and galvanized steel will seek to move a greater fraction of their cold-rolled output downstream. As a result of the dynamics of domestic firms chasing the profits down the product line, the galvanized market squeezes the price of cold-rolled steel.

To reflect these dynamics, the study included both the domestic price of galvanized steel and changes in domestic capacity to produce galvanized steel.

Cold-rolled import market. Imported cold-rolled steel is a substitute for domestically produced cold-rolled steel. Buyers needing cold-rolled sheet can turn to either domestic or foreign suppliers. In other words, a decrease in the foreign price should put downward pressure on the price of the domestic substitute good.

But imports are not a perfect substitute for domestic products. In the case of cold-rolled sheet, the price pressure may not be as large as one might expect for two reasons. First, foreign steel has a far longer delivery time period than domestic steel, typically 120 days. Buyers may simply find such lags unacceptable and purchase their product from domestic suppliers despite the price differences. Alternatively, buyers may demand other concessions from foreign suppliers to offset the delivery lag. In either case, the delivery lag associated with imports will serve to depress the cross-price effect. Second, some buyers only purchase cold-rolled sheet if the product can meet their exacting specifications. In many cases, the existence of low price foreign suppliers is irrelevant since these potential suppliers do not produce steel of sufficient quality.

Despite these caveats, one expects the price and volume of imported cold-rolled to have some impact on the domestic price of steel. The base level model therefore included the price of imports – both subject imports and nonsubject imports – in the model.

4.1.2 FACTORS INFLUENCING THE SUPPLY OF COLD-ROLLED STEEL

Primary inputs. Traditional mills produce cold-rolled steel using an integrated production process where they start with primary inputs (coal, iron ore, coke) or steel slab and produce hot-rolled steel. The hot-rolled steel is then moved "down the line" and cold-rolled. Changes in the price of these key primary inputs can raise or lower the supply curve for the domestic industry. The study therefore included producer price indices for coal, iron ore, and carbon steel scrap.[32] The study also included the price of steel slab as an alternative measure of these inputs.

Hot-rolled steel. Although these primary inputs may have some effect on the price of cold-rolled steel, it must be stressed that the prior stage product – hot-rolled steel – is the actual raw input for cold-rolled steel. Because hot-rolled steel is the essential input for cold-rolled steel, changes in the price of hot-rolled should have a direct effect on the supply of cold-rolled steel. Since hot-rolled steel is such a large component of the cold-rolled product, one expects cold-rolled steel to be particularly price sensitive to changes in hot-rolled prices.

Although all cold-rolled producers will be affected by the price of hot-rolled steel, the mechanism differs according to the type of producer. The impact of the price of hot-rolled is most directly felt by "re-rollers" and "processors." These cold-rolled producers do not start with primary inputs, but rather purchase hot-rolled steel from domestic or foreign suppliers. A fall in the price of hot-rolled steel will immediately translate into

a change in their costs of production and (due to competitive pressures) likely lead to a fall in the price of cold-rolled steel.

The price of hot-rolled steel will also affect traditional integrated producers, but in this case the economics are somewhat different. Integrated producers produce cold-rolled steel sheet starting with the primary inputs. For them, hot-rolled steel is just an intermediate good and therefore it may not be apparent why the price of hot-rolled matters for their cold-rolled price. But there are numerous economic reasons why changes in the price of hot-rolled steel will have the same basic effect on the integrated producers as it does on re-rollers and processors. First, a decrease in the price of hot-rolled steel will intensify the incentive to push product downstream to the cold-rolled and galvanized markets. But this shift simply intensifies competition in the cold-rolled market and pushes the cold-rolled prices down. Second, a fall in the price of hot-rolled means that a set of domestic producers (the re-rollers and processors) can now lower their cold-rolled price and still make a profit. This competition creates pressure for the integrated mills to lower their price exactly as if hot-rolled steel were the key input in their cold-rolled production process. Third, nothing stops the integrated mills from purchasing hot-rolled sheet on the open market.[33] All things considered, given that hot-rolled steel is the raw input for cold-rolled steel, one expects changes in the price of hot-rolled to influence the price of cold-rolled. The study therefore included the price of domestic hot-rolled steel and the price of imported hot-rolled steel.

Capacity. One would also expect the capacity to produce cold-rolled steel to have an effect on the supply of cold-rolled steel. But these effects would be felt over time, as new capacity came on line and the difficulties related to starting up new capacity were worked out over time. The study therefore included domestic cold-rolled capacity.

4.2 The Econometrics

The basic economics provided a theoretical framework for analysis; but the challenge remained how to go from qualitative analysis to quantitative analysis. Again working with very basic textbook theory, the study used two-stage least squares to estimate the supply–demand system.[34]

Simultaneous equations systems, such as any model of supply and demand, occur quite frequently in economics and business contexts. For statistically reliable estimations, however, simultaneous systems require a more complex estimation procedure than single equation models, which can generally be estimated through ordinary least squares (OLS) regression. The most frequently used method of estimating simultaneous systems is the "two-stage least squares" method, which builds upon OLS estimation. Not surprisingly, most standard econometrics texts introduce the method of two-stage least squares with the example of a supply and demand model. The two concepts literally go hand-in-hand.[35]

Although this approach will seem quite routine to economists, it was considered quite radical by the lawyers. Some lawyers might have had an introductory course in statistics, and had some exposure to the basic concept of regression analysis, usually

in the form of simple OLS. Few lawyers had ever heard of “two-stage least squares.” To most trade lawyers, it was some mystical “black box” that somehow produced an answer.

The method was novel, but so was the availability of sufficient data. In many trade cases, the product being attacked is so narrow that one simply cannot find sufficiently detailed time series data to undertake serious econometric analysis. But the steel industry is different. Various consulting firms and trade associations track quantities and market prices for key products, generally on a monthly basis. Like any data set, there are problems through which one must work. But the basic data exists, and can be collected or purchased.

Using two-stage least squares actually to estimate a supply–demand system represents a significant advancement relative to the COMPAS model. Like the COMPAS model, this approach is grounded in microeconomic theory and employs simultaneous equations to address the interaction of supply and demand in the market. Unlike the COMPAS model, however, the parameters emerging from a two-stage least squares model reflect serious analysis applied to time series data. The parameters are actually estimated, not “guestimated” by the analyst. The two-stage least squares method, however, is data intensive; it requires an adequate number of data points for each time series, and there must be a reliable set of time series data available for each significant variable. Although this requirement automatically limits the applicability of this approach in many injury investigations, there were adequate data for its application to cold-rolled steel. For all variables, the study drew on or developed five years or more of monthly data (or at least quarterly) to conduct the analysis.

Moreover, unlike the COMPAS model, two-stage least squares allows for more flexibility in determining the slope and positioning of the supply and demand curves. Specifically, it does not rigidly assume that the demand curve is perfectly inelastic (vertical). Since it also does not require constant elasticities, this approach allows for shifts in the supply and demand curves that come about from changes in other underlying factors.

Moreover, such an approach allows one to carefully distinguish factors affecting the supply curve versus those affecting the demand curve. When the two are blurred, one can easily draw the wrong inferences. For example, suppose that because of import competition, the demand curve contracts (move in), and the equilibrium price falls somewhat. But if at the same time input prices fall, the supply curve will expand (move out), which also will produce a fall in the equilibrium price. If one does not carefully distinguish these two effects, one might erroneously blame import competition for a dramatic price decline, when in fact the prices are falling more because of falling input costs rather than foreign competition.[36]

4.3 The Results

Using the above factors, the study estimated and distinguished the factors determining the price of cold-rolled steel in the US market with a degree of precision and confidence impossible with COMPAS. A base line scenario provided the following econometric results.

Table 3.1 Estimating the effect of price of imports on the domestic price of cold-rolled sheet

Factors	Demand	Supply
Demand		
Domestic cold-rolled sheet shipments	−0.431 (0.252)	0.055 (0.216)
Price-subject imports (lag 4)	0.022 (0.113)	
Price-nonsubject imports (lag 4)	0.158 (0.126)	
Auto assemblies (lag 3)	0.105 (0.101)	
Industrial machinery production (lag 3)	−0.519 (0.090)**	
Domestic price hot-dip galvanized steel (lag 1)	0.675 (0.089)**	
Galvanized capacity (lag 9)	0.481 (0.171)**	
Supply		
Domestic price hot-rolled sheet (lag 1)		0.918 (0.096)**
Import price hot-rolled sheet (lag 1)		0.242 (0.113)*
Coal PPI (lag 1)		0.193 (0.245)
Carbon steel scrap PPI (lag 1)		0.001 (0.118)
Iron ore PPI (lag 1)		0.086 (0.127)
Price steel slab (lag 1)		−0.123 (0.089)
Cold-rolled capacity (lag 9)		−0.014 (0.179)
Constant	6.795 (2.896)*	−1.872 (3.723)
Observations	77	77
R-squared	0.717	0.883

Note: *Standard errors in parentheses; ln–ln estimation; * significant at 5% level; ** significant at 1% level. References to "lag" mean the variable was lagged by that number of months. The study also considered many alternative lag structures; the main result – that imports did not matter – was robust to alternative lag structures.*

This specification provided an excellent benchmark, as it includes all key factors affecting the supply and demand for cold-rolled steel. By standard economic and statistical conventions, the econometric model fit remarkably well. Just as economic theory suggests, the estimates imply that the demand curve slopes downward and that the supply curve slopes upward. One can confirm this by looking at the estimate of "Domestic cold-rolled sheet shipments" in Table 3.1. In the demand equation this parameter is negative, and in the supply equation the parameter is positive. This sensible finding is not imposed on the model, but rather is determined by the data. This is strong indication that the econometric model accurately describes the industry.[37]

One attractive aspect of this estimation approach is that one can quantify the *individual* impact of the different causes. Translating these parameter estimates into dollar per ton estimates in price charges, the study found that:

- The fall in the price of subject imports caused the domestic price of cold-rolled steel to fall by less than $1.50/ton.
- The price of hot-rolled steel sheet was the single most important explanation for trends in the cold-rolled market. A dramatic fall in the price of hot-rolled steel caused the price of cold-rolled steel to fall by $55/ton.

- The fall in the price of galvanized steel caused the domestic price of cold-rolled steel to fall by $28/ton.
- Capacity additions influenced the price of cold-rolled steel. Recent additions to galvanized capacity increased the demand for cold-rolled steel, thereby increasing the price for cold-rolled steel by $35/ton. In conjunction, changes in the galvanized market caused a net increase in the price of cold-rolled steel of $7/ton.
- Additions to cold-rolled steel capacity caused a softening in the price of cold-rolled steel, with the expectation of further softening as capacity ramped up.
- Cold-rolled steel imports from nonsubject countries played a small role in determining the price for cold-rolled steel, but significantly larger than the impact of imports from subject countries. The decrease in the price of nonsubject imports caused the domestic price of cold-rolled steel to fall by $8/ton, also found to be statistically insignificant.[38]

In addition, the effect of import prices was estimated to be statistically indistinguishable from zero. In other words, the fall in the price of subject imports was an economically and statistically insignificant cause of the fall in the domestic price of cold-rolled steel.[39] Several of the other factors, in contrast, were statistically significant at either the 95 or 99 percent level.

The parameters on subject imports were particularly important. A positive sign means that a *decrease* in the price of subject imports will put *downward pressure* on the domestic price of cold-rolled steel. In other words, imported cold-rolled and domestic cold-rolled are substitutes. Crucially, in all the specifications the estimated effect was quite small. For instance, in this base specification, the parameter implies the cross-price effect is only 2.2 percent.[40]

What does the estimate of 2.2 percent mean in easy-to-understand terms? In simplest terms, the estimate reveals how changes in one variable (in this case, the price of subject imports) influence changes in the price of cold-rolled steel. During the period of investigation, the average price of subject imports fell by 14 percent. The model thus predicts that the price of subject imports explains $2.2 \times 14 = 0.308$ percent of the change in the price of domestic cold-rolled steel.

Given the 1996–7 price level of cold-rolled steel was about $470, this means that $1.45 (i.e., $470 × 0.308 percent = $1.45) of the fall in the price of domestic cold-rolled steel is due to the price of subject imports. This small estimate indicates that imports had a very small effect – an economically insignificant effect – on domestic prices. In fact, the study found that the effect of cold-rolled imports was also statistically insignificant. The standard error of 0.113 – the measure of variability of the parameter estimate – was larger than the parameter estimate of 0.022. The study found that other causes explained the cold-rolled steel market. In particular, falling input prices and nonsubject imports accounted for about 90 percent of the change in the domestic price of cold-rolled steel.

Even under the most expansive interpretation of the AD statute – that imports need only be "a" cause of material injury – the facts of this particular case did not support a finding that imports were to blame. If the statistical evidence of a relationship is so weak that one cannot distinguish the effect from zero, then it is impossible to say that imports are "a" cause. So whether one were inclined to weigh causes or simply try

to discern any impact of imports, the model demonstrated that imports just did not matter.

The ITC's Office of Economics could find little to discredit the model. A staff report issued prior to the Commission's determination confirmed the model's economic theory and econometric approach were mainstream and robust.[41] In a very unusual action, this staff memo was then summarized in the final staff report (p. II-15) – the official distillation of facts and analysis provided by the staff to the Commission – a level of attention never before provided in an ITC staff report.[42]

Ultimately, the ITC did in fact rule that the domestic cold-rolled steel industry was not injured by imports, although its opinion was crafted carefully. The opinion discussed only the traditional evidence – trends analysis and qualitative reasons that import prices did not have much effect on domestic prices. In a footnote, the opinion then noted that these findings were consistent with the findings of respondents' econometric model.[43] Again, this footnote, while down-playing the extent of reliance on the model, was unprecedented. We could find no other example of the Commission explicitly relying on an econometric study submitted by one side or the other in an AD or CVD case.

4.4 The Implications

This econometric study in the context of an AD investigation was important for two reasons. First, the study showed how one could use mainstream and relatively simple economic and econometric analysis to demonstrate that imports were having only a negligible effect on the domestic industry. Under the US dumping law, imports need only be "a cause" of material injury. Although there is no clear legal definition, it is hard to argue that an effect that is both economically and statistically negligible could be deemed legally sufficient. If the role of imports is so small as to be lost in the "statistical noise," then imports do not rise to the level of being "a cause" of injury.

Second, the study showed how one could control for different factors. COMPAS also provides a framework for measuring the magnitude of the effect that *imports* are having on domestic prices and quantities. But COMPAS does not allow *other* factors to shift the demand and supply curves, and thus tends to overestimate the effect of imports. Changes in imports might well be blamed for changes occurring because of other factors. In the more flexible structural model, these other factors can be hypothesized and then measured. Essentially, the legal requirement to distinguish causes can be met.

4.5 Why it Worked

In retrospect, the ITC embraced this approach to causation and non-attribution for several reasons. Different Commissioners and staff may have been motivated by different factors, but collectively the following factors were important.

First, the domestic industry underestimated the potential persuasive impact of this approach and responded poorly. At the time, the domestic industry seemed somewhat dismissive of the formal economic argument. Although respondents presented their economic study in a pre-hearing brief, and stressed the study at the public hearing,

the petitioners devoted relatively little time to the economic study at the hearing and never presented an alternative economic study. The petitioners' post-hearing brief was confined to criticizing what respondents had done, but not presenting any alternatives (*Post-hearing Brief of Bethlehem Steel*, et al., 2000, pp. 3–15).[44]

Second, the basic economic and econometric arguments were all very conventional. Petitioners focused their attacks on the respondents' economic study, but the study was hard to attack. The theoretical framework was a very basic supply–demand model, articulating the various factors that might affect either supply or demand. It would be hard for the theory to be more basic. The econometric approach was also quite mainstream. Although two-stage least square estimation may have seemed esoteric to the lawyers, from an economic perspective the approach could not have been more basic. Any well-trained economist would immediately recognize the theory and the empirical approach as very standard and well accepted. Not surprisingly, when presented to the ITC economics staff for analysis and comment, the staff largely embraced the study as theoretically valid and properly done (McDaniel, 2000).

Third, the formal economic arguments meshed very well with more traditional arguments in the proceeding about causation. The respondents stressed the econometric study at the outset in the pre-hearing brief and at the hearing, but then shifted emphasis to the more traditional approach to causation in post-hearing submissions. Thus, in their post-hearing brief, and in final comments on the facts on the record, respondents stressed all of the qualitative evidence that the price of cold-rolled steel imports had no meaningful effect on the domestic price levels. For example, the fact that domestic customers identified domestic steel firms as price leaders, and the fact that domestic steel prices had begun to rise long before import prices began to rise (*Cold-Rolled Steel*, 2000, pp. 22–3), tended to reinforce the findings of the econometric study that import prices did not matter very much. Indeed, the final written submission by respondents relied solely on qualitative evidence – and did not rely on the econometric study at all. Ultimately, the Commission embraced this approach: that the study merely corroborated the other evidence on the record (p. 23, fn. 185).

Finally, the staff seemed genuinely interested in allowing economics to play a more prominent role in the analysis. As professional economists, the ITC staff understood the potential for economics to shed useful light on many issues, such as non-attribution. So the staff seemed willing to depart from tradition and give the economic study serious consideration. This interest can be seen in the unusual one-page discussion of the economic study previously mentioned (p. II-15). This one page reflected a more sustained discussion of a formal economic and econometric model than had ever been included in any other ITC staff report. It may seem strange that such a limited discussion would in fact be path breaking, but such is the state of formal economic analysis at the ITC.

Unlike the COMPAS model, the two-stage least squares approach clearly passes muster under US law and the WTO in terms of non-attribution, as evidenced by the ability to distinguish between causal factors in the cold-rolled steel case. Moreover, the precision with which this approach was able to estimate the effect of each causal factor in the cold-rolled steel case demonstrates that it could help guide decisions on whether injury is "material" under the AD statute, or important and not less than any other cause under section 201.

5 The Backlash: The Political Consequences of Embracing Economic Analysis

The reactions to the ITC acceptance of formal economics were swift and sharp. Although the ITC had tried very hard to downplay its explicit reliance on the formal model, the domestic steel industry seized upon this novel aspect of the case as the most vulnerable to attack.[45] Under the petitioners' view of the world, since everyone "knows" that steel imports drive down domestic steel prices, something must have gone seriously wrong with the ITC analysis.

In the immediate aftermath, the domestic industry severely criticized the decision with dramatic rhetoric. As one newspaper report summarized:

> Roger Schagrin and Alan Wolff, domestic trade lawyers in the case, said that "voodoo economics" and "bad economics" were the overriding factors in convincing the commissioners to vote 5–1 against material injury. Schagrin hoped aloud that the commission had not regressed to relying on faulty analysis in reaching its conclusions, as he said some former commissioners had done. Following up at a trade panel forum on Capitol Hill March 15, Bethlehem Steel Corp.'s Washington lobbyist, Laird Patterson, whose employer was a petitioner in the case, chastised the commission for "embracing rogue econometrics" and flaunting "a fundamental disregard of congressional injury standards." Patterson said the vote went so far that the "objectivity of the commission comes into question" (Kelly, "Potomac pulse," 2000).[46]

Not content merely to complain about "voodoo economics" and "rogue econometrics," the domestic steel industry moved ahead with a multiprong political attack on the Commission. The attack had several dimensions. First, the domestic steel industry's allies in Congress proposed a budget cut for the ITC. Although the actual budget cut failed, the House Appropriations Committee did express concern in its report accompanying the appropriations legislation that the "Commission is expending resources to analyze theoretical constructs of the conditions of competition, rather than focusing on the market realities faced by domestic industries injured by foreign unfair trade practices in accordance with the statutorily mandated factors . . .." Moreover, the Committee sought justification for the Office of Economics, requesting that the ITC provide "a list of positions within the Office of Economics, and the specific functions performed by these positions." Perhaps with specific reference to the Office of Economics, the Committee expressed further concern that the "Commission is conducting studies of questionable relevance to its core mission. The Committee believes that the Commission could better allocate its resources by focusing on practical studies that will enhance its ability to effectively carry out its duties in accordance with the law" (*H.R. Rep.*, 2000, pp. 106–680).

Second, the domestic steel industry used other cases to highlight (for the ITC) Congressional anger with what the ITC had done. A few months later, in a case involving tin mill products, the domestic industry brought forward extensive Congressional testimony, including some rather pointed comments by Senator Byrd and Senator Rockfeller from West Virginia, the home state of the petitioner Weirton Steel:

> I must admit to you I was absolutely shocked that, I think with the exception of one Commissioner, you reached your decision in the cold-rolled case, and even in your unanimous affirmative

preliminary in this case to see that the provision on the conditions of competition have been transformed into a mechanism to deny relief in my judgment, to deny relief, which is what I predict you're going to do, notwithstanding evidence of injury, facts and then whatever it is that goes through your minds and makes you come up with the decision that you come up with.

Let's face it. The numbers are the numbers in these cases. No rocket science in anything I've said. From what I understand, what is occurring here at the Commission is that lawyers and economists representing foreign competition or US buyers, who can't really argue about the numbers that your staff has gathered, are spending all their time before the Commission arguing about conditions of competition and taking the focus away from the statute, which is called the law insofar as I'm aware. I find this deeply disturbing, and I hope, perhaps vainly, that the Commission will seriously reconsider the matter. *There should be no need for Congressional action* (Rockefeller, 2000) (emphasis added).

At the same time, the domestic steel industry recognized that it needed to respond better on a substantive level as well. So in several AD cases after the cold-rolled decision, the domestic industries hired their own economists to do studies "proving" that imports did affect domestic prices. The most noteworthy example was an AD case involving structural steel, for which the domestic industry had Jerry Hausman of MIT testify.

This aggressive response culminated in the political and economic battle in the recent section 201 case on steel. Politically, the section 201 saw the domestic steel industry undertake its maximum effort. During the eight days of hearings on injury in the section 201 case, on the first day alone no less than 17 Representatives, 13 Senators, and two Governors were scheduled to give testimony on the importance of the steel industry to their districts, to their states, and to the country as a whole (USITC Witness List). Many of these members of Congress made national defense arguments, pointing to the September 11 tragedy and the ongoing military response as a reason to protect the domestic steel industry.

Economically, the domestic steel industry turned to Professor Jerry Hausman of MIT to bolster its arguments that imports in fact drove down domestic prices. Having responded too slowly in the prior case, the domestic industry wanted to take the offense in this case. Professor Hausman provided a study that argued imports had a lingering effect in the market that drove down domestic prices.[47] He concluded that imports accounted for the vast majority of the decline in domestic steel prices. Interestingly, even Hausman's own model showed that imports of cold-rolled steel – the product at issue in the earlier dumping case – had little effect on domestic cold-rolled prices and that imports of corrosion-resistant steel had little effect on domestic corrosion-resistant steel prices. Rather, Hausman's model pointed to the significant increase in *hot-rolled* imports in 1998 as having a lingering effect on *all* domestic flat-rolled steel prices: *hot-rolled*, cold-rolled, and galvanized steel.

In contrast, the foreign respondents also used formal economic analysis to show the relatively small role that imports played, relative to other factors.[48] Professor Prusa of Rutgers University developed product-specific studies of *hot-rolled*, cold-rolled, and corrosion-resistant steel. In each, he found that other factors were more important than imports in explaining declines in domestic price levels. In particular, he found that the factors listed in Table 3.2 mattered the most in explaining the price declines.[49]

In these analyses for the section 201 case, Professor Prusa used reduced form models, rather than structural models. Although the foreign respondents recognized

Table 3.2 Magnitude of different factors in explaining domestic price declines

Factor	Hot-rolled (/ton)	Cold-rolled (/ton)	Galvanized (/ton)
Changes in the price of subject imports	$8–10	$2	$1–2
Changes in scrap prices, a key raw material input	$20	$15	NA
Additions of domestic capacity	$7	$10	$90
Additional effect of adding minimill capacity	$20	$50	NA
Declines in domestic demand	$40	$40	$60
Role of intra-industry competition from low cost minimills	$30	$20	NA

Note: *These studies can be found in Joint Pre-hearing Brief of Respondent Producers in Argentina, Brazil, European Union, India, Japan, Korea, Russia, Taiwan, Thailand, and Venezuela before the ITC: Carbon and Alloy Flat Products, Product 6 – Corrosion-Resistant and Other Coated Sheet and Strip, Inv. TA-201-703, September 10, 2000 (filed by the law firm of Sharrets Paley); Respondents' Joint Pre-hearing Brief on Hot-Rolled Steel (Product Category A.3), Inv. TA-201-703, September 10, 2000 (filled by the law firm of Kaye Scholer); Joint Pre-hearing Brief of Respondents: Product Group 4, Cold-Rolled Steel, Inv. TA-201-703, September 10, 2000 (filed by the law firm of Willkie Farr). All these briefs are available on Electronic Document Imaging System (EDIS). Note that the dollar per ton estimates summarized here do not sum to the dollar per ton decline in the domestic price of the various products. The studies included various factors, some raising and some lowering prices. In addition, in some instances the measures should be viewed as alternative ways of looking at related factors, and should not be aggregated. Minimills are not yet major players in the galvanized segment, and thus variables relating to minimills did not apply in the model of galvanized steel.*

that structural models (like that used in the cold-rolled steel AD case) would allow more precise estimates of various parameters, the magnitude of the work and time limitations of the section 201 case did not allow this more refined approach. Besides, the purpose of a section 201 investigation was fundamentally different: rather than assessing the magnitude of imports and showing their individual effect to be negligible, in this legal context foreign respondents needed only to show that some other factor was more important than imports. Given the different legal context, the focus was much more on alternative causes, and less on the magnitude of the effect of imports themselves.

Note that these results from Professor Prusa's studies are quite dramatic: the impact of imports is always small, and usually statistically insignificant.[50] The reduced form models provide quite strong evidence that other factors mattered more than imports in driving down domestic prices. Taken together, both the number of other factors that mattered more and the extent to which these factors mattered more, these alternative causes substantially contradicted the argument being offered by the domestic industry.

A full discussion of these competing economic models is beyond the scope of this chapter, but a few points should be noted. First, Professor Hausman's results depended very much on the structure of his lags. He allowed imports to have a very long-term lingering effect in the market, but did not allow his other factors to have the same effect. In other words, when Professor Hausman's model is adjusted to allow a similar

structure of lag on other variables (e.g., capacity), his finding – that imports mattered more than other factors – collapsed. Second, Professor Hausman's economic model left out or inadequately described various factors that one would expect to have a large effect on market dynamics, such as the role of increasing capacity, the role of intra-industry competition, and the role of changes in aggregate demand (Post-Hearing Brief on behalf of Joint Respondents, 2001).

Given the political controversy over the prior cold-rolled steel case, the Commissioners, not surprisingly, shied away from any serious consideration of the economic models. The staff also took a very different approach. The contrast between the internal staff memos in the two cases is quite dramatic. In the cold-rolled case, the staff economist devoted five pages to a fairly broad-based review of the competing economic arguments. In the section 201 case, the staff economists spent less than three pages to consider a much lengthier and complex set of arguments. Moreover, the section 201 memo is internally inconsistent. The body of the memo is more sympathetic to Professor Prusa's studies than to Professor Hausman's, yet the summary states that both studies are flawed (Koopman, 2001). Interestingly the staff memo in the cold-rolled case was a general evaluation of the economic studies, and was not directed to any specific Commissioner. The staff memo in the section 201 case, in contrast, was prepared in response to a request by Commissioner Bragg, the individual Commissioner most openly hostile to the use of any economic or econometric analysis. Commissioner Bragg had been the one affirmative vote for the domestic steel industry in the cold-rolled AD case.

In the end, the Commission decided to disregard the economic studies. In its final decision, the Commission returned to its pre-cold-rolled steel approach of just dismissing all the studies. The competing studies were not even written up in the final staff report. It was as if there had been no serious economic and econometric arguments presented.

This outcome was particularly disappointing, since the "battle" among the economists had in fact shed some very useful light on the whole issue of causation and non-attribution. Under either Professor Prusa's or Professor Hausman's models, imports of cold-rolled steel and of galvanized steel just did not matter. In that sense, Professor Hausman's own work in fact validated the conclusions from the earlier AD investigation; imports just did not matter.

The ITC tried to finesse this uncomfortable fact by grouping together as one product all flat-rolled steel products, thus blurring the distinction among hot-rolled steel, cold-rolled steel, and galvanized steel. But these issues are still very much in contention. The ITC decision and the President's ultimate decision to impose section 201 relief on steel imports has triggered widespread outrage around the world. The European Union, Japan, Korea, China, Norway, Brazil, New Zealand, and Switzerland have all filed complaints in the WTO challenging these trade restrictions. So the arguments about causation and non-attribution will continue, although this time in a new and more neutral forum.

6 Conclusions

In retrospect, it is clear that the cold-rolled dumping case represented one of the rare moments when the various factors allowed economic analysis to be seriously considered

as part of an ITC decision. The foreign parties were willing to develop the analysis; the domestic parties did not mobilize effective political opposition at the time; and so the staff could actually consider the merits of competing arguments. This exercise showed that more formal economic analysis can be invaluable in fulfilling the mandate – under both US and international law – to consider carefully the role of imports and to distinguish the role of other factors. Econometric estimates allow both the measuring of the individual impact of imports, and that of alternative causes. Whether in an AD context or a safeguard context, the econometric measurement goes to the heart of the legal obligations.

The great irony in recent practice at the ITC is that the agency embraced the cold-rolled econometric study in an AD investigation, despite a less demanding material injury standard, and despite any guidance from the WTO on non-attribution. Indeed, the Commission was operating without the benefit of the WTO Appellate Body's decisions in wheat gluten, lamb meat, and hot-rolled steel. In contrast, Appellate Body's decisions on wheat gluten, lamb meat, and hot-rolled decisions confronting the Commission in the *Steel* 201 case, and with a more demanding serious injury standard requiring a weighing of causes, the Commission hesitated and disregarded economic models offered by the parties. The telling difference between the two investigations was that the domestic industry had significantly ratcheted up the political pressure on the Commission. Therein lies the dilemma.

Even though the applicability and utility of such an approach seems so obvious, there will be severe political opposition. Political forces that want protection understand full well the dangers of economic and econometric approaches: they tend to limit the flexibility of the ITC or other decision makers. A purely subjective process – with qualitative analysis of the facts on a case-by-case basis – is far more susceptible to political influence (or in extreme cases, to political intimidation). Even though economics would never be more than part of the process – a way to better illuminate the underlying market dynamics – domestic protectionist interests are afraid even to open the door a crack to more objective analysis. The success of economics and econometrics in the cold-rolled dumping case was a step forward, even though we have now taken a few steps back.

But the obligation to distinguish causes remains. National authorities will have to come to grips with this obligation, or face legal challenges to all trade restrictions based on sloppy analysis. The WTO dispute settlement process is slow and cumbersome, but over time it has the potential to modify behavior and shift the terms of the debate. Perhaps after some losses in the WTO trigger economically significant retaliation, national authorities will realize the need to begin more serious analysis of these important issues of causation and non-attribution. When that happens, the cold-rolled dumping case presents a perfect case study of the way that economics can shed useful light on these issues.

Acknowledgments

Both authors are international trade attorneys at Willkie Farr & Gallagher LLP. The authors represented foreign steel producers in the antidumping investigations of carbon steel described in this chapter. This chapter benefited enormously from the collaboration with Professor Thomas

J. Prusa of Rutgers University, both during the two investigations discussed in this chapter and in preparing the chapter itself. We also received helpful comments from Bruce Blonigen, Matt Nicely, and Dan Porter. The views expressed here are personal views of the authors, and should not be attributed to Willkie Farr or its clients.

Notes

1 As various commentators have noted, the DOC almost always finds "dumping." Lindsey (1999) provides a general discussion of this trend.

2 This is the same two-step process involved in countervailing duty investigations, although the issue in countervailing duty cases is subsidization, rather than dumping, and its effects on the domestic industry.

3 Note that the material retardation aspect of the law is rarely used.

4 The ITC may consider "other economic factors as are relevant to the determination regarding whether there is material injury by reason of imports" [19 U.S.C. § 1677(7)(B)(ii) (2000)].

5 As noted by House and Senate lawmakers: "The law does not contemplate . . . that injury from (the subject) imports be weighed against other factors which may be contributing to the overall injury to an industry. Any such requirement has the undesirable result of making relief more difficult to obtain for those industries facing difficulties from a variety of sources, precisely those industries that are most vulnerable to dumped and subsidized imports" (H.R. Rep. No. 96-317, 1979; S. Rep. No. 96-249, 1979).

6 In some sense the Statement of Administrative Action (SAA) is actually at odds with itself, since it first notes that the ITC "need not isolate the injury caused by other factors from injury caused by imports," while at the same time directing the ITC to ensure non-attribution. These are seemingly incompatible points. In the statement that "[E]xisting US law fully implements Articles 3.5 and 15.5 . . ." (H.R. Doc. 103-316, 1994, p. 851), the SAA makes clear, however, that the law is intended to be consistent with the WTO Antidumping Agreement, which explicitly states that other factors of injury "must not be attributed to the dumped imports" (World Trade Organization, 1994, art. 3.5, WTO hereafter).

7 Indeed, some ITC Commissioners have been highly suspicious of econometric models: "[W]e must first advise extreme caution on any reliance on the estimations of the econometric models of the price, production, employment, and consumer cost of removal of the import relief programs. To allow theory to substitute for comprehensive and circumspect analysis of the facts is to gamble with the future of American industries . . .. The appropriate use of econometric models is to supplement this analysis . . .." [*Stainless Steel and Alloy Tool Steel*, 1987, pp. 9–10 (views of Commissioners Eckes, Lodwick, and Rohr)].

8 Consider the 1994 assessment of the US Congressional Budget Office: "The move (from Treasury to Commerce) reflected a Congressional desire for more zealous enforcement of AD/CVD laws and for less concern about their being used in a protectionist manner. Its significance goes beyond the difference in institutional sympathies. One of DOC's functions is to serve as an advocate for US firms. Thus, the move placed responsibility for deciding AD/CVD cases in the hands of an advocate of US parties to the cases" (pp. 27–8). Also, see Baldwin (1985).

9 For statistics on the annual number of investigations between 1980 and 2001, visit the Import Administration's internet web page at http://ia.ita.doc.gov/stats/ad-1980-2001.html.

10 For examples of this output, consider Boltuck (1987) (former International Economist at the ITC, article based on paper prepared for Commissioner Cass); Knoll (1989) (former Legal

Advisor to former ITC Chairman Susan Liebler); and Morkre and Kruth (1989) (Morkre was former Economic Adviser to former ITC Chairman Anne Brunsdale, and Kruth was former Counsel to Brunsdale).

11 Professors Pindyck and Rotemberg of MIT were hired by Chilean exporters to prepare the study in this case.

12 Professor Grossman of Princeton University was commissioned by the FTC to study the steel industry for this investigation.

13 The CADIC model is described in Boltuck (1988) (prepared at the direct request of Commissioner Cass and establishing the economic assumptions and methodologies of the CADIC model). For a comprehensive overview of the model, see Francois and Hall (1997).

14 For an exception, the ITC Office of Economics recently developed a series of empirically based Armington elasticities (Gallaway et al., 2000).

15 Because COMPAS employs constant elasticities, it can in some cases indicate dumping effects on US producers' revenues that exceed the total value of subject imports. To compensate for this characteristic, COMPAS incorporates a "But-For Imports" scenario that imposes an upper limit on the price and volume effects of dumping through the assumption that all subject imports are removed from the US market. The "But-For Imports" scenario sets the total revenue reduction caused by the subject imports equal to the base (input) subject import market share and allocates it between the domestic industry and nonsubject imports according to their base market shares. COMPAS then separately calculates the price and volume distribution of the revenue effects on the domestic industry and nonsubject imports based on the supply elasticity for each. The "But-For Imports" scenario therefore reflects the implicit assumption that demand for the subject product is perfectly inelastic to price. In addition to the vertical nature of the demand curve in the "But-For-Imports" scenario, the assumption of constant elasticities means the COMPAS model also rigidly sets the positioning of the demand curve solely according to producers' output. The COMPAS model also provides an invariant supply curve that may or may not have an upward slope, and that is also restricted in terms of positioning. This assumption ignores other determinants of demand and supply that could shift the demand curve or the supply curve to either the left or right.

16 Some individual commissioners, however, embraced analytic approaches that qualitatively went through a similar type of exercise, the most recent example being the analytical framework views expressed by former Commissioner Carol Crawford (*Melamine Institutional Dinnerware from China, Indonesia, and Taiwan*, 1997, pp. 31–2).

17 For overviews, see Knoll (1989); and Morkre and Kruth (1989).

18 In *Certain Cameras*, the ITC found that imports of cameras had increased over the period of investigation and that the petitioner, Keystone Camera, was seriously injured, but imports were not a substantial cause of that injury (1990, p. 3). The Commission was persuaded by compelling facts during the investigation that led it to conclude that "several decisions of Keystone's prior management are more important than increased imports in explaining the serious injury or threat of serious injury," and in particular certain poor investments and an inability to adapt to changes in the camera market (pp. 37–41).

19 The Appellate Body ruling in the wheat gluten case would constitute a more full development of its earlier holding on the same issue (*AB Decision on Footwear*, 1999). In its ruling in *Footwear*, the Appellate Body upheld a basic proposition that an examination of factors operating in the market other than imports must be conducted, so that any injury caused by such other factors can be identified and properly attributed, pursuant to the last sentence of Article 4.2(b) (1999, para. 136).

20 This somewhat confused language on non-attribution was the product of the Appellate Body's wrangling over a related matter – whether imports must be a sole cause of serious

injury. The dispute settlement panel hearing the case had ruled that imports must be a sole cause of serious injury in what was a logically and intellectually defensible interpretation of the Safeguards Agreement. The Appellate Body disagreed, leaving Members with an ill-defined "genuine and substantial relationship."

21 Citing *Lamb Meat* (1999, p. I-24).

22 These six other factors were: the cessation of subsidy payments under the National Wool Act of 1954; competition from other meat products, such as beef, pork, and poultry; increased input costs; overfeeding of lambs; concentration in the packing segment of the industry; and a failure to develop and maintain an effective marketing program for lamb meat.

23 According to the Appellate Body, these four other factors were: cessation of the payments under the National Wool Act of 1954; competition from other meat products; concentration in the packer segment of the industry; and the failure to develop and implement an effective marketing program.

24 Three Commissioners found present serious injury, and two other Commissioners found threat of serious injury.

25 The DOC's procedural history is contained in the ITC report.

26 These factors were: (1) the increase in production capacity of minimills; (2) the effects of a strike at General Motors in 1998; (3) declining demand for hot-rolled steel from the US pipe and tube industry; and (4) the effects of prices of non-dumped imports.

27 The Appellate Body declined to rule on the factual issues raised concerning the ITC's consideration of the four "other" factors because they were not adequately addressed by the dispute settlement panel below (*AB Decision on Hot-Rolled Steel*, 2001, p. 75). However, the ruling made clear that had the full factual record been before the Appellate Body, the US antidumping action would have likely fallen on these grounds. Indeed, the United States essentially admitted it had applied the wrong legal standard to the facts. It was instead found invalid on other grounds (pp. 79–81).

28 Like all public documents filed in ITC proceedings, this study can be found at www.usitc.gov in EDIS. One merely needs to identify the case type and investigation number. See also Prusa and Sharp (2001). The authors both worked with Professor Prusa in developing the model and data sources, as private attorneys representing the foreign exporters in this case.

29 The study also considered alternative specifications based on import volumes rather than import prices, but reached essentially the same conclusions.

30 In an AD context, the difference between those imports "subject" to a case and those not named by the petition is crucial from a legal perspective.

31 This discussion draws heavily on the discussion of the framework provided in Professor Prusa's study.

32 The first two inputs were used by integrated mills, although some minimills use forms of processed iron ore in their production. Scrap is used as the primary input for minimills, although integrated mills sometimes use them in their production process. In a more sophisticated model, one would formally distinguish between minimills and integrated mills. But for purposes of this study, we treated both as domestic firms. In subsequent work in other cases, particularly a safeguard investigation of steel discussed later, the authors and Professor Prusa presented models that measured the separate effect of intra-industry competition between minimills and integrated mills on domestic price levels.

33 Although not an economic factor, we note that in previous AD cases the ITC has consistently valued captive consumption of hot-rolled steel at the market price, indicating that the market price of hot-rolled is precisely the appropriate economic value of hot-rolled output. In other words, the price of hot-rolled exactly captures the opportunity cost of shipping a ton of hot-rolled downstream.

34 This discussion also draws heavily on the discussion of the econometrics provided in Professor Prusa's study. Additional discussion can be found in Gujarati (1995, pp. 678–703) and Kmenta (1997, pp. 651–730).
35 Basically, this econometric method systematically creates instrumental variables to replace the endogenous variables where they appear as explanatory (or right-hand side) variables in a simultaneous system. More specifically, the two-stage procedure works in the following manner: first, an OLS regression is run on the reduced-form equations for each of the endogenous variables that appear as explanatory variables in the structural equations in the system. This procedure will produce unbiased coefficients that can then be used to calculate estimates for each of these endogenous–explanatory variables. Second, the reduced-form estimates for the endogenous–explanatory variables are substituted back into the structural equations and then these revised structural equations are estimated with OLS.
36 This approach assumes that the product is homogenous, and that the industry is competitive. For the cold-rolled steel industry, these assumptions are generally realistic. There are techniques for studying differentiated products, and oligopolistic industry structures, but they are more complex and even more data intensive.
37 The R2 of the model is another commonly used measure of "goodness of fit." Broadly stated, R2 indicates how well the model does predicting the price of cold-rolled steel. In all specifications the R2 was quite high, indicating the model is useful for trying to explain what has happened and for predicting what would have happened.
38 While nonsubject countries as a group played a small role in determining the price for cold-rolled steel, the single largest exporter in that group, South Korea, had an important effect on cold-rolled prices. When isolated, the model showed that the overall $8/ton impact was better interpreted as the weighted average of Korea's impact ($13/ton) and other nonsubject importers' impact ($5/ton).
39 In alternative specifications, the increase in the volume of subject imports caused the domestic price of cold-rolled steel to fall by only $1/ton, also statistically insignificant.
40 The study used the price of subject and nonsubject imports lagged four months in order to capture the fact that delivery of imports takes about four months. The issue of lag choice is important but space limitations limit our discussion. A more complete discussion of the lag length and the robustness of the results to alternative lag lengths is found in the study (Prusa, 2000, pp. 34–56).
41 As explained in the memo, "I reproduced the results, performed sensitivity analysis, and found the main estimation results to be robust. Overall, considering my econometric analysis of the data, the Respondent's response to my questions, and Petitioner's concerns, I have determined that the model is well-specified and the results defensible" (McDaniel, 2000).
42 We conducted a LEXIS search of ITC decisions, and could find only 12 examples where "regression" is even mentioned.
43 The footnote provided: "In addition, respondents produced an econometric model, the results of which suggest that the subject imports had little impact on domestic prices for cold-rolled steel. Joint Respondents' Pre-hearing Brief, Vol. II. We closely examined the model and note that the results are consistent with our findings, and also consistent with the findings of the COMPAS model" (Cold-Rolled Steel, 2000, p. 23, fn. 185).
44 See generally, Post-hearing Brief of Bethlehem Steel Corporation, et al., 2000, pp. 3–15.
45 Possibly no domestic industry is more responsible for existing law and practice under the US trade remedy regime than the US steel industry. For a detailed account of the US steel industry's 30-year campaign of protectionism and efforts to capture US trade law, see Barringer and Pierce (2000).
46 Similar commentary can be found in Kelly, "Economic Heat" (2000).

47 Since Hausman's model using firm specific data for one minimill and one integrated mill, much of his report is business proprietary; although the public version is available for inspection at EDIS online (Pre-hearing Brief on behalf of Bethlehem Steel, et al., 2001, Exhibit 4).
48 These studies can be found in each of the referenced Joint Pre-hearing Briefs submitted in USITC Inv. TA-201-73. All of these briefs are available at EDIS online.
49 The models also included specifications for import volume. For cold-rolled and galvanized steel, volume never mattered. For hot-rolled, the effect of volume depended on whether one included the substantial surge in 1998 in the period being considered.
50 For cold-rolled and galvanized, the coefficients measuring the price and volume effects were statistically insignificant. For hot-rolled, the effect of import prices was sometimes statistically significant and sometimes not.

References

Baldwin, Robert 1985: *The Political Economy of US Import Policy*. MIT Press.

Barringer, William H. and Kenneth J. Pierce 2000: *Paying the Price for Big Steel*. American Institute for International Steel.

Boltuck, Richard 1987: An economic analysis of dumping. *Journal of World Trade Law*, 21, 45–54.

Boltuck, Richard 1988: *Assessing the Effects on the Domestic Industry of Price Dumping, Part II*. US International Trade Commission, Office of Economics, Research Division.

Francois, Joseph E. and H. Keith Hall 1997: Partial equilibrium modeling. In Joseph E. Francois and Kenneth A. Reinert (eds.), *Applied Methods for Trade Policy Analysis*. Cambridge University Press, 135–53.

Gallaway, Michael P., Christine A. McDaniel, and Sandra A. Rivera: *Industry-Level Estimates of US Armington Elasticities*. Office of Economics Working Paper No. 2000-09-A, US International Trade Commission, ftp://ftp.usitc.gov/pub/reports/studies/ec200009a.pdf.

Grossman, Gene M. 1986: Imports as a cause of injury: The case of the US steel industry. *Journal of International Economics*, 20, 201–23.

Gujarati, D. 1995: *Basic Econometrics*, 3rd edn., McGraw-Hill/Irwin 678–703.

Kelly, Nancy 2000: Economic heat behind cold-rolled case. *American Metal Market*.

Kelly, Nancy 2000: Potomac pulse – steel lobbyists, lawyers growing snappish. *American Metal Market*.

Kmenta, J. 1997: *Elements of Econometrics*, 2nd edn., University of Michigan Press, 651–730.

Knoll, Michael S. 1989: An economic approach to the determination of injury under United States Antidumping and Countervailing Duty Law. *New York University Journal of International Law and Politics*, 22, 37–116.

Koopman, Robert B. 2001: Director, Office of Economics, *Memorandum to Commissioner Lynn Bragg*, in the matter of *Steel*, Inv. No. TA-201-73, Memorandum EC-Y-042.

Lindsey, Brink 1999: *The US Antidumping Law: Rhetoric Versus Reality*. CATO Institute Trade Policy Analysis No. 76, http://www.freetrade.org/pubs/pas/tpa-007es.html, April 2002.

McDaniel, Christine 2000: Research Division, Office of Economics, US International Trade Commission *Memorandum to the Commission*, in the matter of *Certain Cold-Rolled Steel Products from Argentina*, et al., Inv. Nos. 701-TA-393 and 731-TA-829–30, 833–4, 836 and 838 (Final).

Morkre, Morris E. and Harold E. Kruth 1989: Determining whether dumped or subsidized imports injure domestic industries: International Trade Commission approach. *Contemporary Policy Issues*, 7, 78–95.

Pindyck, Robert S. and Julio J. Rotemberg 1987: Are imports to blame? Attribution of injury under the 1974 Trade Act. *Journal of Law and Economics*, 30, 101–22.
Prusa, Thomas J. and David C. Sharp 2001: Simultaneous equations in antidumping investigations. *Journal of Forensic Economics*, 14, 63–78.

Administrative and Executive Materials

Certain Cameras, Inv. No. TA-201-807, USITC Pub. 2315, 1990.
Certain Cold-Rolled Steel Products from Argentina, et al., Inv. Nos. 701-TA-393 and 731-TA-829–30, 833–4, 836 and 838 (Final), USITC Pub. 3283, 2000.
Certain Hot-Rolled Steel Products from Japan, 64 Fed. Reg. 33,514 (Int'l. Trade Comm. 1999) (final invest.).
Circular Welded Carbon Quality Line Pipe, Inv. No. TA-201-70, USITC Pub. 3261, 1999.
H.R. Rep. No. 96-317, 1979.
H.R. Doc. 103-316, vol. 1, 1994.
H.R. Rep. 106-680, 2000.
Lamb Meat, Inv. No. TA-201–68, USITC Pub. 3176, 1999.
Melamine Institutional Dinnerware from China, Indonesia, and Taiwan, Inv. No. 731-TA-741–3 (Final), USITC Pub. 3016, 1997.
Post-hearing Brief on behalf of Bethlehem Steel Corporation, Ispat-Inland Inc., LTV Steel Company, Inc., National Steel Corporation, and US Steel Group, submitted January 28, 2000, in the matter of *Certain Cold-Rolled Steel Products from Argentina*, et al., Inv. Nos. 701-TA-393 and 731-TA-829–30, 833–4, 836 and 838 (Final), 3–15.
Post-hearing Brief on behalf of Joint Respondents – Volume 1: Legal and Factual Argumentation, submitted October 1, 2001, in the matter of *Steel*, Inv. No. TA-201-73, 116–19.
Pre-hearing Brief on behalf of Bethlehem Steel, LTV Steel Company, Inc., National Steel Company, and US Steel LLC, submitted September 10, 2001, Vol. II of IV, in the matter of *Steel*, USITC Inv. No. TA-201-73.
Pre-hearing Brief on behalf of Joint Respondent Producers in Argentina, Brazil, European Union, India, Japan, Korea, Russia, Taiwan, Thailand, and Venezuela: Carbon and Alloy Flat Products, Product 6 – Corrosion-Resistant and Other Coated Sheet and Strip, submitted September 10, 2000, in the matter of *Steel*, USITC Inv. No. TA-201-703 (filed by the law firm of Sharrets Paley).
Pre-hearing Brief on behalf of Joint Respondents: Hot-Rolled Steel (Product Category A.3), submitted September 10, 2000, in the matter of *Steel*, USITC Inv. No. TA-201-703 (filed by the law firm of Kaye Scholer).
Pre-hearing Brief on behalf of Joint Respondents: Product Group 4, Cold-Rolled Steel, submitted September 10, 2000, in the matter of *Steel*, USITC Inv. No. TA-201-703 (filed by the law firm of Willkie Farr).
Proclamation 7103, 63 Fed. Reg. 30,359, 1998.
Proclamation 7208, 64 Fed. Reg. 37,387, 1999.
Prusa, Thomas, J., *Econometric Study of Thomas J. Prusa, Ph.D*, Joint Pre-hearing Brief of Respondents, submitted to the International Trade Commission in the matter of *Cold-Rolled Steel Products from Argentina*, et al., Inv. Nos. 731-TA-829–40 (Final), January 2000.
Rockefeller, Senator John D., in the matter of *Tin- and Chromium-Coated Steel Sheet from Japan*, International Trade Commission Hearing Transcript, Washington, DC, June 29, 2000, 55–7.
S. Rep. No. 96–249, 1979.
Stainless Steel and Alloy Tool Steel, Inv. No. TA-203-16, USITC Pub. 1975, 1987.

US Congressional Budget Office 1994: *How the GATT Affects US Antidumping and Countervailing-Duty Policy*.
US Department of Commerce, Import Administration: Antidumping and Countervailing-Duty Statistics, http://ia.ita.doc.gov/stats/iastats1.html.
US International Trade Commission, Witness List, http://www.usitc.gov/steel/wl2_0917.pdf.
Wheat Gluten, Inv. No. 201-TA-67, USITC Pub. 3088, 1998.

Intergovernmental Materials

Argentina – Safeguard Measures on Imports of Footwear – AB-1999-7 – Report of the WTO Appellate Body (AB Decision on Footwear), WT/DS121/AB/R, 1999.
United States – Definitive Safeguard Measures on Imports of Wheat Gluten from the European Communities – AB-2000-10 – Report of the WTO Appellate Body (AB Decision on Wheat Gluten), WT/DS166/AB/R, December 2000.
United States – Anti-Dumping Measures on Certain Hot-Rolled Steel Products from Japan – AB-2001-2 – Report of the WTO Appellate Body (AB Decision on Hot-Rolled Steel), WT/DS184/AB/R, 2001.
United States – Definitive Safeguard Measures on Imports of Circular Welded Carbon Quality Line Pipe from Korea – AB-01-5229 – Report of the WTO Appellate Body (AB Decision on Line Pipe), WT/DS/202/R, 2001.
United States – Safeguard Measures on Imports of Fresh, Chilled or Frozen Lamb Meat from New Zealand and Australia – AB-2001-1 – Report of the WTO Appellate Body (AB Decision on Lamb Meat), WT/DS177/AB/R, 2001.
World Trade Organization 1994: *Agreement on Implementation of Article VI of the General Agreement on Tariffs and Trade 1994*.

CASE LAW

Altx, Inc. v. *United States*, 167 F. Supp. 2d 1353, CIT 2001.

Statutes

19 U.S.C. § 1673.
19 U.S.C. § 1677.
19 U.S.C. § 2251.
19 U.S.C. § 2252.
19 U.S.C. § 2253.

4

Trade-Related Labor and Environment Rights Agreements?

Chantal Thomas

Chapter Outline

This chapter examines the World Trade Organization's (WTO's) "institutional competence" justification for declining to pursue the adoption of affirmative labor and environmental standards. The WTO's incorporation of affirmative intellectual property (IP) standards, in the form of the WTO Agreement on Trade-Related Property Rights (TRIPS), suggests that alternatives to the WTO's current "institutional competence" exclusion of labor and the environment are possible. A case can be made that labor and environmental standards should receive the treatment that IP standards have received within the WTO* namely, formalization, and adoption in WTO Agreements. Labor and environmental standards resemble IP standards in terms of structural relationship to trade flows. Moreover, the reasons for having incorporated IP standards into the WTO (* weak institutionalization and chronic underenforcement in the pre-existing IP regime) also arguably apply to labor and environmental standards under their respective regimes. Finally, conventional distinctions between IP law, labor law, and environmental law, when used to justify the differential treatment of IP law in the WTO, fail to withstand scrutiny. All of these observations point to political will as the barrier to incorporation of affirmative labor and environmental obligations into the international trade regime. To point to political will is not to discount other challenges confronting "linkage" of trade, labor and environmental concerns (such as the cost of specifying broadly phrased labor and environmental obligations) nor is it to insist that the WTO is in fact the best forum for the adjudication of labor and environmental standards. Rather, it is to underscore that the ability, or inability, to overcome these challenges, and to sort out the institutional questions that accompany them, can be traced to the resources made available within and among states for the legal resolution of these issues. The example of the WTO TRIPS Agreement confirms this causality.

1 Introduction

The World Trade Organization (WTO) has stated that the application of affirmative environmental and labor obligations falls outside of its jurisdiction as a body concerned with trade.[1] The argument is one of institutional competence, according to WTO literature: "Other agencies that specialize in environmental issues are better qualified to undertake those tasks"[2]; and, regarding labor issues, "the International Labour Organization (ILO) [is] the competent body to deal with labour standards."[3]

These statements respond to growing requests, mainly by nongovernmental organizations (NGOs), that the WTO affirmatively incorporate labor and environmental standards into its regulation of international trade.[4] Similar pressures at the WTO's inception in 1995 led to the identification of both labor and the environment as issues requiring further study.[5] Despite vocal protests at subsequent meetings of the WTO – particularly at the 1999 Ministerial Conference in Seattle[6] – the WTO has not moved beyond the "institutional competence" argument for excluding labor and environmental standards from the WTO.[7]

This chapter examines the WTO's institutional competence justification for deflecting labor and environmental concerns. The chapter compares the cases for *including* the trade-related field of intellectual property (IP) rights in the WTO with those *excluding* the trade-related fields of labor and environmental rights.[8] While there may be formal legal problems facing the incorporation of labor and environmental standards,[9] the example of IP is instructive for the "trade and . . ." debate.[10]

Section 1 analogizes labor and environmental issues to IP, which is covered by the WTO Agreement on Trade-Related Intellectual Property Rights (TRIPS).[11] Section 2 briefly demonstrates that labor and environmental law, like IP law, are trade-related. Section 3 observes that IP law was included in the WTO due to chronic underenforcement in the previous system and notes that problems of underenforcement also affect international labor and environmental law. In Section 4, the chapter considers what policy arguments might justify including IP law but not labor and environmental law despite their important similarities. The chapter concludes by considering issues of political will that confront the "linkage effort," with special attention to developing-country government resistance to linkage.

2 The Trade versus Nontrade Distinction

A primary objective of the WTO's establishment in 1995 was to extend trade discipline to areas of international economic activity that had been excluded under the previous international trade regime of the General Agreement on Tariffs and Trade (GATT). The WTO retained the existing regime defined by the GATT and significantly expanded it. For example, agreements on IP – TRIPS – and trade in services – the General Agreement on Trade in Services (GATS) – resulted from this effort.

The rationale for a WTO Agreement regulating trade in services is clear: services are a category of trade, just as goods are. By the time the Uruguay Round began in

the 1980s, the importance of internationally provided services justified the inclusion of services in the regulatory scope of international trade law, an area previously concerned only with goods.

The rationale for a WTO Agreement regulating IP is quite another matter. Although IP rights can be traded, in licensing and other rights transferring agreements, the main concern of TRIPS is the way in which IP rights are *affected* by trade, whether traded internationally or traded at all. The value of the IP right increases with the capacity of the right-holder to enforce it and decreases when the information can be accessed without going through the right-holder. The latter may occur because the producer is operating in a region where the right does not exist or is underenforced. It is this relationship between trade and IP rights that caused inclusion of IP rights in the WTO. Indeed, the title of the agreement – *Trade-Related Intellectual Property Rights* – indicates the relationship.[12]

A very similar relationship exists between trade issues and labor and environmental issues. Trade flows are undeniably bound up with labor and environmental protections because labor and environmental standards can affect production costs for goods and services in many ways. Just as the value of an IP right will decrease if it can be evaded through trade, the value of a labor or environmental right will also decrease if evasion through trade is an option. Moreover, the rights themselves may erode: a phenomenon known as the "race to the bottom."[13]

Without considering TRIPS, the above trade versus nontrade distinction might appear to justify exclusion of labor and environmental standards. Considered in isolation, the institutional competence argument that the WTO should not incorporate labor and environmental issues because they are not themselves the subject of trade, even if they clearly influence and are influenced by trade, appears more persuasive. Upon consideration of TRIPS, however, this distinction finds itself on very shaky ground. TRIPS concerns itself primarily with IP rights not as a *subject* of trade, but as a body of standards that *affect* trade flows. In this respect, international IP law is difficult to distinguish from international labor or environmental law in its relation to trade. Given these similarities between IP rights on the one hand, and labor and environmental rights on the other, the question arises as to why one set of rights has been successfully included in the WTO while the others have not.[14]

3 The Problem of Underenforcement Common to the IP, Labor, and Environmental Regimes

The underenforcement of IP rights served as a major justification for integrating them into the WTO.[15] Prior to the Uruguay Round, multilateral instruments for IP protection were administered by the World Intellectual Property Organization (WIPO), a specialized agency of the United Nations.[16] WIPO was widely viewed as ineffectual, however.[17] The relative weakness of the WIPO system, for example, provided a partial justification for US development of forceful unilateral mechanisms to pursue its interests in international IP law, such as the "section 301" mechanism for monitoring foreign IP laws and enforcement.

As the importance of IP-related sectors to Western economies grew, particularly in the United States, the integration of IP rights into the international trade regime presented a solution to the problem of weakness in the freestanding WIPO system.[18] The incorporation of IP into the international trade regime had the potential for three salutary changes. First, the degree of uniformity in IP protections would greatly increase as a result of the accession of all GATT/WTO members to a single undertaking that would include IP.[19] Second, the trade regime's dispute settlement system could address international IP disputes.[20] Third, cross-sector links in the larger system would improve the efficacy of both negotiations and enforcement. For example, if a complainant prevailed on IP grounds, it could impose sanctions on the respondent in non-IP sectors.[21]

Thus, introduction of IP rights into the trade regime was viewed as a way to improve significantly the efficacy of an international IP law system characterized by lax enforcement. The problem of underenforcement is pervasive in international law, however, and in this respect the international IP regime differed little from many other international systems. Certainly, the same problem of underenforcement characterizes international law relating to both labor and the environment.

3.1 Underenforcement of International Labor Law

Established in 1919, The International Labor Organization (ILO) predates the United Nations by several decades, although it now operates under the auspices of the United Nations.[22] The ILO has coordinated more than 180 binding conventions and 185 nonbinding resolutions on subjects ranging from human rights related to labor, to occupational safety, and to specific economic sectors. From the standpoint of enforcement, however, the ILO's record has proven woeful.

The ILO's provisions for enforcement center on fact-finding and reporting. As with IP law prior to TRIPS, compliance with ILO principles is spotty. A 1996 study undertaken by the Organization for Economic Cooperation and Development (OECD)[23] found that, in general, rich countries complied with "core" labor principles (with the notable exception of the finding that the US did not ensure the freedom to engage in collective bargaining), while most poor countries did not.[24]

The lack of uniformity of commitments that characterized international IP law prior to TRIPS still characterizes international labor law in many respects. Differing groups of countries ratify different treaties, creating a patchwork of legal obligations, although it is true that those conventions dealing with "core" labor standards tend to have more uniform ratifications.[25]

3.2 Underenforcement of International Environmental Law

The same problem of underenforcement plagues international environmental law. The institutional scenario is more complicated than that of international labor law: a variety of organizations abound, each of which is in charge of its specific area. While

the number and scope of international economic agreements continue to grow, no institutional mechanism has emerged for the effective enforcement of environmental obligations. The Stockholm Declaration of the United Nations Conference on the Human Environment[26] provided the basis for establishment of the United Nations Environment Program (UNEP), which comes the closest to a central international agency responsible for international environmental law.[27] Yet, its institutional shortcomings are manifold. The UNEP plays a "coordinating and catalytic role," but at the same time "lacks any formal powers."[28] It has no authority to require states to cooperate with its efforts to gather information or to further the progressive development of international environmental law.[29]

Many calls have gone out to improve the institutional basis laid down by the UNEP.[30] Most significantly, in 1992 United Nations Conference on Environment and Development (UNCED) undertook a major review and expansion of international environmental law and sought to lay out a comprehensive blueprint for its development and implementation.[31] While many of the resulting documents were nonbinding, UNCED did generate some formal instruments, an example of which is the Convention on Biological Diversity.[32] Although this Convention benefits from formally binding status, its obligations are generally hortatory, qualified as they are by the phrase "as far as possible and as appropriate."[33] The Convention's institutional mechanisms are sparse.[34] In a communication to the WTO, the Secretariat of the Convention conceded, "The Convention does not have a compliance procedure. Formal assessment of Parties or non-Parties compliance with the Convention has not occurred."[35]

Perhaps the best institutionalized international environmental agreement is the Montreal Protocol on Substances that Deplete the Ozone Layer,[36] because it administers a fund to assist countries in achieving compliance.[37] The Montreal Noncompliance Procedure (NCP)[38] allows disputes between parties to be submitted to an Implementation Committee, which then recommends "appropriate action" to the parties.[39] The NCP has been sparingly used so far, however.[40] Overall, then, resources that exist for enforcement of international environmental law remain sparse.[41]

Thus, without considering TRIPS, the trade versus nontrade argument might serve to justify exclusion of labor and environment standards. Upon consideration of TRIPS, however, this distinction weakens: TRIPS concerns itself primarily with IP rights not as a subject of trade, but as a body of standards that affect trade flows. In this respect, international IP law is structurally analgous to international labor or environmental law in its relation to trade. Moreover, the institutional reasons for incorporating IP – primarily, the problem of underenforcement of international IP law – can also be attributed to international labor and environmental law and therefore argue for their inclusion.

4 Line-Drawing Revisited: Is There Any Good Reason for Excluding Labor and the Environment?

In the shadow of TRIPS, then, justifications for excluding labor and environmental standards must become significantly more nuanced. Observations about the characteristics of all of the bodies of international law under consideration – trade, IP, labor, and

environmental – must draw analogies and distinctions sufficient to justify including one body of law (IP) and excluding others (labor and environmental). In particular, labor and environmental law must be satisfactorily distinguished from trade and IP law to justify their relegation to other realms, that is, there must be special reasons why the WTO is less competent to administer labor and environmental rights, as compared to IP rights.

This section considers what arguments might be made for excluding affirmative labor and environmental obligations from the WTO. It concludes that, upon examination, some of the arguments that might be conventionally made contain serious flaws.

4.1 "Red-Herring" Distinctions

4.1.1 Public–Private

An argument often heard in the past for excluding labor and environmental considerations from the WTO is that the WTO is "private" in orientation. One version of the public–private distinction is that international trade law falls under the rubric of "private international law," whereas labor and environmental standards more appropriately belong to "public international law."[42] A distinction that characterizes GATT/WTO law itself as "private" rather than public, however, is puzzling at best. "Private law" is conventionally understood and accepted to mean legal rules that primarily concern themselves with individual transactions between parties.[43] The law governing contracts, whether domestic or international, is a prime example of such law.[44] By contrast, GATT/WTO law consists primarily of rules governing regulatory measures taken by states.

It is true that those regulatory measures in turn affect individual trading activity. GATT/WTO law itself, however, imposes obligations on states, not individuals. Thus, whereas the Convention on the International Sale of Goods (CISG)[45] might appropriately be characterized as private international law, under the prevalent conception of private law, GATT/WTO law may not. GATT/WTO law is exactly the same as traditional public international law in that it concerns itself with the behavior of states.

To the limited extent the "public–private" distinction is useful, then, it serves as shorthand for differences in orientation and in institutional characteristics: the difference in institutional characteristics can be described as a difference between *pragmatism* and *legalism* and the difference in orientation can be described as a difference between *market* and *nonmarket* or between economic and social. These somewhat more nuanced distinctions, however, suffer from similar analytical and descriptive shortcomings.

4.1.2 Pragmatism versus Legalism

The distinction in perceived institutional characteristics traditionally rested in the pragmatism of the GATT/WTO system as compared to the legalism of public international law.[46] According to the accepted view, for much of the postwar era the GATT deliberately set expectations for compliance low, chose not to expose or respond to blatant departures from GATT discipline, and refused to push states any further than

they were willing to go in developing international trade obligations.[47] Thus, whereas public international institutions explicitly aspired towards universal, and universally obeyed, standards, the GATT was content with compliance that reflected the will of the states.[48] Whereas shortfalls in compliance with public international law were seen as manifest failures of the system, such shortfalls in the GATT/WTO system were essentially accepted as legitimate. This pragmatic approach was part of a self-consciously sophisticated and "cosmopolitan" approach to international law and policy.[49]

This narrative may have partially emerged in hindsight and may therefore seem overdetermined. There is significant truth, however, to the notion that pragmatic institutional values were explicitly in circulation in the GATT at the time. For instance, Olivier Long, who headed the GATT in the 1970s, wrote a testimonial to this concept.[50]

Although this distinction might have underwritten a distinction between international trade law and other, "general" international law in the past, however, it breaks down almost completely in the present. The Uruguay Round of GATT negotiations, which led to the establishment in 1995 of the GATT's successor organization, the WTO, marked a decisive break with the pragmatic strategy of the past.[51] The crowning glory of the WTO is its dispute settlement system, which remains the strongest nonregional multilateral forum by a considerable measure.[52] Rather than taking a voluntaristic approach to compliance, the new international trade regime adopts an enforcement mechanism of unprecedented rigor.[53] As a consequence, usage of the system has dramatically increased.[54]

Indeed, this newly rigorous dispute settlement system is the very mechanism that has prompted appeals to incorporate labor and environmental law. The appeals contain both positive and negative dimensions. On the positive side, the view is that an institution of the WTO's authoritativeness will more effectively promote labor and environmental standards than any other available mechanism.[55] On the negative side, the fear is that without integration, the institutional imbalance will lead to the enforcement of trade standards at the expense of labor and environmental standards.[56]

Ironically, the public international law system now seems far more pragmatic than does the international trade regime. Whereas the WTO regularly and vigorously employs its dispute settlement system, the primary enforcement strategy for public international standards (like labor and the environment) is the so-called "soft law" approach.[57] This approach is characterized by the slow accretion of obligations through customary international law,[58] together with attempts to gain state compliance by influencing public opinion through media campaigns,[59] or through nonbinding consultation and advice.[60]

Thus, the distinction between a pragmatic private approach and a legalistic public approach not only has ceased to be accurate, but also has arguably reversed. If the public international law system has historically been associated with an aspiration toward a formal and consistent ordering of states and norms,[61] in which interstate disputes are resolved by some judicial or quasi-judicial process,[62] then the WTO is more "public" than most public institutions.[63] The WTO itself has begun to recognize that it is more of a public institution, for instance, by explicitly incorporating standards of treaty interpretation from public international law.[64]

4.1.3 Market versus nonmarket or economic versus social

The public–private distinction signaled that international trade law, although it focused on state behavior, was ultimately concerned with *market* activity. By contrast, public international law focused on *social* activity. To the extent that such a distinction can be maintained in any context, it breaks down when considering labor and environmental standards, both of which directly affect and concern market activity. The affected market elements include labor conditions in the marketplace, the environmental consequences of production, and many others.

Any distinction between private market activity and public international law concerns seems hard to maintain in the area of IP rights. IP rights affect the cost and nature of production just as labor and environmental standards do. Moreover, IP rights, labor rights, and environmental rights not only all affect market activity, but also implicate questions of social justice. The distinction does not seem to fit the controversy over accessibility to pharmaceutical drugs in poor countries, particularly those facing HIV and other epidemics.[65]

4.1.4 Open versus closed forum

A related argument is that the WTO forum is less *appropriate* for public decisions than the United Nations and other public bodies. The membership of both the United Nations and the WTO includes mostly governments, but the United Nations has had more participation and input from NGOs.[66] One could argue that the proceedings of UN sessions are more accessible to the public as they occur – i.e. that it is easier for the general public to get information about and respond to UN sessions as they occur. Even if this is true, however, the distinction does not clearly justify keeping labor and environmental concerns out of the WTO, unless one endorses the image of the WTO as a "technocratic" body that engages in rulemaking outside the scope of public attention and concern.[67] The WTO itself seems to be consciously rejecting that criticism, especially in public relations fora such as its website, which "provides an opportunity for the public to comment on the WTO, its activities, and the trading system."[68]

4.1.5 Rigidity versus flexibility

A final distinction departs from any appeal to structural or substantive logic, but appeals purely to pragmatic considerations. Particularly in the wake of the WTO's newfound institutional rigor, many commentators worry about the increased likelihood that the international trade regime will reach a "breaking point."[69] The new WTO sets much loftier ambitions for institution building than did the old GATT.[70] Commentators worry that, while the GATT was flexible and therefore never broke, the WTO may be overly rigid.[71] It may overshoot the mark in seeking to establish a "legalistic" regime.[72]

The result might be a backlash of states that would erode the precious legitimacy that the global regime has built up over the postwar era.[73]

The jury is still out as to whether the WTO is institutionally viable over the long run. The dispute settlement system garnered early acclaim by resolving contentious issues between powerful members with relative success.[74] The first new agreement-making meeting in Singapore went relatively smoothly and succeeded in making anticipated gains.[75] More recently, however, the highly visible failures of the Seattle and Cancun meetings revive the old institutional concerns.[76] Seattle showed that state sovereignty still had bite – the Member States, if they wanted, could refuse to go along with the built-in agenda and could stymie the progressive development of the institution.[77] Thus, the anticipated progress in negotiations on services, agriculture, and other areas failed to materialize.[78]

Particularly in the wake of Seattle, one might argue that at some point institutional modesty is required. The addition of items to the negotiation agenda would be too taxing. There may be some validity to this perception. However, it is not one that justifies the special exclusion of labor and environmental concerns, as opposed to other issues that are placed on the agenda. For example, negotiations on competition policy would require reconciling many different and complex regulatory regimes, but WTO members continue to consider the possibility of framing an agreement on competition. At the conclusion of their 1996 meeting in Singapore, WTO members agreed to "establish a working group to study issues . . . relating to the interaction between trade and competition policy . . . in order to identify any areas that may merit further consideration in the WTO framework."[79] Controversy and disagreement run high over the matter and formal efforts to frame an agreement have not begun.[80]

4.2 Active Distinctions

If conventional boundary-drawing attempts for excluding labor and environmental problems from formal international scrutiny fail, the inquiry must turn to underlying dynamics of the WTO. These following sections argue that the ongoing exclusion of labor and environmental standards has much to do with the perception that labor and environmental standards generally constrain, rather than expand, trade. Ultimately, the exclusion of these standards reflects not only the basic policy orientation of the WTO, but also the political will of the WTO Member States.

4.2.1 Expanding versus Controlling International Trade

The expansion of international trade is the preeminent objective of the GATT/WTO.[81] By contrast, international labor and environmental laws, in many instances, are concerned with reigning in and controlling market activity in order to ensure the fulfillment of certain social and political objectives.

From the perspective of Western actors, the protection of IP rights was a necessary prerequisite to the expansion of international trade.[83] The absence or weakness of IP

enforcement in many poor countries was a major impediment to designs by Western actors in IP-related industries to expand sales and production internationally.[82] This perspective likens IP rights to traditional property rights and asserts that their enforcement is a necessary prerequisite to efficient market function.[84] Of course, one can counterargue that IP enforcement does not expand, but rather impedes, international trade. This argument rejects the analogy of IP rights to traditional property rights, and instead likens IP rights to monopoly rights.[85] As monopoly rights, IP rights tend to exclude poor-country actors from production. Since the monopoly keeps prices high, world demand, and therefore world trade, in these sectors remains lower than it would be under greater competition.[86] In fact, there is a real rift in perspective that reflects the North–South divide (further discussed in the next section), with the North, particularly the United States, pushing to include IP rights in the WTO, and the South losing faith in the entire system.[87]

Despite this divide, the view of IP protection as a facilitator rather than an inhibitor of international trade prevailed, culminating in the TRIPS agreement.[88] When viewed as a facilitator of international trade, the logic for including international IP law becomes clear: to further the central objectives of the GATT/WTO system. By the same token, the logic for excluding international labor and environmental law also emerges. As restraints on international trade, affirmative labor and environmental standards would be anathema to the stated objectives of the system.

Research shows that the interests of trade expansion need not run contrary to labor and environmental protection, however. Recently, the WTO Committee on Trade and Environment commissioned a study on whether environmental regulations impose costs on production that tend to increase prices and thereby decrease trade volume. The resulting report found that, although environmental protection increases production costs, its "competitiveness effect" is ultimately minor.[89] The report found that this minimal effect on competitiveness results, in part, because environmental protection costs are likely offset by cost-saving efficiency gains elsewhere in the production process.[90] In addition, the report found that, with some exceptions, environmental regulations are not of primary importance in international investment decisions and that most industries do not move from developed to developing countries to reduce their environmental compliance costs.[91]

With respect to labor regulation, some commentators argue that many labor protections will have little or no impact on the cost of production. A state's commitment to protecting the freedom of association, for example, imposes no immediate cost on producers.[92]

Even if labor costs increase, however, increased labor costs need not increase overall production costs. On the contrary, improving labor conditions can increase labor productivity such that output per cost unit of labor actually goes up. This focus on increased productivity is a touchstone of labor protection proponents, including the ILO.[93]

The second response to the argument that labor and environmental standards will restrict trade is that, whatever the data, the potential for conflict does not justify segregation of the regimes. Segregation is only logical if "sticking one's head in the sand" is also logical. If potential or actual conflicts exist, the need for examination and coordination increases, rather than decreases. The head of the UNEP, Klaus Töpfer,

expressed this sentiment in his 1998 address to the WTO, arguing for a sustained effort to redress potential conflicts between trade and environmental law so as to create a "win–win agenda" for economic growth that is environmentally sound and socially just.[94]

4.2.2 Political Will (and the North–South Divide)

Conflicts between trade and labor/environment would seem to argue in favor of, rather than against, a concerted effort to integrate the two bodies. Given that the previously discussed distinctions do not hold, the question arises as to why incorporation of affirmative labor and environmental standards into the WTO should not proceed in earnest. Extenuating political circumstances help to explain states' persistent reluctance to address head-on the relationship between international trade, labor, and environmental regimes in WTO affairs. A noted environmentalist put it simply: "Unfortunately, ... there is no political will to take decisions that will give us the tools to do the job."[95]

Developing countries have more clearly voiced resistance than developed countries.[96] There are immediate reasons for this resistance – if enforcement of labor and environmental standards increases due to incorporation into the WTO regime, compliance costs would accrue disproportionately to poor countries, yielding a net competitive marketplace disadvantage to those countries *vis-à-vis* rich countries.

While economic consequences are at the nub of developing countries' opposition, they alone do not explain it. The rancour often present in such objections derives from a widely held belief that the attempt to include labor and environmental standards is little more than disguised protectionism.[97] Rich countries and some of their interest group constituencies are concerned with the market share that they might lose to production from poor countries. Imposing increased costs via labor and environmental standards will decrease or eliminate the competitive advantage production in poor countries.

Unfortunately, this perception of protectionism is not wholly unfounded. For example, industrialized countries have continued to demonstrate great reluctance in reducing trade barriers imposed to protect domestic industries that otherwise would likely lose out to competition from poor countries if unprotected. Textiles, a sector whose production is labor-intensive, and in which countries featuring an abundance of cheap labor may enjoy a competitive advantage, is one example. Agriculture, whose dependence on natural resources also provides poor countries with a potential competitive advantage, is another.

Yet, in both textiles and agriculture, industrialized country protectionism resulted in ongoing contravention of GATT discipline.[98] Agreements to bring textiles and agriculture under trade-liberalizing rules in the Uruguay Round have proven among the least successful in the WTO.[99] The reason is that removing restrictions in those sectors runs contrary to the political interests of industrialized countries in protecting certain domestic industries. These lacunae of protectionism squarely contradict GATT/WTO principles. They also lend credence to the belief that the focus on labor and environmental protection is little more than an attempt by the North to manipulate the international trade regime to its benefit. That the United States, where many of the most influential

linkage proponents reside, has failed to ratify key international labor and environmental instruments only underscores this perception.[100] That issues of developing-country concern, such as migrant labor, have failed to make the "core" list of concerns of international organizations, further underscores it.[97]

In reviewing these political dynamics, one might notice that developing country interests, at least as commonly construed in this particular debate,[101] run counter to IP protection as well as to labor and environmental concerns. At the time of the Uruguay Round negotiations, many developing countries staunchly opposed the incorporation of IP rights. The argument sounded quite similar to those made in the context of labor and environmental debates. Developing countries argued that the administration of IP rights properly belonged to the United Nations organization traditionally charged with that duty: the World Intellectual Property Organization.[102]

Why does developing country opposition seem fatal to the incorporation of labor and environmental standards, when it did not prove fatal to the incorporation of IP standards? First, general opposition to IP standards in developing countries resulted not in an absolute opposition to their incorporation, but rather in a *quid pro quo* approach.[103] Having reconciled themselves to the need to establish a comprehensive multilateral trade regime, many developing countries sought to leverage their willingness to countenance IP negotiations for the purpose of including negotiations on sectors of interest to them. Textiles and agriculture were two such notable sectors.[104] Negotiations on both of these topics were opposed by trade groups in the United States that had an interest in maintaining the status quo of high protectionist barriers to imports.[105]

Were labor or environment negotiations to begin, the *quid pro quo* approach would undoubtedly reemerge. That is, if industrialized countries wanted better labor and environmental practices in poor countries, they would likely have to support the process of acquiring those practices. This approach has been a staple of recent international environmental law, in which the principle of environmental protection is now wedded to the goal of development in poor countries.[106] Many environmental agreements include provisions for technology transfer and funding mechanisms to aid developing countries in compliance efforts.[107] Developing country resistance to linkage is likely to continue as long as proponents of trade–labor/environment linkage are not successful at constructing cost-sharing arrangements, as in the case of environmental regulations, or present only partial arrangements, as in the case of labor regulations.[108]

5 Conclusion

Environmental and labor standards, like IP standards, can affect trade flows. Despite these similarities, only IP standards have been incorporated into the WTO. Many of the conventional reasons for excluding labor and environment standards can be cleared away, leaving more genuine issues of political will and orientation. The resolution of these political questions would still leave significant challenges for drafting enforceable legal standards.[109] Such difficulties could be overcome, however, were it deemed desirable to do so. Resolving the relationship of "nontrade" areas to WTO law remains one of the challenges for international law and politics.

Acknowledgments

I would like to thank Christopher McCrudden, Robert Howse, Katherine Franke, Sheila Foster, Sonia Katyal, Terry Smith, Jim Fleming, and Matthew Diller for their helpful perspectives and comments.

Notes

1 'The WTO is only competent to deal with trade.' World Trade Org., Trading into the Future 46 (1999).

2 *Id.* The WTO has stated that its competence in environmental matters is limited to assessing the impact of environmental matters on trade flows, and the relationship between WTO obligations and "specific trade obligations" in Multilateral Environmental Agreements. Broader environmental obligations are outside this scope. See, e.g. WTO Doha Ministerial Declaration, 20 Nov. 2001, WT/MIN(01)/Dec/1; WTO Environmental Backgrounds (2004).

3 *Id.* at 51.

4 See Chantal Thomas, *Constitutional Change and International Government*, 52 Hastings L. J. 1, 5 (2000) (discussing the failure of the WTO to consider social goals such as environmental protection, employment security, and consumer safety). This chapter focuses on the cases for incorporation of "affirmative" obligations. Current WTO law (e.g., GATT Article XX) includes "exception" provisions that allow member states to depart from general rules to implement certain kinds of social policies.

5 Subsequent to this identification, the WTO committee preparing an agenda for the first WTO ministerial conference was charged with considering labor and environmental issues. The result was a single paragraph incorporated in the larger Ministerial Declaration of the WTO:

> We renew our commitment to the observance of internationally recognized core labour standards. The International Labour Organization (ILO) is the competent body to set and deal with these standards, and we affirm our support for its work in promoting them. We believe that economic growth and development fostered by increased trade and further trade liberalization contribute to the promotion of these standards. We reject the use of labour standards for protectionist purposes, and agree that the comparative advantage of countries, particularly low-wage developing countries, must in no way be put into question. In this regard, we note that the WTO and ILO Secretariats will continue their existing collaboration.

WTO, Singapore Ministerial Declaration, adopted December 13, 1996, WT/MIN(96)/DEC ¶ 4 (December 18, 1996), available at http://www.wto.org, reprinted in 36 I.L.M. 218, 221 (1997). This Declaration can be contrasted with the more concrete institutional linkages the WTO has built with the International Monetary Fund and the World Bank. See Press Release, International Monetary Fund, WTO, and IMF Sign Cooperation Agreement (December 9, 1996); Press Release, WTO, and World Bank Sign Cooperation Agreement (April 28, 1997) 1, at 43–4 (1999). During the Uruguay Round of WTO negotiations, a formal working group on trade and the environment was established – the Committee on Trade and the Environment. Trade and Environment Decision of April 14, 1994, WT/CTE/1, Annex I (November 12, 1996), available at http://www.wto.org, reprinted in 33 I.L.M. 1125, 1267 (1994). The Committee was charged with determining the relationship between international trade law and international environmental

law. *Id.* The 1994 decision establishing the Committee laid down a mandate including directives "(a) to identify the relationship between trade measures and environmental measures . . . [and] (b) to make appropriate recommendations on whether any modifications of the provisions of the multilateral trading system are required." *Id.* (quotations omitted). The Committee has, among other things, kept track of the extent to which international environmental agreements provide for the use of trade sanctions, as well as the general enforcement mechanisms of those agreements. See, e.g., The Convention on Biological Diversity and its Relation to Trade, Communication from the Executive Secretary of the Convention on Biological Diversity, WT/CTE/W/64 (September 29, 1997).

6 See John Burgess and Steven Pearlstein, *Protests Delay WTO Opening – Seattle Police use Tear Gas; Mayor Declares a Curfew*, Wash. Post, December 1, 1999, at A1. President Clinton, partially in response to the concerns of labor and environmental groups and partially in keeping with appeals to put a "human face on the global economy," proposed that the WTO formally revisit the question of trade and labor linkages by establishing a formal working group. See *Clinton's Plea: "Open the Meetings,"* NY Times, December 2, 1999, at A17. The response was met with opposition by many governments of developing countries. See Clyde Summers, *The Battle in Seattle: Free Trade, Labor Rights, and Societal Values*, 22 U. Pa. J. Int'l Econ. L. 61, 64–5 (2001) (describing the historic divide between developed and developing countries on whether or not trade should be free from restraint); Steven Greenhouse and Joseph Kahn, *US Effort to Add Labor Standards to Agenda Fails*, NY Times, December 3, 1999, at A1 (quoting the Pakistani representative as warning, "We will block consensus on every issue if the United States proposal goes ahead.").

7 Since Seattle, the WTO has firmly any discussion of labor issues. *Trading into the Future*, the recent WTO publication, went so far as to assert, in the context of a brief discussion of the relationship between labor and trade: "Strictly speaking, this should not be mentioned here at all because there is no work on the subject in the WTO, and it would be wrong to assume that it is a subject that 'lies ahead."' World Trade Org., *Trading into the Future* 51 (1999). Although the WTO's stance toward environmental concerns has been less unequivocal than toward labor, the bottom line seems to come out in the same place. In 1999, the WTO convened its "High-Level Symposium on Trade and the Environment." Representatives of WTO member states, international and non-governmental environmental organizations, and scholars debated various approaches. No clear mandate resulted from this symposium, although the WTO Director-General did suggest the formation of an environmental organization that would operate in parallel fashion to the WTO – a "World Environment Organization." Renato Ruggiero, Director-General, WTO, *Opening Remarks to the High Level Symposium on Trade and the Environment* (March 15, 1999), available at http://www.wto.org/english/tratop_e/envir_e/dgenv.htm (last visited June 30, 2003). The Director-General's remarks reinforce the institutional competence reasoning for excluding environmental concerns while at the same time acknowledging that international environmental law is insufficiently institutionalized.

8 For a shorter discussion of this analogy, see Chantal Thomas, *Trade-Related Labor and Environment Agreements?* 5 Journal of International Economic Law 791 (2002).

9 See Chantal Thomas, *Trade-Related Labor and Environment Agreements?* 5 Journal of International Economic Law 791, 799–816 (2002).

10 See *Symposium: The Boundaries of the WTO*, Jose E. Alvarez (ed.), 96 Am. J. Int'l L. 1–158 (2002); *Symposium, Global Trade Issues in the New Millennium*, 33 Geo. Wash. Int'l L. Rev. (2001); *Symposium, Linkage as Phenomenon: An Interdisciplinary Approach*,

19 U. Pa. J. Int'l Econ. L. 201 (1998); Andrew Guzman, *Global Governance and the WTO* (manuscript on file with author); see also *Fair Trade and Harmonization: Prerequisites for Free Trade*? (Jagdish Bhagwati and Robert Hudec eds., 1996).

11 Agreement on Trade-Related Aspects of Intellectual Property Rights, April 15, 1994, WTO Agreement, Annex 1C [Treaties Binder 1, Treaties Booklet 1] L. & Prac. World Trade Org. (Oceana) 383 (March 1995), 33 I.L.M. 81 (1994) [hereinafter TRIPS Agreement].

12 A play on the acronym made by the distinguished trade scholar Jagdish Bhagwati – humorously asserting that "TRIPS" could also be understood to mean the "Tangentially Related Intellectual Property" issues – also makes the point. See Robert P. Merges, *Battle of Lateralisms: Intellectual Property and Trade*, 8 B.U. Int'l L.J. 239, 240 n. 2 (1990) (quoting Jagdish Bhagwati's play on words).

13 Joel P. Trachtman, *International Regulatory Competition, Externalization, and Jurisdiction*, 34 Harv. Int'l L.J. 47, 49 (1993).

14 The article does not directly address the question of incorporating human rights into the WTO, but many labor and environment rights may be deemed to be human rights. For a discussion of this issue, see Chantal Thomas, *Poverty Reduction, Trade and Rights*, 18 Am. U. Int'l L. Rev. 1399 (2003).

15 World Trade Org., *Trading into the Future* 25 (1999) ("New internationally agreed trade rules for IP rights were seen as a way to introduce more order and predictability, and for disputes to be settled more systematically.")

16 Convention establishing the World Intellectual Property Organization, July 14, 1967, 21 U.S.T. 1749, 828 U.N.T.S. 3 [hereinafter WIPO].

17 See Brent W. Sadler, *Intellectual Property Protection Through International Trade*, 14 Hous. J. Int'l L. 393, 400–1 (1992) ["WIPO is of little practical use in protecting intellectual property rights . . . (because) it lacks meaningful enforcement provisions such as are found in the GATT."]; see also Robert P. Merges, *Battle of Lateralisms: Intellectual Property and Trade*, 8 B.U. Int'l L.J. 239, 240 (1990) (describing an "increasingly lethargic and bureaucratized WIPO").

18 See Sadler, *supra* note 17, at 393–4 ("Many countries do not protect intellectual property rights to the extent desired by the United States. Intellectual property laws differ from nation to nation, both in scope of protected rights and enforcement. Existing multinational agreements do not effectively protect intellectual property rights of United States citizens and industries in the global market.")

19 World Trade Org., *supra* note 1, at 25 ["The extent of protection and enforcement of (IP) rights varied widely around the world; and as intellectual property became more important in trade, these differences became a source of tension in international economic relations."].

20 Robert W. Kastenmeier and David Beier, *International Trade and Intellectual Property: Promise, Risks, and Reality*, 22 Vand. J. Transnat'l L. 285, 296 (1989) ("The key to a strong international intellectual property system within the GATT is . . . a strong dispute resolution mechanism.").

21 *Cf.* Robert P. Merges, *Battle of Lateralisms: Intellectual Property and Trade*, 8 B.U. Int'l L.J. 239, 240 n. 2 (1990) (noting, by contrast, that "[t]here is little potential for horsetrading in the WIPO context").

22 Constitution of the ILO, Treaty of Versailles, June 28, 1919, Part XIII, 49 Stat. 2713, 225 Consol. T.S. 189, 373 [hereinafter ILO Constitution].

23 The OECD is an intergovernmental organization whose 30 current members are the high-income countries together with a few medium-income countries such as Mexico and Korea.

24 See Steve Charnovitz, *Trade, Employment and Labour Standards: The OECD Study and Recent Developments on the Trade and Labor Standards Debate*, 11 Temp. Int'l & Comp. L.J. 131, 137 (1997) (noting that in the surveyed set of countries, 22 of 24 OECD countries ensured freedom of association, while nine of 67 non-OECD countries surveyed did, and that 20 of 24 OECD countries, the United States not among them, met the standard of freedom of collective bargaining, while only 15 of 67 non-OECD countries did); see generally, Organization for Econ. Co-operation & Dev., *Trade, Employment, and Labour Standards: A Study of Core Workers' Rights and International Trade* (1996) (describing results of the 1996 study of compliance with core labor principles).

25 The ratifications of "core" ILO agreements are as follows: 1948 Freedom of Association and Protection of the Right to Organise Convention has been ratified by 141 countries including Argentina, Austria, and Uruguay, but not Zimbabwe. The 1957 Abolition of Forced Labour Convention has been ratified by 159 countries, including Argentina, Austria, Uruguay, and Zimbabwe. The 1973 Minimum Age Convention has been ratified by 121 countries including Argentina, Austria, Uruguay, and Zimbabwe. The 1999 Worst Forms of Child Labour Convention has been ratified by 134 countries, including Argentina, Austria, Uruguay, and Zimbabwe. The record for noncore agreements is much different: The 1981 Occupational Safety and Health Convention has been ratified by 39 countries, including Uruguay, but not including Austria, Argentina, and Zimbabwe. The 1991 Working Conditions (Hotels and Restaurants) Convention has been ratified by 12 countries including Austria and Uruguay, but not including Argentina and Zimbabwe. The 1946 Food and Catering (Ships' Crews) Convention has been ratified by 24 countries, including Argentina, but not including Austria, Uruguay, and Zimbabwe. The 1969 Labour Inspection (Agriculture) Convention has been ratified by 40 countries, including Argentina, Uruguay, and Zimbabwe, but not Austria. The current lists of ratifications of these conventions is found at http://www.ilo.org/ilolex/english/convdisp1.htm.

26 Declaration of the United Nations Conference on the Human Environment, adopted June 16, 1972, UN Doc. A/CONF.48/14, reprinted in 11 I.L.M. 1416 (1972) [hereinafter Stockholm Declaration].

27 See G.A. Res. 2997, UN GAOR, 27th Sess., Supp. No. 30, at 43, UN Doc. A/8730 (1972), reprinted in 12 I.L.M. 433 (1973) (establishing UNEP).

28 Geoffrey Palmer, *New Ways to Make International Environmental Law*, 86 Am. J. Int'l L. 259, 260, 261 (1992). Other agencies have also worked on parallel tracks. For example, "although the international community's emphasis during the 1970s was on pollution of water, air, and soil, the International Union for Conservation of Nature and National Resources expanded the community's concern to endangered wildlife and played a big part in negotiating the 1973 Washington Convention on International Trade in Endangered Species of Wild Fauna and Flora." Andronico O. Adede, *The Treaty System from Stockholm (1972) to Rio De Janeiro*, 13 Pace Envtl. L. Rev. 33, 36 n.17 (1992). See generally Convention on International Trade in Endangered Species of Wild Fauna and Flora, March 3, 1973, 27 U.S.T. 1087, 993 U.N.T.S. 243.

29 Notwithstanding this lack of formal authority, the UNEP has succeeded in encouraging states to negotiate new environmental treaties elaborating on such core principles as marine conservation. Among other things, the UNEP negotiated the 1976 Barcelona Convention for the Protection of the Mediterranean Sea Against Pollution, the 1978 Kuwait Regional Convention for Co-operation on the Protection of the Marine Environment from Pollution, the 1981 Abidjan Convention for Co-operation in the Protection and Development of the Marine and Coastal Environment of the West and Central African Region, the 1981 Lima Convention for the Protection of the Marine Environment and Coastal Areas

of the South-East Pacific, the 1982 Jeddah Regional Convention for the Conservation of the Red Sea and Gulf of Aden Environment, the 1983 Cartegena Convention for the Protection and Development of the Marine Environment of the Wider Caribbean Region, the 1985 Nairobi Convention for the Protection, Management and Development of the Marine and Coastal Environment of the Eastern African Region, and the 1986 Noumea Convention for the Protection of the Natural Resources and Environment of the South Pacific Region. *Id.* The UNEP's efforts have also proven key to the establishment of treaties on ozone depletion. See Palmer, *supra* note 28, at 261 ("In my opinion, without UNEP, the system to prevent ozone depletion now in place would not have been developed.").

30 In a document entitled the Hague Declaration on the Environment, several countries called for the "new and more effective decision-making and enforcement mechanisms." The Hague Declaration on the Environment, done March 11, 1989, 28 I.L.M. 1308, 1309. In 1988, the President of the World Federation of United Nations Associations proposed that a new mandate be issued to the United Nations Trusteeship Council to protect the "planetary systems on which our security and survival depends, as well as over the global commons." Durwood Zaelke and James Cameron, *Global Warming and Climate Change – An Overview of the International Legal Process*, 5 Am. U.J. Int'l L. & Pol'y 249, 280 (1990). The World Commission on Environment and Development, an expert group established by the United Nations, has called for the establishment of a centralized mechanism headed by a United Nations High Commissioner for the environment. Such a centralized mechanism would be responsible for investigating and making reports on state compliance with core principles of international environmental law. See The World Commission on Environment and Development, Our Common Future 308–51 (1987) (suggesting establishment of a mechanism for investigating and reporting on compliance). More recently, commentators have called for the establishment of an International Court for the Environment. Amedeo Postiglione, *A More Efficient International Law on the Environment and Setting up an International Court for the Environment Within the United Nations*, 20 Envtl. L. 321, 323 (1990). A number of other international environmental scholars have also weighed in with their own proposals for a central institution for the monitoring and enforcement of international environmental law. See Palmer, *supra* note 28, at 278–82 (proposing an International Environment Organization).

31 See generally Greening International Law (Philippe Sands ed., 1993) (discussing UNCED and the development of international environmental law). A few years prior to the UNCED's review, the UN General Assembly had set forth a framework, the System-Wide Medium-Term Environment Programme, for the integration and implementation of international environmental law. G.A. Res. 32/197, UN GAOR, 32nd Sess., Supp. No. 45, at 121, UN Doc. A/32/45 (1977); United Nations, UNEP, System-Wide Medium-Term Environment Programme for the Period 1990–5, UN Doc. UNEP/GC/SS.I/7/Add.1 (1988). Agenda 21 of UNCED, for example, outlined an extensive set of recommendations for the institutionalization of international environmental law. See Edith Brown Weiss, *Introductory Note*, United Nations Conference on Environment and Development, 31 I.L.M. 814, 814–15 (1992) (describing Agenda 21).

32 Convention on Biological Diversity, done June 5, 1992, 1760 U.N.T.S. 79, 31 I.L.M. 818 [hereinafter Convention on Biological Diversity]. The Convention was negotiated out of a separate initiative by the UNEP. See Proceedings of the Governing Council at its Fifteenth Session, UNEP Governing Council, 15th Sess., Agenda Item 12 at 165, UN Doc. UNEP/GC.15/12 (1989) (authorizing *ad hoc* working group to negotiate international agreement on biological diversity).

33 Marc Pallemaerts, *International Environmental Law in the Age of Sustainable Development: A Critical Assessment of the UNCED Process*, 15 J.L. & Com. 623, 660 (1996) (quotation omitted). Pallemaerts observes:

> The Convention is no doubt legally binding, but is it any more "authoritative" than [nonbinding statements], in the sense of imposing on the contracting parties precise rules of conduct . . .? [ ] Indeed, all of the substantive obligations of the parties relating to the identification and monitoring of the elements of biological diversity (Article 7); to *in situ* conservation (Article 8) and to *ex situ* conservation (Article 9); to sustainable use (Article 10); to incentives (Article 11), and to environmental impact assessment (Article 14) are qualified by the phrase "as far as possible and as appropriate."

Id. These would be called "illusory promises" in ordinary contract law.

34 The only regular institutional mechanism is that of regular meetings of the state parties. Convention on Biological Diversity, art. 23, 1760 U.N.T.S. at 157, 31 I.L.M. at 832. Implementation reports are called for, but there is no regular reporting requirement. Rather, parties are to present reports "at intervals to be determined by the Conference of the Parties." *Id.* art. 26, 1760 U.N.T.S. at 159, 31 I.L.M. at 834. This contrasts with the regular reporting requirements laid down as part of the WTO. Trade Policy Review Mechanism, April 15, 1994, WTO Agreement, Annex 3 (Treaties Binder 1, Treaties Booklet 3) L. & Prac. World Trade Org. (Oceana) 1 (March 1995). For a brief description of the Trade Policy Review Mechanism, see Amelia Porges, *Introductory Note*, GATT: Multilateral Trade Negotiations Final Act Embodying the Results of the Uruguay Round of Trade Negotiations, 33 I.L.M. 1125, 1129 (1994). A similar laxity characterizes the Convention's dispute settlement provisions. The Convention allows, but does not require, parties to submit a dispute to compulsory arbitration or compulsory submission to the International Court of Justice. Convention on Biological Diversity, art. 27, ¶ 3, 1760 U.N.T.S. at 159, 31 I.L.M. at 834. If the parties have not both so agreed, and cannot otherwise agree on a dispute settlement process, the dispute is to be addressed through the Convention's "conciliation" process. *Id.* art. 27, ¶ 4, 1760 U.N.T.S. at 160, 31 I.L.M. at 834. This process results in a "proposal for resolution of the dispute, which the parties shall consider in good faith." *Id.* Annex II, part 2, art. 5, 1760 U.N.T.S. at 169, 31 I.L.M. at 841.

35 Communication from the Secretariat of the Convention on Biological Diversity to the Committee on Trade and the Environment, WT/CTE/W/116, ¶ 3 (June 28, 1999), available at http://www.wto.org.

36 Protocol on Substances that Deplete the Ozone Layer, opened for signature September 16, 1987, 26 I.L.M. 1541 [hereinafter Montreal Protocol]; see generally Elizabeth R. DeSombre, *The Experience of the Montreal Protocol: Particularly Remarkable, and Remarkably Particular*, 19 UCLA J. Envtl. L. & Pol'y 49 (2000/2002) (discussing Montreal Protocol).

37 See *id.* at 70 (describing the Multilateral Fund).

38 The Montreal Protocol explicitly addressed the issue of enforcement by calling on the parties to consider and approve enforcement mechanisms in the future. See Montreal Protocol, art. 8, 26 I.L.M. at 1556 ("The Parties, at their first meeting, shall consider and approve procedures and institutional mechanisms for determining noncompliance with the provisions of this Protocol and for treatment of Parties found to be in noncompliance.").

39 See Communication from the Secretariat for the Vienna Convention and the Montreal Protocol to the WTO Committee on Trade and the Environment, UNEP, WT/CTE/W/115, ¶ 17 (June 25, 1999), available at http://www.wto.org (stating that potential actions include "assistance, caution or suspension of specified rights and privileges under the Protocol"

and that "[t]hat the measures for noncompliance include assistance, on the assumption that a Party's noncompliance is not deliberate but only due to its inability, is a novel provision introduced by the Parties to the Montreal Protocol").

40 See *id.* ¶ 18 ("So far, only the case of the Russian Federation and other Republics of the former Soviet Union are before the Committee for noncompliance, on the basis of the Secretariat's report.").

41 One patently frustrated NGO bemoaned an "unfamiliar network of haphazardly coordinated [agencies] . . . a fantasm, with mirage-like powers, a creaking and fragmented process for deciding policy, and a surfeit of bureaucratic fiefdoms that consistently muster inadequate resources to meet even the most urgent challenges." United Nations Ass'n of the USA, Uniting Nations for the Earth 33 (1990). Another commentator remarked, "The problem is that rhetoric about the environment is far easier to produce than action, and international forums tend on occasion to degenerate into 'rhetoric-fests,' where world leaders spout all the proper phrases but then go home and often fail to implement their internationally-formulated promises." Ranee K. L. Panjabi, *From Stockholm to Rio: A Comparison of the Declaratory Principles of International Environmental Law*, 21 Denv. J. Int'l L. & Pol'y 215, 216 (1993).

42 See generally Paulsen K. Vandevert, *The Uruguay Round and the WTO: A New Era Dawns in the Private Law of International Customs and Trade*, 31 Case W. Res. J. Int'l L. 107 (1999) (discussing private and public international law). More sophisticated commentators hasten to reject this distinction. See, e.g., John H. Jackson, *International Economic Law: Reflections on the "Boilerroom" of International Relations*, 10 Am. U. J. Int'l L. & Pol'y 595, 597 (1995) ("International Economic Law (IEL) can not be separated or compartmentalized from general or 'public' international law. The activities and cases relating to IEL contain much practice which is relevant to general principles of international law, especially concerning treaty law and practice.")

43 The public–private distinction in and of itself has come under extensive scrutiny. See, e.g., Robert L. Hale, *Coercion and Distribution in a Supposedly Non-Coercive State*, 38 Pol. Sci. Q. 470, 470–94 (1923) (arguing that the distinction is invalid); Duncan Kennedy, *From the Will Theory to the Principle of Private Autonomy: Lon Fuller's "Consideration and Form,"* 100 Colum. L. Rev. 94, 107 (2000) ("Classical legal thought should be associated . . . with a particular ordering of substantive principles around the public–private distinction, iterated and reiterated at every level of doctrine."); see also Thomas C. Grey, *Langdell's Orthodoxy*, 45 U. Pitt. L. Rev. 1, 48–9 (1983) (describing how the classical legal structure imposes a value-laden and questionable distinction between those types of lawmaking it describes as "private" and those it describes as "public"). For example, contract law is normally viewed as private law, whereas securities law is viewed as administrative law. Yet, both ultimately govern individual economic activity, and both are formulated and enforced by public authorities including courts, legislatures, and agencies. A difference might be that securities law primarily sets a framework for agency regulation, whereas contract law primarily sets a framework for judicial regulation. This difference clearly does not follow a literal public–private distinction, although it is the difference that most reliably predicts how a particular body of economic law will be characterized. Although critiques of the public–private distinction are quite warranted in my view, it is not the point I want to make here. Rather, I am arguing here that even accepting this distinction on its own terms, it has not been accurately applied to international trade law.

44 Hence the promulgation by the International Institute for the Unification of Private Law of its guideline, Principles of International Commercial Contracts (the UNIDROIT Principles). International Institute for the Unification of Private Law, Principles of International Commercial Contracts (1994). For a discussion of the UNIDROIT Principles,

see Joseph M. Perillo, *Force Majeure and Hardship under the UNIDROIT Principles of International Commercial Contracts*, 5 Tul. J. Int'l & Comp. L. 5, 14–26 (1997); see also Luiz Olavo Baptista, *The UNIDROIT Principles for International Commercial Law Project: Aspects of International Private Law*, 69 Tul. L. Rev. 1209, 1209–24 (1995) (explaining the UNIDROIT Principles). For a more general discussion, see Harold H. Koh, *The Globalization of Freedom*, 26 Yale J. Intl L. 305 (2001).

45 Convention on Contracts for the International Sale of Goods, opened for signature April 11, 1980, 1489 U.N.T.S. 3, 19 I.L.M. 668.

46 See, e.g., John H. Jackson, *Book Review*, 82 Am. J. Int'l L. 653, 653–4 (1988) ("[T]he GATT has had to exist without the many institutional clauses found in normal organization charters. These deficiencies are noted by the author, who describes many of the fascinating circumstances of practice and evolution that have allowed the GATT to exist as a pragmatic and reasonably effective legal order.")

47 See Robert E. Hudec, *Developing Countries in the GATT Legal System* 30 (1987) (describing developing-country compliance as a situation of "pro forma"). It should be noted, however, that the principle of special and differential treatment enjoys a solid basis in international law, both in form and substance. See Chantal Thomas, *Balance-of-Payments Crises in the Developing World: Balancing Trade, Finance and Development in the New Economic Order*, 15 Am. U. Int'l L. Rev. 1249, 1259–60 (2000) (describing the emergence of the principle of special and differential treatment and its operation in the GATT).

48 The shortfalls in compliance were numerous. Industrialized countries, for their part, engaged in substantial departure from GATT disciplines on textiles and agriculture. See Olivier Long, *Law and its Limitations in the GATT Multilateral Trade System* 56 (1985) (highlighting the practical enforcement techniques of the GATT). If developing countries had a less glaring problem with formal compliance, it was because they argued that the principle of "special and differential treatment" should except them from many GATT rules.

49 See generally David Kennedy, *The International Style in Postwar Law & Policy*, 1994 Utah L. Rev. 7 (discussing postwar attitudes toward international law).

50 See Long, *supra* note 48, at 94–8 (highlighting the practical enforcement techniques of the GATT).

51 See G. Richard Shell, *Trade Legalism and International Relations Theory: An Analysis of the WTO*, 1995 Duke L.J. 829, 833 ("[T]he new WTO system represents a stunning victory for international trade 'legalists' in their running debate with trade 'pragmatists' . . .."); Chantal Thomas, *Balance-of-Payments Crises in the Developing World: Balancing Trade, Finance and Development in the New Economic Order*, 15 Am. U. Int'l L. Rev. 1249, 1250 (2000) ("[I]nstitutionally, the order has moved away from 'pragmatism' and toward 'legalism.'"); see generally Kenneth W. Abbott, *The Many Faces of International Legalization*, 92 Am. Soc'y Int'l L. Proc. 57 (1998) (discussing causes and effects of legalization in international politics); Miquel Montañà i Mora, *A GATT with Teeth: Law Wins over Politics in the Resolution of International Trade Disputes*, 31 Colum. J. Transnat'l L. 103 (1993) (discussing resolution of trade disputes).

52 Sylvia Ostry, Reinforcing the WTO, Group of Thirty, Occasional Paper No. 56 (1998), http://www.utoronto.ca/cis/wtogp30.pdf (last visited July 7, 2003) (describing the dispute settlement mechanism as "the 'jewel in the crown' of the WTO").

53 The WTO dispute settlement mechanism features the following novel reinforcements: (1) a dispute resolution timeline of approximately 15 months (whereas disputes in the GATT were not subject to any timeline, and the longest one ran for over a decade); (2) an appellate review body (no such body existed in the GATT); (3) the virtually automatic adoption of findings (findings can be rejected only by unanimous decision – by contrast,

GATT findings were infrequently adopted because unanimous consent was required for their adoption rather than their rejection); (4) an increase in the ease and likelihood of concrete sanctions for noncompliance through the authorization of retaliation by the adversely affected party (such retaliation had occurred only once in the decades-long operation of the GATT dispute resolution mechanism). See John H. Jackson et al., *Legal Problems of International Economic Relations*, ch. 7 (3rd edn., 1995), for a general discussion of the GATT/WTO dispute settlement system.

54 As of November 2002, the WTO dispute settlement mechanism had received 273 requests for consultation. See WTO, Dispute Settlement Body, Annual Report (2002), Addendum, Overview of the State-of-Play of WTO Disputes, WT/DSB/29/Add.1, sec. I (November 29, 2002) & WT/DSB/29/ Add.1/Corr.1 (December 10, 2002) (listing requests for consultation), available at http://www.wto.org. By contrast, barely over 100 panel reports were adopted under the GATT in its 40 years of operation. GATT Panel Reports are available from the official GATT reporter, Basic Instruments and Selected Documents (BISD). They can be viewed on the WTO website at http://www.wto.org/english/tratop_e/dispu_e/gt47ds_e.htm.

55 For a general discussion of this issue, see Gregory Shaffer, *WTO Blue–Green Blues: The Impact of US Domestic Politics on Trade–Labor, Trade–Environment Linkages for the WTO's Future*, 24 Fordham Int'l L.J. 608 (2000).

56 See, e.g., Chantell Taylor, *NAFTA, GATT, and the Current Free Trade System: A Dangerous Double Standard for Workers' Rights*, 28 Denv. J. Int'l L. & Pol'y 401, 401 (2000) (voicing the concern that a "transfer of power" to "trade bureaucracies" is "producing dire consequences for the environment, human rights, social welfare, agriculture, food safety, workers' rights, national sovereignty, and democracy itself").

57 See Gunther F. Handl et al., *A Hard Look at Soft Law*, 82 Am. Soc'y Int'l L. Proc. 371, 371–3 (1988) (explaining concept of "soft law").

58 See Paul H. Brietzke, *Insurgents in the 'New' International Law*, 13 Wis. Int'l L.J. 1,7 (1994) (describing a process in which those who seek to reform or expand public international law must often take "the 'soft law' route, in the hope that a binding customary international law . . . will emerge eventually"); see generally R. R. Baxter, *International Law in "Her Infinite Variety,"* 29 Int'l & Comp. L.Q. 549 (1980) (discussing variation of norms in customary international law); Michael Bothe, *Legal and Non-Legal Norms – A Meaningful Distinction in International Relations?*, 11 Neth. Y.B. Int'l L. 65 (1980) (discussing distinction between legal and nonlegal norms in international relations); Oscar Schachter, *The Twilight Existence of Nonbinding International Agreements*, 71 Am. J. Int'l L. 296 (1977) (discussing obligations under nonbinding agreements).

59 See Stephanie Farrior, *The International Law on Trafficking in Women and Children for Prostitution: Making it Live up to its Potential*, 10 Harv. Hum. Rts. J. 213, 252 (1997) ("Concerted action and close cooperation among a range of governmental and nongovernmental bodies are integral to curbing trafficking for prostitution."); Tamlyn Hunt, *People or Power: A Comparison of Realist and Social Constructivist Approaches to Climate Change Remediation Negotiations*, 6 UCLA J. Int'l L. & Foreign Aff. 265, 277 (2001) (describing a "soft social constructivism" that "looks to the role of nonstate actors, in the form of transnational advocacy networks (TANs), epistemic communities (ECs or transnational epistemic communities, TECs), or NGOs in defining international law"). Examples of such efforts abound in both international labor law and international environmental law. See, e.g., Ruth Greenspan Ball, *Developing a Culture of Compliance in the International Environmental Regime*, 27 Envtl. L. Rep. 10402, 10406 (1997) (arguing that the publicity campaigns of NGOs, in part resulting from lawsuits, play an important role in increasing compliance with international environmental law); Katherine Van Wezel Stone, *To the*

Yukon and Beyond: Local Laborers in a Global Labor Market, 3 J. Small & Emerging Bus. L. 93, 121 (1999) ("Soft regulation refers to a regulatory regime that is comprised of an interwoven network of private sector actors and voluntary associations acting within the context of market forces to establish and enforce rules of conduct. In the area of labor rights, one would look to labor activists, trade unionists, enlightened firm managers, academics, and progressive political leaders to fashion such a regime.").

60 The fact-finding and reporting capabilities of the United Nations Human Rights Commission, the International Labor Organization, and the UNEP all exemplify this approach.

61 There was much rumination about these aspirations during the fiftieth anniversary year of the United Nations. See, e.g., Saul H. Mendlovitz and Burns H. Weston, *The United Nations at Fifty: Toward Humane Global Governance*, 4 Transnat'l L. & Contemp. Probs. 309, 311 (1994) (noting trend from "'geopolitics' to 'geogovernance'" in the world order); Carolyn L. Willson, *The United Nations at Age Fifty: A Legal Perspective*, 91 Am. J. Int'l L. 762, 762–4 (1997) (book review) (commenting on attempts to assess effectiveness of United Nations); see generally, *The United Nations at Age Fifty: A Legal Perspective* (Christian Tomuschat ed., 1995) (reflecting on accomplishments and failures of the United Nations).

62 In the UN system, the International Court of Justice (ICJ) performs this role. See Statute of the International Court of Justice, art. 1, 1946 U.N.Y.B. 843, 843, UN Sales No. 1947.I.18 (establishing the ICJ as the United Nations' judicial organ). For further discussion, see *Fifty Years of the International Court Of Justice* (Malgosia Fitzmaurice and Vaughan Lowe eds., 1996).

63 See Kenneth W. Abbott, *GATT as a Public Institution: The Uruguay Round and Beyond*, 18 Brook. J. Int'l L. 31, 33 (1992) (suggesting that GATT is a public institution with public functions).

64 See Adelle Blackett, *Whither Social Clause? Human Rights, Trade Theory and Treaty Interpretation*, 31 Colum. Hum. Rts. L. Rev. 1, 60 (1999) (noting that the new dispute settlement mechanism calls for interpreting the WTO agreement in accordance with customary public international law rules of interpretation).

65 See Robert Howse and Makau Mutua, *Protecting Human Rights in a Global Economy: Challenges for the WTO* 19 (Int'l Centre for Human Rts. & Democratic Dev., 2000) (arguing that human rights law and humanitarian law require an interpretation of WTO law, particularly that concerning patents in TRIPS, that eases access to AIDS drugs in countries, like South Africa, facing epidemics).

66 See, e.g., Press Release, WTO, 647 Non-Governmental Organizations Eligible to Attend the Doha Ministerial (August 13, 2001) (celebrating widespread NGO interest in Doha conference, but lamenting that logistical concerns limited each NGO to only one representative), available at http://www.wto. org/english/news_e/pres01_e/pr240_e.htm.

67 See, e.g., Paul B. Stephan, *The New International Law – Legitimacy, Accountability, Authority, and Freedom in the New Global Order*, 70 U. Colo. L. Rev. 1555, 1560 (1999) (noting that one example of international obligations forcing domestic agencies to alter their substantive rules would be technocrats from the WTO compelling Americans to "tolerate dirtier gasoline and, ultimately, breathe dirtier air"); Sara Dillon, *Fuji-Kodak, the WTO, and the Death of Domestic Political Constituencies*, 8 Minn. J. Global Trade 197, 207 (1999) ("The dull, technocratic language of WTO law serves to conceal its essential radicalism. A number of the principal cases recently decided at the WTO, however, demonstrate that the WTO is now capable of brusquely undoing the legislative will of national constituencies as expressed in domestic legislation.").

68 See WTO, Community/Forums Page, at http://www.wto.org/english/forums_e/forums_e.htm (last visited July 7, 2003) (containing information and links to discussions about

WTO issues and NGO-related activities at the WTO); see also General Council, Guidelines for Arrangements on Relations with Non-Governmental Organizations, WT/L/162, ¶ 2 (July 23, 1996) (adopting guidelines on relations with NGOs that "recognize[] the role NGOs can play to increase the awareness of the public in respect of WTO activities"), available at http://www.wto.org.

69 See Robert E. Hudec, *The New WTO Dispute Settlement Procedure: An Overview of the First Three Years*, 8 Minn. J. Global Trade 1, 14 (1999) (anticipating that, because international legal systems depend on voluntary member government compliance, the WTO legal system cannot expect unanimous governmental compliance with its legal rulings).

70 See *id.* at 3–4 (noting boldness of the WTO dispute settlement initiative and legal rigor of new procedures).

71 See Montañà i Mora, *supra* note 51, at 178 (explaining that many scholars prefer informalism in international economic law rather than rigid mechanisms).

72 See Hudec, *supra* note 69, at 12 (displaying skepticism that member countries will be receptive to greater legal discipline in the WTO system).

73 See, e.g., Montañà i Mora, *supra* note 51, at 151–3 (voicing misgivings about the plausibility of the WTO appellate review system).

74 See, e.g., Hudec, *supra* note 69, at 14 (suggesting that the first three years of the WTO and its dispute settlement system was a "considerable initial success").

75 See *Success in Singapore, WTO Focus* (World Trade Org., Geneva, Switz.), January 1997, at 1, 1 (reporting Chairman Yeo Cheow Tong's view that the conference was successful in reaching agreements on multilateral trading system and a plan of action for underdeveloped countries), available at http://www.wto.org/ english/res_e/focus_e/focus15_e.pdf.

76 See John H. Jackson, *The Perils of Globalization and the World Trading System*, 24 Fordham Int'l L.J. 371, 375 (2000) (focusing on the "institutional" causes for the "failure of Seattle"). But see Timothy M. Reif and Viji Rangaswami, *Joltin' Joe has Left and Gone Away – Embracing Change: The Way Forward for US Trade Policy and the WTO*, 32 Law & Pol'y Int'l Bus. 427, 437–46 (2001) (arguing that the Seattle debacle occurred because institutional reform of the WTO has not gone far enough in increasing public input, heightening sensitivity to fairness, and implementing legislative processes).

77 See Joseph Kahn, *Swiss Forum has its Focus on Memories from Seattle*, NY Times, January 29, 2000, at C1 (noting that "government officials have stressed that the failure of trade talks owes more to negotiating positions taken by WTO members than to the influence of demonstrators").

78 See Daniel Pruzin, *WTO: Diplomats Say Seattle Scars too Tender to Permit Big Push for WTO Trade Round*, BNA Int'l Trade News Daily, December 15, 1999 (noting that the breakdown of talks in Seattle preventing WTO members from fulfilling their commitment made "at the end of the Uruguay Round in 1994 to begin new negotiations on agriculture and services by 2000"); see also Daniel Pruzin, *WTO: Moore Outlines Incremental Approach Towards New Round of WTO Talks*, BNA Int'l Trade Daily News, February 3, 2000 (describing collapse of meeting after members could not resolve differences over the framework and objectives for agriculture talks, refusal of the United States to consider new negotiations on antidumping rules, and inclusion of labor standards on the WTO's work agenda).

79 Ministerial Conference, Singapore Ministerial Declaration, para. 20, WT/MIN(96)/DEC (December 18, 1996), available at http://www.wto.org, reprinted in 36 I.L.M. 220, 226 (1997).

80 See *WTO: Trade Officials Seek Doha Breakthrough, but Few Changes Occur in WTO Stances*, International Trade Daily, June 7, 2001 (stating that 35–40 of the WTO's 141 members support negotiations on competition, with support strongest among Europe and

Latin America and with opposition strongest from the United States, among others); see also *WTO: Business Groups Say too Early for WTO to Set Framework for Competition Policy*, BNA Int'l Trade Daily News, September 14, 1999.

81 See GATT 1947, pmbl., 61 Stat. at A11, 55 U.N.T.S. at 194 (avowing a commitment to "expanding the production and exchange of goods"); WTO Agreement, pmbl. [Treaties Binder 1, Marrakesh Declaration Booklet] L. & Prac. World Trade Org. (Oceana) at 7, 33 I.L.M. at 1144 (1994) (establishing a commitment to "expanding the production of and trade in goods and services").

82 See, e.g., Office of the United States Trade Representative, Report on Trade Expansion Priorities Pursuant to Executive Order 13116, 66 Fed. Reg. 23064, 23066–7 (May 7, 2001) (identifying 12 recent international trade disputes, including five IP issues).

83 See, e.g., Meredith A. Harper, *International Protection of Intellectual Property Rights in the 1990s: Will Trade Barriers and Pirating Practices in the Audiovisual Industry Continue?*, 25 Cal. W. Int'l L.J. 153, 162 (1994) ("To remain a dynamic force in today's global economy, the United States must abolish trade barriers and foster the growth of the audiovisual industry while simultaneously guaranteeing effective protection for owners of intellectual property rights.").

84 See F. Scott Kieff, *Property Rights and Property Rules for Commercializing Inventions*, 85 Minn. L. Rev. 697, 735 (2001) (asserting that "the power to restrict output is paradoxically essential to providing incentives for the market to generate output" and that "[r]ights of exclusion facilitate efforts for the social ordering and bargaining around inventions that are necessary to generate output").

85 See Laurinda L. Hicks and James R. Holbein, *Convergence of National Intellectual Property Norms in International Trading Agreements*, 12 Am. U. J. Int'l L. & Pol'y 769, 771 (1997) ("There is an inherent conflict . . . between the free circulation of goods and services across countries' borders and the exclusive right of intellectual property owners to explore their creation at the exclusion of others, thus restricting the free circulation of goods and services within the common market."); see also Paul S. Grunzweig, *Note, Prohibiting the Presumption of Market Power for Intellectual Property Rights: The Intellectual Property Antitrust Protection Act of 1989*, 16 J. Corp. L. 103, 103–4 (1990) (noting that "[a]t their core, the policies that drive the patent laws and the antitrust laws of the United States always will conflict").

86 See, e.g., Ruth Gana Okediji, *Copyright and Public Welfare in Global Perspective*, 7 Ind. J. Global Legal Stud. 117, 163–4, 183 (1999) (suggesting that the push for heightened IP protection focuses too much on domestic concerns and largely ignores global economic effects, including disincentives to buy protected products).

87 See, e.g., Jan D'Alessandro, *Note, A Trade-Based Response to Intellectual Property Piracy: A Comprehensive Plan to Aid the Motion Picture Industry*, 76 Geo. L.J. 417, 453–4 (1987) (describing efforts of countries, including the United States, to include IP protection in GATT); Horacio Teran, *Intellectual Property Protection and Offshore Software Development: an Analysis of the US Software Industry*, 2 Minn. Intell. Prop. Rev. 1, 58–9 (2001) (concluding that Westerners' long-term exploitation of indigenous knowledge and resources has reduced the confidence of indigenous peoples in Western system of IP); see also Joel R. Reidenberg, *Trade, TRIPS and NAFTA*, 4 Fordham Intell. Prop. Media & Ent. L.J. 283, 283 (1993) ("While the expansion of intellectual property protection around the world can be attributed to American trade pressure, the trade framework will constrain any country's ability to take unilateral measures against infringements of intellectual property rights."); Marshall A. Leaffer, *Protecting United States Intellectual Property Abroad: Toward a New Multilateralism*, 76 Iowa L. Rev. 273, 282 (1991) (explaining that, although they seek access to IP as important to their development, developing countries

resist allocating resources to enforcement of IP rights because doing so devotes too many resources to protecting foreign interests).

88 See Evelyn Su, *The Winners and the Losers: The Agreement on Trade-Related Aspects of Intellectual Property Rights and its Effects on Developing Countries*, 23 Hous. J. Int'l L. 169, 171 (2000) (noting adoption of TRIPS agreement at 1994 Marrakesh summit).

89 See Håkan Nordström and Scott Vaughan, *WTO Special Study, Trade and Environment* 36–8 (1999) (concluding that "superior environmental performance" does not always reduce profitability).

90 See *id.* at 51–2 (discussing the "tradeoff between production of goods and environmental quality" that changes with increases in income level).

91 See *id.* at 40–6 (suggesting that environmental regulations are not of primary importance in competitiveness and in location decisions).

92 See Clyde Summers, *The Battle in Seattle: Free Trade, Labor Rights, and Societal Values*, 22 U. Pa. J. Int'l Econ. L. 61, 67 (2001) (arguing that "[o]bservance of these core labor rights" such as those prohibiting forced labor, child labor and discrimination "would have minimal impact on labor costs"). Increased freedom of association may well lead to increased union formation, which will then lead to more success in the introduction of certain concessions to labor, such as increased wages and safety measures and shorter workweeks. These concessions would almost certainly raise the marginal cost of labor. A better argument, therefore, is that increased productivity stemming from improved employment conditions will increase profit more than any increased labor costs.

93 See *Organization for Econ. Co-operation & Dev., Trade, Employment, and Labour Standards: A Study of Core Workers' Rights and International Trade* 80–2 (1996) (discussing how core labor standards can improve efficiency of market outcomes); see also Blackett, *supra* note 64, at 49 (suggesting that "higher labor standards may increase competitiveness and productivity").

94 See Klaus Töpfer, *Statement to the WTO Symposium on Trade, Environment and Sustainable Development* (March 17, 1998) (stating that trade can influence world conditions to encourage environmental sustainability and social justice), available at http://www.grida.no/news/index.cfm?requestedItemId=588 (last visited July 7, 2003).

95 Palmer, *supra* note 28, at 259.

96 This statement should not be mistaken for the proposition that there is a monolithic point of view held by the developing world. A wide range of views on linkage exist, of course, both within developing countries and between them. See Chantal Thomas, *Should the WTO Incorporate Labor and Environmental Standards?*, Washington & Lee Law Review (forthcoming). Notwithstanding this complexity, the most salient voices have been those of developing-country governments strongly opposed to linkage of trade with labor and environmental standards within the WTO.

97 See *Third World Intellectuals and NGOs Statement Against Linkage (TWIN-SAL)*, available on Professor Jagdish Bhagwati's university webpage, www.columbia.edu/~jb38/TWIN_SAL.pdf.

98 See Robert E. Hudec, *Developing Countries in the GATT Legal System* 143–5 (1987).

99 See Brenda A. Jacobs, *Integration of the US Textile Quota Program into WTO Rules, in Practicing Law Inst.*, The Commerce Department Speaks on International Trade & Investment 359, 361 (1998) ("Despite all the fiery rhetoric and loud laments by the US domestic industry, three years after implementation of the WTO Agreement on Textiles and Clothing (ATC), a point by which the gradual process should be yielding some tangible effects, in fact import restraints in this sector are tighter than ever."); Mark Ritchie and Kristin Dawkins, *WTO Food and Agricultural Rules: Sustainable Agriculture and*

the Human Right to Food, 9 Minn. J. Global Trade 9, 10 (2000) {noting that "the intensity and sensitivity of agricultural issues is greater than in the past and that agricultural issues would be the dominant and most difficult issue" in upcoming WTO negotiations [citing Secretary of Agriculture Dan Glickman, Speech to the Center for Strategic and International Studies (June 16, 1999)]}.

100 For instance, the United States has failed to approve several key labor and environmental agreements. See John J. Fialka, *Nations Approve Rules for Kyoto Pact Without US*, Wall St. J., July 24, 2001, at A2 (reporting US objections to international environmental agreement on climate change); see also Virginia A. Leary, *Workers' Rights and International Trade: The Social Clause (GATT, ILO, NAFTA, US Laws)*, *in* 2 Fair Trade and Harmonization: Prerequisites for Free Trade? – Legal Analysis 177, 188 (Jagdish Bhagwati and Robert E. Hudec eds., 1996) (describing the "shameful record that places the United States among the countries that have ratified the fewest ILO conventions").

101 Of course, there are many reasons to believe that poor countries, or at least significant populations therein, stand to benefit from such protections.

102 Indian Proposal says Developing Countries should get Patent, Trademark Concessions, 6 Int'l Trade Rep. (BNA) 953, 953 (July 19, 1989) (reporting that Third World countries, led by Brazil and India, argued that IP "should be dealt with by the World Intellectual Property Organization" rather than be included in GATT).

103 See Chantal Thomas, *Balance-of-Payments Crises in the Developing World: Balancing Trade, Finance and Development in the New Economic Order*, 15 Am. U. Int'l L. Rev. 1249, 1260–1 (2000) (discussing political trends leading to deferential approach by industrialized countries toward Third World nations in applying GATT with cognizance of political issues facing those states).

104 See Agreement on Agriculture, April 15, 1994, WTO Agreement, Annex 1A [Treaties Binder 1, Treaties Booklet 1] L. & Prac. World Trade Org. (Oceana) 27, 27 (March 1995) (calling on developed countries to take into account the "particular needs and conditions" of developing countries in increasing markets for their agricultural products); Agreement on Textiles and Clothing, April 15, 1994, WTO Agreement, Annex 1A [Treaties Binder 1, Treaties Booklet 1] L. & Prac. World Trade Org. (Oceana) 77, 77 (March 1995) (recalling agreement that "special treatment should be accorded to the least-developed countr[ies]").

105 See Edmund W. Sim, *Derailing the Fast-Track for International Trade Agreements*, 5 Fla. Int'l L.J. 471, 484–5 (1990) (noting that "[e]mbittered textile lobbyists threatened to 'do whatever is necessary"' to forestall concessions on textiles, and that others were "implacable opponents of the Uruguay Round," although some agricultural sectors supported trade negotiations in the hopes of increasing their overseas market share).

106 See Weiss, *supra* note 31, at 814–15 (outlining areas of concern and priorities relating to sustainable development and implementing economic progress while using resources in clean and efficient manner).

107 See Montreal Protocol, arts. 10, 12, 26 I.L.M at 1557–9 (discussing the funding and technical assistance mechanisms of the Montreal Protocol). For a discussion of the cost of compliance, framed as transfer of "administrative technology," see Chantal Thomas, *Transfer of Technology in the Contemporary International Order*, 22 Fordham Int'l L.J. 2096 (1999).

108 However, it is important to note that, in many cases, integrating trade, labor, and environmental regulations would impose *no new obligations* on member states. Particularly with respect to the most salient international agreements on labor and the environment, most rich and poor countries have reached agreement. See ILO, International Labour Standards Ratifications, at http://webfusion.ilo.org/public/db/standards/normes/appl/ (listing current

status of labor agreement ratification); see also UNEP, Environmental Conventions, at http://www.unep.org/SEC/env3.htm (listing current status of environmental convention ratification). The fact that poor countries have already formally assumed these obligations seems to weaken any legitimate protest on their part to integrating into a regime that might provide more effective enforcement. See Clyde Summers, *The Battle in Seattle: Free Trade, Labor Rights, and Societal Values*, 22 U. Pa. J. Int'l Econ. L. 61, 67 (2001) ["It is ironic that Egypt, Brazil, Indonesia, and Pakistan, which were among those most vocal in opposing the (proposal to incorporate labor discussions into WTO talks in Seattle) have ratified conventions on all of these subjects, with the exception of Pakistan's failure to ratify a convention on child labor. They expressed outrage that (President) Clinton would suggest (in his proposal on labor) that they should be required to observe the conventions that they had ratified."] As Jose Alvarez has argued, however, state accession to international law instruments is often conditional precisely on their relatively weak enforceability and the relatively high autonomy in determining compliance that results. See Jose E. Alvarez, *How not to Link: Institutional Conundrums of an Expanded Trade Regime*, 7 Widener L. Symp. J. 1, 1 (2001) (noting that the WTO is rare among international law regimes because it "secures at least procedural (if not always substantive) compliance"); see also Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 Yale L.J. 1935 (2002).

109 See Chantal Thomas, *Trade-Related Labor and Environment Agreements?* 5 Journal of International Economic Law 791, 799–816 (2002).

5

A Comparative Analysis of Compliance Institutions in International Trade Law and International Environmental Law

Brett Frischmann

Chapter Outline

This chapter compares the strategic, institutional approaches taken in international trade and international environmental agreements and explains the role of "compliance institutions" in facilitating coordination of State behavior. After a brief introduction, the chapter sets forth a broadly applicable, game theoretic framework for analyzing and comparing the ways in which States deal with international compliance dilemmas in different contexts. The second part of the chapter applies the framework in two sections: the first applies the analytic framework to international trade, focusing on the GATT/WTO system and its compliance institutions. The second applies the analytic framework to international environmental law, focusing on the international regime that regulates ozone-depleting substances and its compliance institutions. Attention is given to these regimes and compliance institutions because they are often considered as models for the development of institutions in related areas of international law and are increasingly the focal point of interdisciplinary legal issues.

1 Introduction

States cooperate and create international agreements to deal with problems that arise as a result of their interdependence. To the extent that the actions and decisions of one State (or its citizens) may affect the welfare of another State (or its citizens), negatively or positively, there is room for negotiation and potential cooperation among the two States to minimize joint costs or maximize joint benefits. International trade and international environmental law involve economic, ecological, technological, and social interdependencies that make international cooperation desirable. If transaction costs were zero, information was perfect, and States were of equal bargaining power, we might expect effective agreements to be reached frequently. In the real world, however, there are impediments to cooperation, including *transaction costs* associated with identifying interdependencies and opportunities to cooperate, communicating preferences, obtaining "approval" of domestic constituents, negotiating and implementing a workable agreement, monitoring, facilitating, and enforcing compliance; *uncertainties* associated with payoffs, time horizons, and information generally; and *strategic considerations* associated with static and dynamic bargaining positions, asymmetric information, asymmetric preferences, and opportunism through actual or threatened intentional noncompliance. In the absence of a supranational government, States often find themselves in need of governance structures, or "institutions," to facilitate cooperation, prevent and resolve conflicts, and generally serve as a locus for coordination, information sharing, and other important activities that assist States in overcoming these impediments.

States often cooperate by creating international law and by complying with the commitments they have undertaken. We do not live in a world dominated by anarchy, but rather we live in a world in which international rules, norms, and other products of cooperative decision making effectively constrain the behavior of States and private actors in many areas. Even when States are not perfectly constrained and find it necessary to breach or not comply fully with an obligation, the existence of international law influences the manner in which and degree to which States deviate from the rule. Moreover, beyond the law itself, other institutions play a significant role in discouraging a breach or encouraging compliance and affecting the manner in and degree to which States breach or comply with international commitments. This chapter develops an analytical framework for comparing legal regimes and their evolution, with a particular emphasis on the role of compliance institutions.

At the outset, it is necessary to explain what is meant by the term "compliance." Compliance is the degree to which a State behaves in a manner that conforms to its legal obligations (Jacobson and Brown Weiss, 1998a; Victor et al., 1998). Scholars have drawn important distinctions between treaty implementation, compliance, and effectiveness (Jacobson and Brown Weiss, 1998a, pp. 1, 4–5 and table 1). Generally, implementation refers to the (domestic) actions that States take to give effect to international commitments, and effectiveness refers to the degree to which the treaty objectives are met (Jacobson and Brown Weiss, 1998a, pp. 1, 4–5 and table 1). Although strict compliance with treaty obligations alone does not ensure that the underlying problem motivating international cooperation will be effectively addressed,[1] this chapter

nonetheless focuses on compliance, adopting "the widespread belief that an agreement is likely to be more effective the greater the degree to which its parties comply with its obligations" (Knox, 2001, p. 1, n. 1). Moreover, as described below, "compliance institutions" are broadly defined to encompass institutions designed (by parties) to modify parties' incentives to comply with obligations, including those institutions that modify parties' obligations over time so as to facilitate compliance and improve the effectiveness of the regime.[2]

States entering into an international agreement perceive *ex ante* an *ex post compliance dilemma*. In simple terms, given the lack of supranational adjudicative and enforcement bodies to "guarantee" compliance,[3] there is a risk to State A (and other parties) that State B will sign on but not comply with its obligations. Of course, State B faces the risk that State A (and other parties) will do the same. The perception of this dilemma prior to negotiations may affect negotiations primarily in two dimensions. First, the perceived dilemma may affect the substantive commitments that States are willing to undertake – possibly leading to overblown commitments that some parties do not intend to honor fully or to minimal commitments that may easily be met. On the one hand, opportunistic States may attempt to push the level of commitments higher (than would be the case in the absence of the compliance dilemma) so as to free ride on the actions of others or simply to experience relative gains by not bearing the same costs as others.[4] On the other hand, nonopportunistic States may push the level of commitments lower to counteract such pressures. Of course, the degree to which strategic misrepresentations and attendant reactions cause negotiations to diverge from the commitment levels that would result in the absence of the compliance dilemma depends on the information held by the negotiators regarding each other's preferences, intentions, and capabilities. Second, in addition to affecting substantive commitments, perception of the compliance dilemma may prompt States to set forth dispute resolution and/or other compliance procedures in international agreements. Some such institutions dynamically alter both the incentives to comply with existing commitments and the commitments themselves. The approach taken in various international agreements for dealing with the compliance dilemma varies considerably.

This chapter begins to compare the strategic, institutional approaches taken in international trade and international environmental agreements and attempts to explain the role of "compliance institutions" in facilitating coordination of State behavior. Accordingly, after a brief introduction, the chapter sets forth an analytic framework that provides a useful tool for comparing the ways in which States deal with the compliance dilemma in different contexts. The chapter then applies the framework to compare compliance institutions in international trade and international environmental regimes. This latter part of the chapter is necessarily a preliminary assessment; a complete analysis is simply beyond the scope of this chapter and is left for future work.

International trade and international environmental law reflect considerably different evolutions of international law in part due to evolving institutional approaches to compliance. For example, international trade law has evolved into a relatively *strong* version of public international law in large part because of the evolution of innovative compliance institutions, particularly the WTO dispute settlement understanding. This compliance institution has enabled States to effectively coordinate their behavior around explicit substantive legal obligations with a credible means of detecting and

evaluating violations and prospectively sanctioning on-going violations. This institution embodies a legalistic approach to international cooperation, which works well in the trade context for various reasons associated with the nature of the underlying dilemma faced by States.

Recent innovations in international environmental law have also enabled States to effectively coordinate their behavior, but not typically in the legalistic manner accomplished in trade. Instead, progress in international environmental "regimes"[5] involves varying degrees of binding and nonbinding commitments and multifaceted compliance systems. The compliance systems tend to be facilitative and directed at encouraging compliance through positive incentives rather than discouraging noncompliance through the possibility of sanctions. This "management-oriented" approach to international cooperation works well in the environmental context for various reasons associated with the nature of the underlying problem motivating cooperation among States.

Interestingly and importantly, both international trade law and international environmental law are still evolving and, in a sense, "learning" from each other and adopting successful institutional innovations made in one issue area to work to the benefit of the other. For example, the WTO regime increasingly relies on managerial-type institutions to encourage cooperation in pursuit of trade liberalization by existing and potential members. These institutional developments range from the Trade Policy Review Mechanism, which periodically evaluates domestic implementation and compliance with trading rules, to capacity-building programs and financial and technical assistance.[6] Furthermore, international environmental regimes increasingly include enforcement-oriented institutions, ranging from legalistic adjudication to compliance panels with authority to impose sanctions of various types.

The structure of this chapter is as follows. Section 2 sets forth a game theoretic framework for analyzing State cooperation in framing, forming, and complying with international agreements. Section 3 then applies the framework in two subsections: the first applies the analytic framework to international trade, focusing on the GATT/WTO system and its compliance institutions. The second applies the analytic framework to international environmental law, focusing on the international regime that regulates ozone-depleting substances (the "Ozone regime") and its compliance institutions. Attention is given to these regimes and compliance institutions because they are often considered as models for the development of institutions in related areas of international law and are increasingly the focal point of interdisciplinary legal issues.

2 Analytic Framework

2.1 Introduction

Game theory provides a useful framework for analyzing international cooperation and institutions.[7] It provides a relatively straightforward model of decision making by entities in situations where their decisions are interdependent and they face conflicting strategic incentives. The primary variables are the players (States), possible actions

(cooperate/comply or defect/breach), and outcomes (payoffs), which together constitute the rules of the game. Other important factors include information, strategies, and the number of iterations (how many times will the game be played). When the rules of the game are set, one can predict an equilibrium (or set of equilibriums), in terms of expected actions and outcomes. International cooperation has often been analyzed as a Prisoners' Dilemma or some variant thereof, described below.

As with any theoretical framework, it is important to recognize that there are a number of assumptions in play. In fact, consideration of the assumptions themselves provides insight into the complexities that shape international cooperation. Game theoretic analysis focuses on States as "players" and assumes that States are unitary rational actors acting to maximize their "individual" welfare. Of course, this assumption builds on additional assumptions that may not hold up perfectly in all cases. Although the point here is not to debate the fine details, it is important to recognize that (1) the actions taken by States are the products of complicated domestic and international processes involving public and private actors and institutions; (2) numerous public actors within State governments have active roles in international negotiations, implementation of agreements, decisions whether and to what extent to comply with commitments, and other State actions;[8] and (3) defining a State's "individual welfare" is itself a complicated task, particularly given the fact that States sometimes act in the best interest of a particular domestic group rather than the country as a whole.

Acknowledging these complications, this chapter nonetheless folds them into the game theoretic framework. In reality, States are the dominant (although by no means exclusive) actors in international relations and the primary sources of international law. The international agreements considered in this chapter are agreements among States to undertake commitments and to create compliance institutions. Thus, as far as commitments and compliance are concerned, it seems reasonable to say that ultimately, "the State" undertakes commitments and "the State" either complies or does not comply with its commitments. Although not discussed in detail here, international agreements create complicated webs of legal (and nonlegal) relationships such that compliance with and the effectiveness of an agreement may depend on the actions of both States and private actors.[9] While actions by domestic (public or private) actors may lead to State noncompliance, States remain "responsible" for the commitments undertaken in an international agreement (of course, what "responsibility" means depends on the compliance institutions involved).

States make decisions based on some notion of expected payoffs (i.e., cost–benefit or welfare analysis). It is clear that decisions are not made solely on the basis of "national welfare" or some objective evaluation of aggregate individual preferences within a country, because domestic processes are imperfect and often involve a competition among domestic preferences.[10] Moreover, because States act through people, such as politicians and bureaucrats, the individual preferences and biases of those people may skew perceived payoffs from particular decisions away from the socially optimal ideal.[11] Again, the point here is not to fully model the processes by which payoffs are perceived by States (or delegated decision makers). It should be sufficient for the purposes of this chapter to highlight that States make decisions based on an estimation of expected payoffs and that ultimate payoffs are uncertain and are subject to continuous reevaluation by relevant decision makers. The extent to which the relevant measure of

payoffs is "national welfare" as opposed to "politicians' welfare" (or the most powerful domestic interest's welfare) varies among States and across issue areas. In fact, as described below, the payoff structure in both international trade and international environmental games depends on both measures: the payoffs associated with cooperation are generally diffuse, widespread, long term, and arguably more closely linked to perceptions of national welfare while the payoffs associated with defection are generally concentrated among particular industry groups, short term, and more closely linked to the welfare of politicians and special interests. Rather than belabor the point that there are important limitations in the underlying assumptions that States are unitary rational actors that make decisions with the intent to maximize national welfare, it is sufficient for the purposes of this chapter to "soften" the assumptions, flexibly apply the game theoretic framework, and keep in mind its limitations.[12]

2.2 The Basic Games

The Prisoners' Dilemma is a commonly told story that describes the difficulty of getting two parties to cooperate when the action taken by each party affects the welfare of the other (i.e., interdependency).[13] Suppose that two parties, **i** and **j**, face a decision regarding whether or not to cooperate with each other and that the following payoffs are expected by each:

Prisoners' Dilemma		**Party j**	
		Cooperate	Defect
Party i	Cooperate	3, 3	1, 4
	Defect	4, 1	2, 2

If both parties cooperate, each receives a payoff of 3 for a total of 6. If both parties do not cooperate (or defect), then each only receives a payoff of 2, for a total of 4. Finally, if one party cooperates and the other defects, then the cooperator receives a payoff of 1 and the defector receives a payoff of 4, for a total of 5. The mutually beneficial solution is for both to cooperate and receive joint benefits of 6. However, each party knows that it can get a better individual payoff if it defects and the other party cooperates. Moreover, each party knows that if it cooperates and the other party defects, then the cooperator will only get a payoff of 1, and therefore it would have been better off had it also defected. In other words, *if* each party knows what the payoff structure looks like *ex ante*[14] and must decide whether or not to cooperate, each party will reach the conclusion that it is in the party's best interest to defect – if party **i** cooperates, party **j** is better off defecting ($4 > 3$), and if party **i** defects, party **j** is better off defecting ($2 > 1$). Thus, in the end, each party has a *dominant strategy to defect* (Carlson, 2000; Axelrod, 1984). Achieving cooperation in the face of this incentive structure requires mutual assurance in some form or other. Where an enforceable agreement between the parties can be reached (e.g., a contract between two individuals enforceable under domestic law), cooperation is possible, although not guaranteed (transaction

costs, uncertainty, opportunistic behavior, strategic bargaining, and other factors may inhibit effective cooperation by contracting). Other forms of assurance mechanisms include reputation and the expectation of future interactions between the same parties as well as linking the outcome in a particular game to actions taken in other unrelated games between the parties. Importantly, assurance mechanisms are intended to make cooperation the dominant strategy by altering the expected payoffs, whether by expected damages for breach of contract, reputational harm, or costs inflicted in future iterations or other linked games in the event of defection, or by expected side-payments for contractual performance, reputational gains, or benefits gained in future iterations or other linked games (perhaps through the promise of continued cooperation) in the event of cooperation.

The counterpart to the Prisoners' Dilemma game is a game called Harmony.[15] Again, suppose that two parties, **i** and **j**, face a decision regarding whether or not to cooperate with each other but that the following payoffs are expected by each:

Harmony

		Party j	
		Cooperate	Defect
Party i	Cooperate	4, 4	2, 3
	Defect	3, 2	1, 1

In contrast to the Prisoners' Dilemma, both parties have the dominant strategy to cooperate, hence, the name Harmony. The primary impediment to cooperation where the interdependence among States gives rise to a Harmony game is identification of such a situation. In other words, once identified, States should be able to reach an agreement without much concern over (non)compliance. Coordination or standardization problems give rise to a Harmony game. International aviation rules are often cited as an example (Raustiala, 2000, p. 400; Chayes et al., 1998). Once the rules are established and "parties understand the rules, no actors have incentives to violate them" (Raustiala, 2000, p. 400; Chayes et al., 1998). Moreover, due to potential network externalities, nonparties have an incentive to adopt the rules as well.

The Prisoners' Dilemma and Harmony represent two extremes where players have symmetric dominant strategies and incentives to negotiate a cooperative solution.[16] There are variations of the basic two-player scenario that involve asymmetries in strategies. For example, in the game called Suasion, party **j** has a dominant strategy to defect but party **i** does not have a dominant strategy.

Suasion

		Party j	
		Cooperate	Defect
Party i	Cooperate	4, 3	2, 4
	Defect	5, 1	1, 2

If party **i** knows what the payoff structure looks like though, it will cooperate because it expects party **j** to defect. In order to maximize joint benefits through mutual cooperation, party **i** must make a side-payment to party **j** to adjust the payoff structure. Finally, in the game called Chicken, neither party has a dominant strategy.[17] Although maximizing joint benefits requires that at least one party cooperate, each party would prefer to defect and allow the other party to cooperate and thereby realize a relative gain.

These simple games highlight the importance of interdependent decision making and the need for identifying problem structures and opportunities to devise cooperative solutions. Of course, the real world is not so easily simplified. The rules of the game – number and identity of participants, range of options, payoff structures, number of iterations (or times that the game is played in succession) – are often considerably more complex, as are the mechanisms used to coordinate behavior.

First, the number and identity of (potential) participants are particularly important considerations for compliance purposes. In some cases, the scope of the interdependency problem may require the cooperation of only a small number of States while in other cases a large coalition of States may be necessary. Generally, it is easier to coordinate behavior among a smaller group, in part because it is easier to bargain and in part because it is easier to monitor compliance and overcome the compliance dilemma. Apart from the numbers, the importance and identity of particular participants may stifle or facilitate cooperation. The participation of a particular State (or group of States) may be essential in successfully dealing with a particular problem, in which case that State gains leverage that can be exploited during negotiations and potentially re-exploited thereafter depending on the nature of the compliance institutions adopted. In addition, the presence of a State (or group of States) with the resources, power, and will to encourage/coerce participation during negotiations; to provide administrative, financial, or technical assistance; or to threaten/impose unilateral sanctions in response to noncompliance may facilitate cooperation.[18] In some cases, a paradox of sorts may arise in putting together an international coalition to address a particular problem. On the one hand, maximizing the number of participants may be the primary objective of negotiations. This may be due to the global scope of the problem, as in the case of global warming or ozone depletion, where solving the problem may be impossible without broad-based participation or where nonparticipants can effectively hold up participating States.[19] On the other hand, States may find it necessary to constrain the number of participants to a manageable number for compliance-management purposes.[20] As the number of States increase, it may become more difficult to detect (potential) defections and either encourage compliance through facilitative measures or retaliation. Moreover, transaction costs and uncertainties may increase with the number of participants and thereby inhibit cooperation, for example, by making it more difficult to estimate payoffs due to more complicated interdependent welfare functions. The equilibrium position between these opposing objectives will vary depending upon the nature of the underlying problem motivating international cooperation, the particular States involved and their intent and capacity to comply, and their willingness/ability to coordinate *ex ante* in a sufficiently credible manner.

Second, the range of options or possible actions that participants in a game may take are not always binary (i.e., comply or not comply) and may vary along many

dimensions. There may be numerous obligations of varying degrees of importance (to the coalition as a whole and to a particular party) within a negotiated treaty structure; compliance with particular obligations may be a higher priority for some participants than for others; and parties may lack the capacity (administrative, economic, scientific, political, or otherwise) to comply with particular commitments. In the end, compliance is often a matter of degree along a continuum. Moreover, there is a significant difference between intentional noncompliance (or opportunistic noncompliance), which some scholars have argued is less frequent than presumed to be the case, and unintentional noncompliance, which may result from the lack of capacity (Chayes and Chayes 1995).

Third, and related to variability in possible options available to parties, the payoff structures in the real world differ in important ways from the two-player games outlined above. First, the magnitude and relative value of payoffs may vary considerably. Certain players may benefit more than the others from mutual cooperation, perhaps by orders of magnitude. This may lead to concerns over the equitable distribution of gains among cooperators as well as opportunities for strategic holdout (i.e., refusal to cooperate unless gains are shared). In the international environmental area, for example, distributional concerns are particularly acute for a few reasons: not all countries have consumed or polluted natural resources in the same manner or to the same degree; the present welfare status of countries and their citizens is not uniform but rather is markedly differentiated; and the value systems and cultural preferences of domestic constituencies varies considerably among countries. These factors, among others, bring distributional issues to the forefront of international environmental law, as demonstrated prominently by Principle 7 of the Rio Declaration on Environment and Development that, "In view of the different contributions to global environmental degradation, States have common but differentiated responsibilities."[21] Second, the interdependencies among players may complicate payoff structures considerably based on the number of players and the relative importance of each player's actions to every other player's welfare. This may lead to the emergence of numerous conflicting strategies and/or the formation of internal coalitions with strategic advantages.

Finally, the number of iterations (or as Robert Axelrod has dubbed the notion, the "shadow of the future") is also an extremely important consideration for compliance purposes (Axelrod, 1984). Generally, participation in an international legal regime involves a continuous series of decisions regarding compliance such that the game is repeated indefinitely unless and until a party withdraws completely. The "shadow of the future" refers to the fact that players' expectations regarding the future – specifically, future iterations of the game – affect decisions made in the present depending, of course, on discount factors and the possession of information regarding future payoffs and that players may adopt strategies that depend on other players' performance in previous rounds ("reciprocal strategies"). Perhaps the most famous reciprocal strategy (used in the Prisoners' Dilemma context) is "tit-for-tat," where a player cooperates during the first round and thereafter commits to do whatever the other player did in the previous round (Axelrod, 1984). By punishing defection and rewarding cooperation, the tit-for-tat strategy facilitates cooperation when the other player sufficiently values future payoffs because it allows players to signal to each other their strategy and their willingness to reciprocate either cooperation or defection. It is important to note that there are numerous other strategies that players may adopt, including, *inter alia*,

penance strategies, where after one player defects, other players remain in defection with respect to the initial defector until the defector cooperates (at which point the players return to cooperation),[22] grim trigger strategies, where all players remain in noncompliance forever after noncompliance by a single player is observed,[23] and trigger strategies with (exogenous or endogenous) periods of penalty following an instance of noncompliance.[24]

2.3 Framing, Forming, and Playing Games

International cooperation is the product of a complex, dynamic process that is rife with collective action problems, strategic behavior, transaction costs, and uncertainty. This process may be summarized as follows: (1) States identify an interdependency problem, delineate its contours and possible solutions ("framing the game"); (2) negotiate an agreement and create compliance institutions ("forming a game"); and (3) implement the agreement ("playing the game").[25] At each stage, there are repeated, dynamic interactions between domestic and international, public and private actors.

2.3.1 Framing the Game

The first step toward cooperation involves identification of the interdependency problem (or set of problems), potential participants, options, expected payoffs, duration (number of iterations), and other factors that define the rules of the game. This part of the process may take years and may involve a host of different actors, especially in the environmental area, where problem identification and definition may be extremely complex and may involve uncertainty in various dimensions (scientific, economic, political) relevant to decision making. It is during this stage that domestic and international policy makers, bureaucrats, NGOs, businesses, and other interested parties home in on an issue and begin to debate the merits of potential (international) solutions. In some cases, problem identification, issue framing, and informal negotiations may occur at the sub-State level for years before formal diplomatic negotiations take place. When an issue has risen among national priorities to the point where a State is willing to invest the time and necessary resources to fully explore a cooperative solution, negotiations may begin.

2.3.2 Game Formation

Once an interdependency problem has been identified and potential participants decide that taking action as opposed to maintaining the status quo may be necessary, States negotiate some form of agreement, for example, a treaty,[26] establishing the rules of the game. This step affects the expected payoffs for participants by virtue of the commitments undertaken, concessions made and compliance system established.[27] The negotiation of an agreement is a dynamic process in itself that redefines the rules of the game: while the preliminary game structure may be framed simply in terms of two options (maintaining the status quo or doing something), negotiations modify the initial payoff structure by bringing to light the full range of options with associated

costs, benefits, timeframes, and distributional patterns. Thus, negotiations give form and substance to potential commitments expanding the range of options along various dimensions, such as, *inter alia*, the depth of substantive commitments (e.g., to liberalize trade or to regulate emissions of environmentally harmful substances by specific quantitative measures) and the time frames over which commitments must be satisfied. Notably, international negotiations tend to be quite different from legislative negotiations at the domestic level because international negotiations generally aim toward consensus rather than majority approval.

When States negotiate over the details of an agreement, a wide range of considerations must be taken into account.[28] While the substantive commitments often deserve the most attention, institutional design is an incredibly important aspect of negotiations. Most importantly for our consideration in this chapter, States are fully aware of the noncompliance risks prior to undertaking substantive commitments[29] and may work out the details of a compliance system while negotiating substantive commitments or condition their acceptance of substantive commitments on the development of an acceptable compliance system. Of course, as described below, what States consider to be an "acceptable compliance system" depends on the agreement and varies across issue areas.

2.3.3 Play

Once the game is framed and formed, "play" begins. Participants cooperate by complying with their commitments and defect by not complying with them. Games are iterative, meaning that participants repeatedly find themselves faced with the question of whether or not to comply, unless they choose to withdraw, are suspended, or otherwise forced to withdraw. Importantly, depending on the underlying dilemma, games may reframe and reform dynamically during play, affecting payoff structures, the opportunities for opportunism, and the bargaining/strategic positions of players. In other words, these three "stages" are not independent, and the cooperative process is not necessarily linear. First, a better understanding of the underlying problem may emerge, for example, through advancements in scientific understanding of an environmental threat. This may alter the perceived payoff structure and prompt parties to revisit the existing agreement.[30] Second, the expected costs (benefits) for implementing a particular obligation may be less than expected, perhaps promoting parties to ratchet up (down) existing commitments. Third, exogenous factors may alter the existing payoff structure or the relative priority of a given commitment. The degree to which games reframe and reform during play (i.e., the "dynamicism" of the process) is an important factor in the design of compliance institutions, which is highlighted below in the context of multilateral trade and environmental agreements.

The advantage of looking at international agreements through the lens of game theory is that the strategic considerations embodied in the agreement and its constituent commitments can be explored. During negotiations of an agreement, each party must consider the effects of its and other countries' compliance with the agreement on its citizens, as a whole and taking into account distributional implications, and also must contemplate both the likelihood and potential impact of opportunistic defection. In some cases, the *ex ante* balancing of expected costs and benefits may lead to outright rejection

of any agreement because the costs and opportunism risks vastly outweigh the benefits; in other cases, the opposite may occur – for example, easy coordination problems. In the middle, there is a class of situations where the balancing suggests that agreement may be feasible provided that a compliance system is integrated into the agreement, which of course entails its own cost–benefit analysis. Analysis of the strategic considerations, i.e., the rules of the game, in international trade and international environmental regimes provides insights into the choices made regarding compliance institutions, i.e., the means by which the rules of the game are manipulated to encourage cooperation or discourage defection.

2.4 Compliance Strategies

Recognizing *ex ante* that potential compliance problems may plague efforts to address an interdependency problem, States may reach agreement on the design of a compliance system that institutionalizes a suite of compliance strategies. Compliance systems can be thought of as a systematic method of adjusting the rules of the game in either an evolving game or a series of games. As the tremendous volume of literature discussing monitoring and enforcement institutions indicates, the strategies discussed below may be implemented through a wide range of institutions. The focus here is primarily on function rather than institutional form.

Recent research on national compliance with international commitments suggests that national *intent* and national *capacity* to comply are critical variables for actual compliance. "The level of State compliance depends on having the leaders and the citizenry understand that it is in their self-interest to comply and on their acting on this understanding. While external pressures and assistance can push a country toward compliance, there is no substitute for engaging 'self-interest.' In some cases, the issue is one of prioritization: How much of the country's national resources to devote to complying with particular agreements. This is especially difficult when compliance requires States to coordinate actions among several powerful ministries, with provincial and local governments, or with powerful business and industry organizations" (Brown Weiss, 2000b, p. 458). Furthermore, the capacity to comply with international agreements depends on numerous assets, such as "[a]n honest and effective bureaucracy, economic resources, technical expertise, and public support."[31] Importantly, these variables change over time, and the variables are subject to endogenous and exogenous influence. Thus, cooperation depends in part upon adjusting the individual benefit–cost structures faced by States to affect the intent of parties and in part on building the capacity of participants so that compliance is feasible.[32] As Edith Brown Weiss explains, "[p]arties should have a suite of compliance methods available that can be tailored to meet the needs of particular countries for particular agreements" (Brown Weiss, 2000b, p. 458; Brown Weiss, 1999, p. 1555).

While compliance strategies may be grouped and classified in a number of different ways,[33] the theory developed in this chapter suggests that compliance strategies may be analyzed based on the aspect of the game targeted by the strategies. Specifically, compliance strategies may target (1) expected payoffs, or the payoff structure itself ("Type I strategies"); (2) the cooperation-inhibiting factors, such as transaction costs

and uncertainties, that are not contingent upon the decisions of players ("Type II strategies"); and (3) dynamic adjustment (reframing and reforming) of the game in response to exogenous events unrelated to the decisions of players ("Type III strategies").

Through Type I strategies, States may directly target the payoff structure itself by making concessions or side-payments to encourage cooperation and offset the perceived benefits of defection ("positive incentives") or by committing parties to retaliate against defection through some form of sanction ("negative incentives"). Positive incentives tend to influence both national capacity and national intent, for example, by lowering the costs of implementation, while negative incentives tend to influence only national intent. Thus, where a potential participant lacks the capacity to cooperate, positive incentives become an essential ingredient in the mix of institutional mechanisms incorporated into an agreement or negotiated on the side. Importantly, the promise of side-payments (or threat of sanction) itself effectively adjusts the expected payoffs and thus acts as an inducement (deterrent). Thus, both positive and negative incentives are aimed at making compliance with commitments more attractive than noncompliance; in addition, adjusting the payoff structure may be aimed at making the distribution of net payoffs equitable in the eyes of participants.

Generally, incentive-based strategies depend on compliance over time (over the course of a series of games, referred to as a "supergame") and a necessary component to the success of such strategies is information regarding parties' performance.[34] To facilitate such strategies, States may develop monitoring systems to improve the likelihood that cheating will be detected, thereby also adjusting the expected payoffs. An important source of information regarding State behavior is the State itself. For example, most international environmental agreements that have a monitoring system rely on national reporting. Like the substantive obligations undertaken by States, reporting obligations similarly raise compliance problems. That is, States do not always provide full and accurate reports. Of course, this is not surprising. What distinguishes one monitoring system from another is the degree to which the information can be verified and reviewed.[35] Beyond national reporting obligations, international monitoring systems can be designed in many different ways ranging from the creation of a formal intergovernmental organization ("IGO"), delegation of monitoring responsibilities to an existing organization (IGO or NGO), or other institutional mechanisms. The scope of monitoring responsibilities delegated to an organization may vary considerably. For example, parties may limit the scope to gathering and disseminating information to the parties and/or the public generally. Or parties may expand the scope to include a more substantive role in assessing and verifying compliance. Informal monitoring may also arise independently through the actions of nongovernment actors. For example, in the trade context, importers and exporters regularly monitor the State actions as a necessary consequence of doing business. And in the environmental context, environmental NGOs frequently play an important role by voluntarily monitoring the compliance of States with their international commitments and publicizing reports of noncompliance. Transparency requirements further assist compliance monitoring. For example, pursuant to GATT Article X, WTO members must publish laws, regulations, judicial decisions, and administrative rulings of general application that pertain to classification, valuation, duty rates, taxes or other charges or requirements, restrictions or prohibitions on import or export; or that affect the sale,

distribution or transportation, insurance, warehousing, processing mixing or other use of products.

Through Type II strategies, States seek to reduce transaction costs and uncertainty that inhibit cooperation. In contrast with Type I strategies, the focus is on reducing the risk of unintentional noncompliance due to capacity problems, ambiguity regarding what constitutes compliance, and changing circumstances. Nonetheless, with respect to reducing transaction costs, there is some overlap with the positive incentives component of Type I strategies. The impetus behind capacity-building programs may be facilitative and premised on the managerial belief that States have a propensity to comply, and at the same time, such programs affect incentives, especially when financial, technical, or other support is contingent upon compliance.

With respect to reducing uncertainty, there may seem to be an overlap with the informational component of Type I strategies; however, Type II strategies should be viewed more broadly than compliance monitoring (in the assurance and verification sense). For example, States often create organizations, such as a secretariat, to coordinate information gathering, information sharing, communications among parties, and other activities that facilitate cooperation and reduce uncertainties associated with problem identification, possible solutions, payoffs, time horizons, and other variables that affect States' willingness and capacity to participate in negotiations in the first instance, to comply with commitments undertaken in a negotiated agreement, and to continue to cooperate more broadly (beyond express commitments) as the regime evolves. Institutionalizing this strategy may lessen the risk of sudden dynamic change triggered by exogenous events. Depending on the issue area, it may facilitate renegotiation in a new bargaining stage or the operation of institutions implementing a Type III strategy (discussed below).

Consider, for example, the role of an adjudicative (or legalistic) dispute settlement institution. In addition to facilitating a payoff-altering, incentive-based strategy by providing a basis for imposing sanctions, an adjudicative dispute settlement institution may reduce uncertainty regarding the interpretation, scope, and applicability of various obligations and exceptions within a treaty as the regime evolves. To the extent that the line between compliance and noncompliance itself is blurry, adjudicative institutions (and other institutions performing the same function) identify problem areas, provide useful guidance, and reduce uncertainty for future iterations. As seen below, this appears to be the primary function of the legalistic, adjudicative approach taken in the multilateral trading regime. Furthermore, there may be an important value associated with the expression of evolving legal norms through such institutions.

Finally, States may adopt Type III strategies by incorporating institutional mechanisms for dynamically adjusting commitment levels, payoff structures, or other institutional features (Brown Weiss, 2000b; Chayes and Chayes, 1995). Once commitments are undertaken and an expected payoff structure is in place, States may find it necessary to do more than reward compliance or punish noncompliance. The underlying commitments may need to be relaxed, made more stringent, or changed completely depending on the circumstances. Particularly where the number or identity of participants is important or where the expected benefits of cooperation in future rounds outweigh comparable benefits in a current round, parties may prefer to forgive a party's noncompliance and readjust the party's commitments in a manner that improves

the likelihood that it will comply in the future. Furthermore, as Alan Sykes has argued persuasively with respect to the "escape clause" in GATT, the inclusion of such institutions in an agreement may facilitate bargaining over substantive commitments in the formation stage.

Notably, such mechanisms are not always responsive to concerns over opportunism. States may create institutions to facilitate the adjustment of commitments (or other aspects of the agreement) for the purpose of attaining collective goals without having to reform a new agreement (Chayes and Chayes, 1995). Such adjustment may be accomplished through structured bargaining conventions, as in the case of ratcheting down tariffs in the GATT/WTO regime, or through institutional mechanisms delegated authority to adjust commitments in a less formal manner, as in the case of ratcheting up production and consumption limitations for ozone-depleting substances (ODSs). Interestingly, as discussed below, both the GATT/WTO system and the Ozone regime work toward their respective goals of liberalized trade and significantly limited consumption of ODSs by successive ratcheting of commitments made possible through institutional mechanisms.[36] In both cases, the parties anticipated dynamic change and chose to participate in an evolving game rather than a series of iterated games.

Furthermore, beyond adjusting commitment levels, States may wish to modify existing institutions or create new ones as necessary to respond to unforeseen developments. Of course, States may hesitate to delegate the power to make such "adjustments" as a general matter, but they also may hesitate when contemplating the prospect of repeated multilateral negotiations over institutional details, particularly in an issue area where States expect a significant degree of dynamic change and the expected bargaining costs are high.

International law and international relations scholars have suggested two general approaches to designing compliance systems: an enforcement-oriented approach based essentially on the threat and/or use of sanctions as a means of deterring noncompliance, and a management-oriented approach based essentially on reducing ambiguity regarding obligations themselves, creating positive incentives to comply prior to an incidence of noncompliance and adapting treaties to changing conditions (Chayes and Chayes, 1995). The enforcement-oriented approach is consciously focused on the problem of opportunistic defection and the solution of altering States' individual incentives and aligning them with the mutually desired cooperative outcome. It relies on the imposition of "hard" sanctions (such as trade or other economic sanctions) in the event of noncompliance. The WTO dispute settlement mechanism is often considered a good example of an institution embodying this approach, although the analysis below suggests that the institution primarily serves Type II and Type III compliance functions rather than Type I functions.[37] The management-oriented approach de-emphasizes the problem of opportunistic defection and largely rejects the use of "hard" sanctions and other forms of enforcement (Chayes and Chayes, 1995). The management-oriented approach tends to rely primarily on positive incentives, "soft" negative sanctions (such as publication of noncompliance reports and suspension of privileges), commitment-adjustment, and other mechanisms aimed at encouraging continued participation, building capacity, increasing public awareness and support. International environmental regimes are often considered good examples of this approach. The framework set forth in this part provides a means for integrating both approaches. As this part has necessarily been

rather abstract in setting forth a general theory, the next part applies the theory and demonstrates its utility for evaluating compliance system design.

3 Applying the Analytical Framework to International Trade and International Environmental Regimes

This part briefly applies the analytic framework set forth above to two regimes: the GATT/WTO regime within international trade law, and the "Ozone regime" within international environmental law. For brevity's sake, more attention has been given to applying the analytical framework to the GATT/WTO regime than the Ozone regime; a more detailed analysis of various international environmental regimes, which involve a wide range of different problems and institutions, and a comparison of various international environmental regimes with international trading regimes are left for another day. Essentially, the discussion of the Ozone regime provides a single point of comparison with the GATT/WTO regime as well as a point of entry into the international environmental field. Further analysis would need to broaden the focus to include other international environmental regimes. Even the international trade discussion is GATT-focused, a bit abbreviated, and could be expanded significantly. The purposes here are to (1) illustrate the applicability of the game theoretic framework for interdisciplinary analysis; and (2) shed light on the compliance strategies involved in these regimes.

3.1 International Trade Law and the GATT/WTO Regime

3.1.1 The Underlying Interdependency Problem and the Need for International Cooperation

Since World War II, the level of economic interdependence among nations has increased dramatically and national economies have become increasingly dependent upon international trade (Jackson, Davey, and Sykes 2002, pp. 6–7). International trade is understood to be a positive, welfare-enhancing activity that benefits both importing and exporting States. There are extensive theoretical and empirical economic studies supporting this general proposition. Liberal economists argue that assuming competitive markets, the absence of production or consumption externalities, and the absence of economies of scale, *unrestricted* international trade maximizes global welfare for a fixed quantity of resources and technology due to efficient specialization (i.e., the law of comparative advantage) (Yarbrough and Yarbrough, 1992, pp. 7–13). Furthermore, even when the aforementioned assumptions are relaxed, the argument goes, the case for unrestricted international trade generally remains quite strong because of the dynamicism of markets. As natural barriers to international trade become less formidable, for example, when technological advances lead to reductions in transportation and/or communication costs or when cultural differences are overcome through increased social interdependency and relationship building, foreign entry into noncompetitive markets may challenge incumbents and lead to the development of competitive markets over

time. Moreover, the possibility of such a dynamic serves as an incentive to invest in technologies and social relations that facilitate international trade.

Essentially, the argument in favor of international trade can be simplified as follows: (1) competitive markets are generally the most efficient way to provide private goods to consumers; (2) international trade simply involves competition between domestic and foreign firms in domestic markets; (3) government-imposed restrictions on international trade – whether tariffs, subsidies, quotas, or regulatory measures – artificially raise the costs of foreign firms, reduce the scope of entry and reduce competition; and (4) removing existing restrictions on international trade ("trade liberalization") makes markets more competitive, is generally efficient, and is thus economically desirable.[38]

It is well understood that international cooperation on trade liberalization is needed because of the conflict between efficiency gains associated with minimizing artificial, government-made barriers to foreign competition ("liberalization") and the political gains associated with protecting domestic industry from foreign competition ("protectionism"). Ignoring the transaction and adjustment costs of trade liberalization, it can be seen as a positive sum game. Classic economic theory suggests that States acting in their own individual self-interest would unilaterally liberalize trade (i.e., adopt a cooperative strategy regardless of what others do) and happen to make the whole world better off while doing so. The primary explanation for why this does not occur[39] and why domestically constructed barriers to trade persist is typically associated with the distributional consequences of liberalizing trade. Liberalizing trade provides dispersed benefits, such as lower prices to consumers, at concentrated costs to domestic industries that are "hurt" by competition. In essence, the "political economy of trade" biases domestic policy away from trade liberalization and toward protectionism because politicians may maximize the sum of social welfare and campaign contributions (and other personal payoffs), rather than simply maximize welfare.[40]

Economists have also pointed out that, in certain special cases, strategic trade policy may encourage protectionist measures as a means to obtain/retain market dominance in imperfectly competitive, winner-take-all markets. Consider, for example, the simple illustration presented by Paul R. Krugman in his article *Is Free Trade Passe?* (Krugman, 1987). Krugman explains how in a winner-takes-all game where the stakes are high, States may find it in their best interest to subsidize (or otherwise provide a regulatory benefit to) a domestic producer with a chance at the stakes. Krugman uses the example of Boeing and Airbus and suggests that the hypothetical payoff structure absent government intervention would be (−5, −5) if both companies produce a particular plane, (100, 0) if one company produces and the other does not (the producer gets the entire surplus), and (0, 0) if neither company produces.

Hypothetical Payoff Matrix without Government Intervention

		Airbus	
		Produce	Not produce
Boeing	Produce	−5, −5	100, 0
	Not produce	0, 100	0, 0

There is not a unique solution to this game; if one company was able to decide first, it would choose to produce and then the other would not. Absent a head start, it is essentially a race between the companies to commit to producing in order to discourage the other company from doing so. Krugman explains through his example how a subsidy from Europe to Airbus of 10 would adjust the payoff matrix, give Airbus a dominant strategy to produce and thereby deter Boeing from producing, provided, of course, that Boeing is made aware of the altered payoff structure.[41]

Hypothetical Payoff Matrix with Government Intervention

		Airbus	
		Produce	Not produce
Boeing	Produce	−5, 5	100, 0
	Not produce	0, 110	0, 0

Thus, for a cost of 10, Europe may help secure the surplus for domestic industry. Of course, in reality, both Europe and the United States are capable of and have the incentive to subsidize their respective producer. Action by Europe is likely to spur retaliation on the part of the United States. Much like the race to first commit that the companies would find themselves in the absence of a head start, the States similarly may find themselves in a race to subsidize or even a trade war.[42] Such a dynamic is particularly troublesome economically where the game is iterated and the companies (and host States via protectionist measures) are competing for a dominant market position. The importance of this simple, strategic trade policy example is that it illustrates first a situation in which protectionist policy may be in the national interest (by "securing" a surplus for domestic industry, with attendant jobs and tax revenue, at a "small" cost), and second additional impetus for international coordination in order to avoid wasteful races to subsidize and trade wars.[43]

A third reason why States do not unilaterally tear down all domestic barriers to international trade is that some domestic measures that act as barriers, in the sense that the measures impose "artificial" (government-induced) costs on foreign firms, are directed at legitimate domestic ends. As noted above, it is well understood that domestic "intervention" into the market is (economically and socially) justified where positive or negative externalities (or other forms of market failure) are prevalent.[44] Basically, where markets fail (e.g., to produce public goods or prevent overconsumption of commons resources), *the government* is legitimately expected to step in and regulate, tax, stimulate, subsidize, or otherwise promote the public welfare and ensure that it is not sacrificed at the expense of private welfare.[45] The problem in the international trade context is that on one hand, *governments* take a wide range of actions aimed at achieving legitimate regulatory objectives rather than protectionism, even though the measures may have protectionist effects by giving competitive advantages to domestic firms, and on the other hand, governments may mask protectionist measures under the guise of legitimate regulatory objectives for political or strategic reasons. Perhaps the most difficult situation that arises along these lines in the international trade context is where a domestic regulation addresses a valid nontrade concern and is applied in the

same manner on domestic and foreign firms but imposes a significant burden on foreign firms that is not felt by domestic firms (e.g., where domestic firms are technologically equipped to meet the burden without any additional incremental cost while foreign firms are not). In the end, the "legitimacy" of domestic measures is difficult to define, much less measure, because it hinges on the motives and objectives of sovereign States.

Thus, in framing the trade liberalization game and delineating the underlying problem and potential solutions, it is important to recognize the difficulty that arises in distinguishing between legitimate and illegitimate (intentionally protectionist) restrictions. It is also important to note that the existence of this difficulty is one further reason for international cooperation on trade liberalization.[46] Unilateralism leaves it to States to individually assess the "legitimacy" of other States' domestic programs and decide whether to retaliate through similar measures or other means (e.g., unilateral trade sanction). International cooperation, particularly through a legal and institutional framework, can provide both a set of principles and a forum for assessing legitimacy. Such cooperation may take place within the trade regime, other international regimes, or both. The GATT/WTO regime provides what might be considered, depending on one's perspective, default principles and a (default) forum for assessing the legitimacy of various trade-inhibiting measures. International environmental regimes provide another (arguably more direct) mechanism for establishing the legitimacy of nontrade objectives on a multilateral basis.

Although envisioning the "trade liberalization game" through classic economic theory tends to invoke images of a Harmony game where each State has a dominant strategy to cooperate by liberalizing trade with every other State, the political benefits of "defecting" by maintaining or creating barriers to protect domestic industries may transform the game into a Prisoners' Dilemma (or another game, e.g., Suasion, that does not lead to joint cooperation without coordination among players).[47] In the first instance, trade liberalization was seen as a two-player game because trade barriers, such as tariffs and quotas, tended to be bilateral when actually applied (i.e., the barrier directly affects a foreign firm). Of course, in addressing the problem of trade liberalization, States have entered into bilateral, regional, and multilateral agreements, linking and integrating two-player games. The most-favoured nation principle of the GATT/WTO regime has effectively made international trade a multiplayer game for its members. The payoff structures across sets of countries vary considerably and are not amenable to simple generalization. It is worth noting, however, that beyond the internal (or domestic) distributional consequences of liberalizing trade, the magnitude of payoffs from cooperating to liberalize international trade varies considerably based on differences in, *inter alia*, the size and economic development of potential cooperators, and the existing level and types of restrictions (e.g., gaining access to US markets on better terms may be considerably more valuable than gaining equally liberal access to the markets of a smaller or less-developed country).[48] Moreover, as evidenced by the formation of various bilateral, trilateral, regional, and other coalitions, the identity of participants in a particular trade liberalization game is an important factor because of the complex economic interdependencies across different sets of countries. (The economic interdependencies themselves arise not simply from government policies but also from, *inter alia*, social and cultural interdependencies. Such complexities are, however, beyond the scope of this chapter.) Finally, the game is iterated and

expected to be played indefinitely (cessation of trade relations altogether is not generally foreseeable).

3.1.2 THE GATT/WTO REGIME

The most important institutional arrangement of the post-World War II trading system is the GATT/WTO. The GATT/WTO regime is a multilateral regime that originated with the General Agreement on Tariffs and Trade ("GATT 1947"),[49] evolved through many trade rounds, and culminated with the creation of the WTO.[50] Although the evolution of the regime is quite complex and would be an interesting topic to analyze more thoroughly through the lens of the foregoing analytical framework, the treatment here is brief. The analysis below focuses primarily on the manner in which the regime has evolved with respect to substantive obligations/principles,[51] commitments, and compliance institutions. Notably, the GATT 1947 has been superseded by the GATT 1994 and is no longer in effect. Nonetheless, because the focus of this chapter is the evolving GATT/WTO regime,[52] the chapter generally uses the GATT to refer to the operative agreement at a given time and distinguishes between the two agreements only where necessary. As stated in the GATT 1947 preamble:

> Recognizing that their relations in the field of trade and economic endeavour should be conducted with a view to raising standards of living, ensuring full employment and a large and steadily growing volume of real income and effective demand, developing the full use of the resources of the world and expanding the production and exchange of goods.
>
> Being desirous of contributing to these objectives by entering into reciprocal and mutually advantageous arrangements directed to the substantial reduction of tariffs and other barriers to trade and to the elimination of discriminatory treatment in international commerce

The parties agreed to a multilateral cooperative effort generally aimed at liberalizing trade through a series of tariff reductions. Four principles set forth in the GATT 1947 constituted the most important (although by no means exclusive) substantive obligations between parties. The core principle of the GATT system, originally set forth in Article II of GATT 1947, is that parties will not charge a tariff for a particular product above the level agreed to and set forth in the tariff schedules. To be clear, tariffs themselves are legal, and parties may charge whatever tariff they like for products not listed in the schedules. But once a "tariff binding" has been negotiated for a particular product, parties are bound to charge at or below that level.[53] This simple, legal obligation provided the basis for gradually "ratcheting" down artificial barriers to trade through negotiations and reciprocal commitments, and a relatively easy obligation to police for compliance purposes. First, the obligation itself is rather clear and not open for renegotiation[54] or broad interpretation, and the tariff schedules themselves provided a precise, detailed, quantitative baseline for gauging compliance. Second, with respect to detecting noncompliance, "governments and international organizations collect mountains of trade statistics; tariff lists are publicly available; and the possibility of tracing goods through customs exists" (Yarbrough and Yarbrough, 1992).

A second principle, originally set forth in Article I of the GATT 1947, is *General Most-Favoured-Nation Treatment* (MFN). This nondiscrimination principle, which "has long been a cornerstone of international trade law,"[55] requires that each party grant

to every other party treatment with respect to any imports and exports no less favorable as it grants to any other nation. Thus, MFN ensures that any trading advantages given by one party to another nation, whether a party or not, will also be given to each party. In essence, this principle "multilateralizes" the trade liberalization game (at least with respect to obligations and negotiations) and provides a significant incentive for nonmembers to join.

A third principle, originally set forth in Article III of GATT 1947, is *National Treatment on Internal Taxation and Regulation*. This nondiscrimination principle requires that with respect to internal taxation and regulatory measures, each party treat imports from other parties no less favorably than domestically produced goods. Finally, a fourth principle, originally set forth in Article XI of GATT 1947, is that "No prohibitions or restrictions other than duties, taxes or other charges, whether made effective through quotas, import or export licenses or other measures, shall be instituted or maintained by any contracting party on the importation of any product of the territory of any other contracting party or on the exportation or sale for export of any product destined for the territory of any other contracting party."[56] In essence, this principle, the national treatment principle, and a number of other GATT clauses limiting government actions affecting imports and exports are aimed at restricting domestic measures that could be substituted for tariffs as protectionist tools and at protecting the potential benefits of binding and then reducing tariffs. While the GATT treaty system is extremely complex, incorporating hundreds of agreements and protocols, incredibly detailed schedules of tariff bindings and many important exceptions (e.g., Article XIX: escape clause allowing temporary use of import restraints when imports cause serious injury to domestic industry; Article XX: health and safety regulation; Article XXI: national security; Article XXIV: allowing customs unions and free trade areas; Article XXV: waiver authority), these four principles create the basic framework for multilateral trade liberalization through reciprocal tariff concessions. By most accounts, the framework has been extremely successful in reducing tariffs on most goods.

The GATT 1947 itself did not set forth a comprehensive, integrated compliance system. It imposed reporting and publication obligations on parties: for example, Article XVI states "[i]f any contracting party grants or maintains any subsidy, including any form of income or price support, which operates directly or indirectly to increase exports of any product from, or to reduce imports of any product into, its territory, it shall notify the contracting parties in writing of the extent and nature of the subsidization, of the estimated effect of the subsidization on the quantity of the affected product or products imported into or exported from its territory and of the circumstances making the subsidization necessary;" and Article X states that virtually all "[l]aws, regulations, judicial decisions and administrative rulings of general application" pertaining to imports and exports "be published promptly in such a manner as to enable governments and traders to become acquainted with them." As noted above, governments and international organizations collect "mountains" of trade information; most importantly, in the end, the experience of importers and exporters naturally exposes barriers to trade.

The GATT system generally relied on the parties to initiate and participate in dispute settlement procedures generally beginning with consultations and shifting into arbitration if necessary. The major dispute settlement provision in the GATT 1947 was Article XXIII, *Nullification or Impairment*, although other compliance-oriented

clauses permeate the treaty text – for example, numerous clauses in the GATT 1947 require consultation between parties,[57] permit compensatory withdrawal or suspension of concessions, and delineate exceptions. Article XXIII generally requires parties first to consult in the event that a party believes a "benefit accruing to it under [the] Agreement is directly or indirectly nullified or impaired or the attainment of any objective of the Agreement is being impeded as the result of" another party's (1) breach of its obligations under the agreement or (2) application of some measure, or as the result of any other situation.[58] According to the text of Article XXIII, in the event that consultations are unsuccessful, "the matter may be referred to the Contracting Parties." Further, "[i]f the Contracting Parties consider that the circumstances are serious enough . . ., they may authorize a contracting party or parties to suspend the application to any other contracting party or parties of such concessions or other obligations under this Agreement as they determine to be appropriate in the circumstances." Article XXIII does not set forth comprehensive procedures for determining when "circumstances are serious enough to justify such action," or what actions would "be appropriate in the circumstances." Rather, the dispute settlement procedures (in contrast with the substantive rules) have evolved over time into a detailed practice where a panel of arbitrators preside over the dispute and issue a report recommended for adoption by the parties. The procedures used were codified in the 1979 Tokyo Round Understanding and subsequently modified through various understandings, interpretations, and declarations.[59]

Two important aspects of the GATT dispute settlement process significantly limited its effectiveness: first, panel reports (i.e., final decisions of an arbitration panel) were not adopted by the parties and thus made "binding" on the participants unless the parties reached a consensus; the obvious, debilitating result of this rule was that the losing party could block adoption of the report. This veto power restricted the dispute settlement institution's deterrent effect because a party contemplating noncompliance as a strategic decision would know *ex ante* that it could rely on the veto power to avoid a formal finding of noncompliance.[60]

Second, even if a report established, for example, that one party's breach had caused another party's nullification or impairment and the report is adopted by the parties, the effect of the report was limited.[61] Specifically, the report would establish the existence of a violation (or situation leading to nullification or impairment) and would require the offending party to bring itself into compliance. In the event that the offending party failed to do so and the parties' failed to negotiate a compromise solution, the parties could authorize the injured State to utilize trade measures to accomplish adequate compensation/retaliation[62] on a prospective basis to offset future harms from continued noncompliance;[63] *an injured party was not authorized to seek compensation for past harms or to implement punitive measures*.[64] Arguably, this limitation severely constrained the effectiveness of the institution from an enforcement- or deterrence-oriented compliance perspective because it could not adequately alter the expected payoffs for noncompliance in the first place, in other words, from an *ex ante* perspective where a nation was deciding whether to defect on the basis of the expected payoffs (e.g., to intentionally not comply by implementing a protectionist government measure in conflict with a particular GATT obligation). The expected benefits of noncompliance depend primarily upon the length of time (or number of iterations) that a State can maintain noncompliance without surrendering payoffs thereby obtained if caught.

Because the GATT system first provided a noncompliant State with the opportunity to bring itself into compliance with no penalty, except collateral consequences such as a loss of goodwill or reputation, and second would only withhold benefits on a prospectively compensatory basis,[65] States intending to act opportunistically would weigh the expected benefits only against collateral costs that could occur upon detection.

On its face, the compliance strategy undergirding the GATT dispute settlement system appears to have been a Type I strategy aimed at the payoff structure, in the sense that it relied on an adjudicative dispute settlement process that authorizes an injured party to prospectively withhold benefits that would otherwise flow to the other party. To the extent that one focuses on the GATT dispute settlement institution's capability to effectuate alterations in the *ex ante* incentives to defect, the institution might appear weak because of the pitfalls described above. Besides the difficulty created by the consensus rule, the inability to seek compensation for past harm (much less punitive retribution) severely limited the effectiveness of the institution in deterring intentional noncompliance through alterations in the *expected* payoff structure.

Perhaps the reason why the GATT 1947 did not expressly authorize retroactively compensatory or punitive measures is that disputes generally did not concern intentional noncompliance, which would warrant such measures from the perspective of deterrence, but rather concerned either real disputes over the scope and interpretation of the rules or capacity problems, either of which would not warrant such measures from the perspective of deterrence because noncompliance would not be the result of a party's lack of intent to comply. Of course, once an adopted panel report established the existence of a violation, parties could work out a cooperative solution if a capacity problem existed; thereafter, continued noncompliance would likely be a problem of intent and thus properly sanctioned on either a compensatory or even punitive basis. In the end, therefore, it appears that the institution was not directed primarily at deterring intentional noncompliance, except on an ongoing basis after the noncompliance had been identified and adjudicated.[66]

It seems more appropriate to analyze the effectiveness of the institution with respect to Type II and Type III compliance strategies, which respectively concern developing institutional mechanisms for reducing transaction costs and uncertainties that inhibit cooperation and for dynamically altering commitments. The consultations and dispute settlement process may alleviate uncertainty with respect to the interpretation and proper application of the rules and the scope of the parties' obligations. To the extent that disputes between parties did not originate from intentional noncompliance but rather from an actual dispute over the rules or obligations, or from a capacity problem, the process would provide significant benefits to the disputants as well as the GATT regime as a whole. The fact that the settlement rate prior to a panel ruling was high and that most panel reports were eventually adopted by the losing party provides support for this conclusion. With respect to dynamic readjustment of commitments, when an adopted panel report was not implemented by the offending party, the injured party might be authorized to readjust its commitments relative to the other party.

In addition to setting forth the basic principles for liberalizing trade through reciprocal tariff concessions and the dispute settlement process, GATT served as a negotiating forum, most importantly, in a series of trade negotiating rounds. To date, there have

been eight rounds, summarized as follows (Jackson, Davey, and Sykes, 2002, p. 227):

Location/common name	Year(s)	Number of countries participating
1 Geneva, Switzerland	1947	23
2 Annecy, France	1948	33
3 Torquay, England	1950	34
4 Geneva	1956	22
5 Geneva, “Dillon Round”	1960–1	45
6 Geneva, “Kennedy Round”	1964–7	48
7 Geneva, “Tokyo Round”	1973–9	99
8 Geneva, “Uruguay Round”	1986–94	120+

Notably, the first six rounds focused on negotiating tariff reductions; one of the goals of the Kennedy Round was to deal with nontariff barriers, but in the end, the results were tariff reductions (Jackson, Davey, and Sykes, 2002, p. 227). By the seventh round, the Tokyo Round, the parties recognized that nontariff barriers (NTBs) needed to be addressed (along with continued reductions in tariffs) and that the existing broad principles set forth in the GATT 1947 were insufficient. In addition to protocols regarding tariff reductions, the Tokyo Round led to the completion of special, side-agreements or “codes” that essentially operate as independent treaties by binding signatories and creating independent obligations and compliance mechanisms. The side-agreements addressed a wide range of issue areas not directly touched on by the GATT 1947, such as technical barriers to trade and government procurement.

The creation of side-agreements illustrates the dynamicism (or reframing and reforming) of cooperation in an iterated game. After identifying an interdependency problem and crafting a cooperative solution through an international agreement, parties may need to return to the framing stage and reframe particular aspects of the problem that are not adequately addressed under the existing arrangement.[67] The inadequacy of the GATT 1947 (in terms of applying, interpreting, and enforcing existing obligations relating to NTBs) was exposed in part as a result of its success in reducing tariffs and in part as a result of its inability to adapt internally.

> [B]y the 1970s, it became clear that tariffs were not likely to be the main problem for trade liberalization. As tariffs decreased (especially for industrial goods imported to industrialized countries), many special sector interests began to seek other ways to reduce competition from imports, turning to “non-tariff barriers” (NTBs). These are myriad, and many are the types of things for which human ingenuity can perpetually develop new devices. Most such barriers are internal measures and not border measures, and therefore are often most relevant to the national treatment clause of GATT Article III. In this respect, the GATT was essentially forced to address nation-states’ internal economic regulatory measures, or run the risk of becoming almost totally irrelevant to the need for international cooperative mechanisms to resolve thousands of international tensions and problems related to trade, or to keep some movement toward trade liberalization.
>
> ***(Alvarez and Jackson, 2002)***

While Article III applies to NTBs on its face, enforcing that obligation through the dispute settlement mechanism was difficult. While a common law-like evolution of the GATT 1947 principles, particularly Article III, to address NTBs was theoretically possible (even though dispute settlement reports were not formally subject to the doctrine of stare decisis), such an approach did not prove to be practically feasible. Evolution through amendment was even less feasible because an amendment would not be binding on a particular party unless that party accepted the amendment. As Jackson and his colleagues explain, the near impossibility of amending the GATT 1947 was a major "birth defect:" the "delay required by the treaty acceptance process, the shift in bargaining power involved under the amending procedure in the context of a large membership and the fact that even when an amendment is effective it will not apply to countries which do not accept it, are all reasons why the amending procedure had fallen into disuse . . . caus[ing] a certain rigidity and inability to develop rules to accommodate the many new developments in international trade and other economic interdependence subjects" (Jackson, Davey, and Sykes, 2002, pp. 214–15). While designed to supplement the GATT 1947 and address its deficiencies, the side-agreements created a host of other institutional problems. For example, it was not clear whether benefits given in the side-agreements were subject to the MFN provision until the GATT parties adopted a decision suggesting that they were. The GATT 1947's uniform "multilateralization" of the trade liberalization game gave way, in part, to a balkanized system of side-agreements.

In addition to the growing importance of addressing NTBs, a second major development pushed the evolution of the multilateral trading system developed under the GATT 1947 toward today's WTO system. Specifically, as summed up succinctly by Charles Tiefer:

> Initially, the United States accepted the GATT as the chief means of ensuring that these countries reciprocated US openness. However, by 1962, Congress began to reflect a domestic political conviction that other countries, particularly Japan and members of the European Economic Community (EEC), had not honored US rights. In response, it enacted section 252 of the 1962 Act authorizing retaliation against other countries' import restrictions of either illegal or "unreasonable" nature. In the 1974 Act, Congress further followed these domestic political convictions, exacerbated by a sense of the shortcomings of the GATT itself and the Kennedy Round agreement. Specifically, Congress created section 301, [19 U.S.C. § 2411,] allowing measures to be taken against countries that maintained unreasonable barriers to US trade. In 1988, Congress expanded the section 301 system significantly. In response to the persistent foreign infringement on intellectual property rights, the 1988 Act [Omnibus Trade and Competitiveness Act of 1988, Pub. L. No. 100- 418,] added the "Special 301" provision. [19 U.S.C. § 2242.] It provided for identification of countries that infringed on intellectual property rights and imposed unilateral US trade sanctions if those countries did not mend their ways. US sanctions could take the form of suspension of trade agreement benefits or an increase in tariffs or non-tariff barriers against imports from the violating countries. [19 U.S.C. § 2411(c)(1)(A)–(B); see also § 2416(b).] Also, the 1988 Act added the "Super 301" provision, which allowed the USTR to identify foreign country practices that blocked US exports and designate them as priorities. [19 U.S.C. § 2420.] The Super 301 provision [was and] remains the backbone of unilateral US efforts to force trade liberalization on foreign countries.
>
> ***(Tiefer, 2001, pp. 63–4)***

While various scholars debate the relative importance of the role of US hegemony (or its perceived decline) in pushing the multilateral trading system's evolution (a debate not engaged in here), there is little doubt that the increased unilateralism and threat thereof on the part of the United States as a means to "enforce" its trade-related rights was an important factor in getting nations to explicitly reframe the underlying interdependency problem and reform the existing set of obligations (primarily through significant expansion) and compliance institutions.

The evolution of the GATT 1947 institutional framework culminated in the Uruguay Round. At the close of the Uruguay Round, the parties had negotiated a complex package of agreements that resulted in a comprehensive overhaul of the existing GATT regime. Parties were not given the option of picking and choosing among agreements but rather were bound to accept or reject the entire package; this allowed parties to engage in a complicated process of negotiations that reframed, reformed, and integrated games across many issue areas. Among many other things, the parties (1) created the WTO – a formal international organization (2) significantly broadened the scope of the trading system beyond goods to encompass services and intellectual property as well as previously exempted areas such as agriculture and (3) created a more comprehensive compliance system including more effective dispute settlement procedures. Ultimately, the GATT 1947 regime was completely subsumed by the WTO regime. Technically, the GATT 1947 is no longer in effect because it was replaced with GATT 1994, although the substantive obligations are essentially the same. With respect to commitments and compliance, the WTO package of agreements significantly expanded the range of commitments undertaken by States and significantly deepened them as well. The Marrakesh Agreement, which established the WTO, includes four Annexes. The first Annex contains a series of multilateral agreements that bind all WTO members and thus comprised the "package deal" during negotiations. Annex 1 contains, among others, the GATT 1994, the General Agreement on Trade in Services ("GATS");[68] and the Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPS").[69] Annex 2 contains the dispute settlement rules that apply to all members, and Annex 3 establishes the Trade Policy Review Mechanism. Annex 4 contains "plurilateral agreements" that are optional.[70] The analysis here will not focus on the range of commitments undertaken via these agreements[71] but instead on the compliance system.

The WTO regime involves a comprehensive, integrated compliance system. In particular, the dispute settlement system is seen as the most advanced and powerful, yet created under international law. Although, as many have pointed out, it still has its flaws, it represents a significant advancement in the international law field, is the "backbone" of the WTO trading system,[72] and has been touted as the model institution for other developing areas of international law. The WTO dispute settlement system builds on the preexisting GATT dispute settlement system, brings together and formalizes many of the procedures set forth in various interpretations and declarations, and creates a formal dispute settlement organization called the Dispute Settlement Body (DSB).[73] The DSB is formally delegated authority to administer the detailed rules of adjudication contained in the Understanding on Rules and Procedures Governing the Settlement of Disputes ("Dispute Settlement Understanding" or "DSU") (Annex 2).[74] The DSB orchestrates the dispute settlement process from beginning to end – it establishes panels, adopts Panel and Appellate Body reports, and even monitors and

oversees implementation of such reports. Importantly, the WTO dispute settlement system prevents forum shopping among procedures set forth in side-agreements, for example, and acts as the exclusive adjudicative institution for all agreements under the WTO umbrella. In some cases, the adjudication rules set forth in the DSU are modified by provisions in some of the Annex 1 agreements, but the modifications are generally tailored to the particular area covered (e.g., intellectual property).

Besides integrating and formalizing various procedures into a comprehensive set of binding rules and creating the DSB to administer the process,[75] the DSU reflects a number of important changes to the dispute settlement system. First, a losing party can no longer block the adoption of a panel report; instead, a panel report is automatically adopted unless there is a consensus not to do so. A second related reform was the introduction of appellate procedures and the creation of an Appellate Body, consisting of seven members that are appointed to four-year terms. The Appellate Body may only consider issues of law and legal interpretations decided by the panel.[76] As with panel reports, the Appellate Body report is adopted automatically unless there is consensus not to do so. Together, these modifications significantly depoliticize the dispute settlement process and give weight to the panel's and Appellate Body's interpretation and application of the rules and determinations with respect to whether a particular party is complying with its obligations. A third important change is that the DSB monitors whether a losing party implements the relevant recommendations and brings itself into compliance; if the party fails to do so within a reasonable period of time, authority to suspend concessions is automatically granted absent consensus to the contrary and the level of suspension is calculated pursuant to a specified procedure.

The clear preference in the DSU is for withdrawal by the losing party of its offending measure,[77] but the sanction for failing to do so basically remains the same as it was for the GATT: *prospective trade measures intended to offset the prospective harm imposed on the injured party*.[78] Article 22 of the DSU sets forth various rules and procedures for compensation or suspension of concessions, which are "temporary measures available in the event that the recommendations and rulings are not implemented within a reasonable period of time." The DSB expressly states that compensation is voluntary.[79] It further states "[t]he level of the suspension of concessions or other obligations authorized by the DSB shall be equivalent to the level of the nullification or impairment."[80] As with the GATT 1947 Article XXIII language, it not clear whether equivalency includes retroactive effects, although both the text of the DSU and past practice strongly suggest that remedies must be determined on a prospective basis.[81] The focus of the DSU is clearly prospective and aimed at inducing compliance once a violation has been established (Schwartz and Sykes, 2002, p. 12). As Article 22 ¶ 8 states, "suspension of concessions or other obligations shall be temporary and shall only be applied until such time as the measure found to be inconsistent with a covered agreement has been removed, or the Member that must implement recommendations or rulings provides a solution to the nullification or impairment of benefits, or a mutually satisfactory solution is reached." As the arbitrators in the *EC-Bananas* case reasoned, the "*temporary* nature (of suspension of concessions or other obligations) indicates that it is the purpose of countermeasures to *induce compliance*."[82] Furthermore, in that case, the arbitrators assessed the level of proposed suspension of concessions in relation to

the measures taken by the European Communities ("EC") in order to comply with the recommendations and rulings of the DSB rather than the original measures found to be inconsistent with the EC's WTO obligations.[83] Compensation for past harm caused by the original measures was not a relevant consideration in determining whether the proposed countermeasures were appropriate. Importantly, the types of concessions that may be suspended may include obligations or concessions under any of the WTO multilateral agreements, unless there is a provision precluding such suspension.[84] This development arguably "levels the playing field" so to speak because developing nations that would have had little bargaining power under the GATT system may suspend obligations that matter to developed countries such as those set forth in TRIPS.[85]

In the end, WTO dispute settlement appears to maintain the same orientation as GATT dispute settlement and appears to implement the Type II and Type III strategies in the same manner as the GATT, see *supra*, although much more effectively and credibly because of the detailed, integrated framework of rules and the depoliticization of the process (e.g., by reversing the consensus rule, creating the DSB, and providing for appellate review of legal determination). The rule-oriented, adjudicative process established in the DSU and the creation of an appellate review process substantially improve the quality, consistency, predictability, and overall utility of the dispute settlement process as a means for reducing uncertainty regarding the scope, interpretation, applicability, and complex interrelationship among the sets of rules in the various agreements.

As with GATT dispute settlement, the prospective orientation of remedies available through WTO dispute settlement limits the institution's effectiveness as a deterrent to intentional noncompliance from an *ex ante* perspective. The institution does serve a Type I function in the sense that once the process has been invoked and noncompliance has been established, the possibility of suspended concessions may induce compliance and may act as an effective deterrent to intentional noncompliance of the same kind between the same Members in the future. Furthermore, as a result of the improvements made in the evolution from GATT dispute settlement to WTO dispute settlement, the expected duration of undetected noncompliance may shorten (in part because the institutions are more efficient, and in part because the scope of the trading rules and obligations becomes more precise and certain over time), which reduces the expected benefits of intentional noncompliance.

Yet, the fundamental limitations of the dispute settlement system, from an enforcement-oriented perspective – an opportunity to return to compliance without penalty and the prospective nature of available remedies (explored above with respect to the GATT regime) – strongly suggest that the institution is not directed at deterring intentional noncompliance, and therefore, that the institution does not implement a collective strategy to alter the expected payoffs in a manner that makes cooperation more attractive *ex ante* than defection. Instead, it appears to implement a collective strategy to facilitate internal evolution of the existing (multilateral) game, and thereby avoid collapsing into sequential games (whether bargaining games or trade wars on a bilateral or multilateral basis).

In addition to dispute settlement, the parties also created another Type II compliance institution, the Trade Policy Review Mechanism (Annex 3).[86] As set forth in its

statement of objectives,

> [t]he purpose of the Trade Policy Review Mechanism ("TPRM") is to contribute to improved adherence by all Members to rules, disciplines and commitments made under the Multilateral Trade Agreements and, where applicable, the Plurilateral Trade Agreements, and hence to the smoother functioning of the multilateral trading system, by achieving greater transparency in, and understanding of, the trade policies and practices of Members. Accordingly, the review mechanism enables the regular collective appreciation and evaluation of the full range of individual Members' trade policies and practices and their impact on the functioning of the multilateral trading system. It is not, however, intended to serve as a basis for the enforcement of specific obligations under the Agreements or for dispute settlement procedures, or to impose new policy commitments on Members.

The TPRM establishes the Trade Policy Review Body (referred to herein as the TPRB) to periodically review the trade policies and practices of all Members; the frequency of the reviews is based on the impact of a particular country on the trading system. Each Member is obligated to submit a report to the TPRB on a regular basis as well. Specifically, the TPRM provides that:

> In order to achieve the fullest possible degree of transparency, each Member shall report regularly to the TPRB. Full reports shall describe the trade policies and practices pursued by the Member or Members concerned, based on an agreed format to be decided upon by the TPRB.... Between reviews, Members shall provide brief reports when there are any significant changes in their trade policies; an annual update of statistical information will be provided according to the agreed format. Particular account shall be taken of difficulties presented to least-developed country Members in compiling their reports. The Secretariat shall make available technical assistance on request to developing country Members, and in particular to the least-developed country Members. Information contained in reports should to the greatest extent possible be coordinated with notifications made under provisions of the Multilateral Trade Agreements and, where applicable, the Plurilateral Trade Agreements.

While not oriented towards dispute settlement, the TPRM serves an important Type II compliance function in that it collects and makes publicly available information regarding trade policies of the parties and provides an important feedback mechanism for evaluating and understanding the payoff structure.

The WTO has recently begun to focus on the capacity problems that developing countries may face and has been developing capacity-building and technical assistance programs to enable developing countries that lack the resources to fully participate in the WTO. In November 2001, at the Fourth Ministerial Conference in Doha, Qatar, a broad mandate was declared concerning negotiations on, *inter alia*, numerous implementation issues, capacity-building and technical assistance programs, technology transfer, and special and differential treatment provisions. The Doha Development Agenda stresses that WTO trade assistance must be viewed as part of an overall development and poverty reduction strategy. The WTO's recent 2003 Technical Assistance/Capacity Building Plan, for example, focuses on the complementary goals of technical assistance and capacity-building for effective participation in negotiations, implementation, and trade integration. To that end, the WTO Secretariat will be engaged in the coordination of a number of technical assistance and capacity-building activities (such as internships, trade policy courses, and sustainable capacity-building), which coherently link national activities with regional activities. Least-developed countries, which have

the most urgent and acute trade development needs, will be given priority focus. As a precursor to technical assistance, diagnostic studies are being prepared to identify structural weaknesses and constraints impeding the countries development into the multilateral trading system. In the near future, capacity-building and technical assistance institutions will likely become an important part of the WTO compliance framework. The emergence of these institutions can be explained as a response to the bubbling threat of destabilization posed by exogenous events, such as health crises, and by the capacity problems facing many developing countries.

Looking at the GATT/WTO regime's evolution through the lens of the game theoretic framework set forth in this chapter helps us to focus on the aspects of State decision-making and behavior that are affected by compliance institutions. While international trade law has evolved into a relatively strong version of public international law, the strength of the current WTO regime does not appear to derive from strict enforcement-oriented institutions aimed at deterring intentional noncompliance through the threat of sanctions. This view is contrary to the view of most trade experts who believe that the dispute settlement process is, in fact, enforcement-oriented and backed by hard-edged sanctions. The analysis in this section has shown that, despite its adjudicative, rule-based orientation, the WTO dispute settlement institution appears to be management-oriented and facilitative, rather than enforcement-oriented, because it primarily implements Type II and Type III compliance strategies and implements Type I strategies only on a limited prospective basis. The underlying purpose of the WTO compliance system is to maintain regime stability.

The next section applies the analytical framework to the Ozone regime, which addresses a fundamentally different interdependency problem than discussed above with respect to the international trading regime. While the Ozone regime also appears to be primarily management-oriented and facilitative, it does not rely heavily on dispute settlement but rather utilizes a host of compliance institutions to implement all three compliance strategies.

3.2 International Environmental Law and the Ozone Regime

3.2.1 The Underlying Interdependency Problem and the Need for International Cooperation

Today, the interdependency problem motivating international cooperation to reduce the emission of manmade, ozone-depleting substances (ODSs) is well understood. Accordingly, a detailed analysis is not necessary here. To summarize, ozone in the stratosphere (the "Ozone layer") shields all of us living on Earth from ultraviolet radiation emitted from the sun. There are significant health risks associated with increased exposure to ultraviolet radiation. When certain manmade chemicals are emitted and make their way into the stratosphere, ultraviolet radiation breaks down the chemicals, releasing, among other things, chlorine or bromine atoms, which act as catalysts in destroying ozone. Chlorofluorocarbons (CFCs) and other ODSs, such as halons, methyl bromide, carbon tetrachloride, methyl chloroform, and hydro-chlorofluorocarbons (HCFCs), have been used in numerous consumer and industrial applications.

International cooperation is necessary to address the problem because doing so involves a tradeoff between the risk of long-term environmental harm and attendant health effects and the short-term economic costs of developing and shifting to alternative chemicals. The Ozone layer can be analyzed as either a global public good or common pool resource depending on the manner in which the problem is framed.[87] On the one hand, an intact Ozone layer performs a service by shielding everyone on Earth from ultraviolet light. Because this service is both nonrivalrously consumed and nonexcludable, it can be characterized in economic terms as a classic public good (Oakland, 1987, pp. 485–6; Cooter and Ulen, 1997, pp. 40–1). On the other hand, when framed in terms of its capacity to act as a sink for ozone-depleting chemicals, the Ozone layer remains nonexcludable while its capacity is rivalrously consumed (Downie, 1999, pp. 97–121). Under either view, sustaining the Ozone layer may be analyzed as a multi-player Prisoners' Dilemma.[88] Robust international cooperation is especially necessary because the Ozone layer may be depleted by ozone-depleting emissions regardless of where on the Earth the emissions come from. This fact is particularly important because it means that broad international cooperation is essential to protecting the Ozone layer. A State (or group of States) that opts not to participate, either by not committing to anything or by not complying with commitments undertaken, may undermine the efforts of cooperating States by simply emitting ODSs in a quantity sufficient to deplete the Ozone layer. This characteristic of the interdependency problem provides some States with an opportunity to hold out and demand concessions from others.

3.2.2 The Ozone regime

The Ozone regime is a popular case study for international environmental cooperation because it has been quite successful in a number of respects (Brown Weiss, 1998, pp. 135–57). First, and most importantly, participants have significantly reduced (and in some cases completely banned) the production and consumption of various ODSs. Second, participants have collaborated on identifying additional chemicals that pose a risk to the Ozone layer. Third, newly identified ODSs have been added to the list of chemicals targeted for "ratcheting" down production and consumption quotas. Fourth, in response to institutionalized incentives, countries that were not original participants have gradually joined the regime. Of course, there are additional successes and some shortcomings. The analysis below focuses primarily on the manner in which the regime has evolved with respect to substantive obligations/principles, commitments, and compliance institutions. Given the fact that the Ozone regime and its compliance institutions have been extensively studied elsewhere[89] and that the Ozone regime provides only an introduction to the wide range of international environmental regimes, the treatment here is somewhat brief in comparison with the previous treatment of the multilateral trading regime.

In 1974, a series of scientific papers suggested the CFCs could destroy stratospheric ozone (Molina and Rowland, 1994, pp. 810–12; Stolarski and Cicerone, 1974, pp. 1610–15; Cicerone, et al., 1974, pp. 1165–8). In the mid-to-late 1970s, scientific and public policy debates ensued and various nations, including the United States, established domestic controls on the use of CFCs in aerosol sprays. However, the European Community, Japan, the Soviet countries, and various large developing countries would

not do so, in part because of uncertainties regarding various aspects of the environmental problem. Accordingly, in the late 1970s, United Nations Environment Program (UNEP) was charged with promoting research to develop a better understanding of the problem and possible solutions. In 1981, UNEP established an Ad Hoc Working Group of Legal and Technical Experts to draft a global framework convention, and in 1985, States adopted the Vienna Convention for the Protection of the Ozone Layer.[90] The Vienna Convention was signed by 20 countries plus the European Community, and it entered into force on September 22, 1988.[91]

Through this framework convention, parties did not undertake specific binding controls on their production or consumption of ODSs but instead agreed to cooperate on (1) conducting research and scientific assessment regarding the problem; (2) information exchange; and (3) adopting "appropriate measures" to deal with the problem. Although quite general in nature, these obligations served as the foundation for significant collaborative efforts to better understand the problem and evaluate appropriate cooperative solutions. This was a crucial first step toward international cooperation because it broadly framed the relevant issues at a time where the underlying problem and potential solutions were relatively uncertain and States were understandably hesitant to undertake specific commitments. In essence, the broad framework convention allowed parties to coordinate their behavior while continuing to reframe the problem and potential solutions, as nonbinding agreements frequently do in international environmental regimes.[92]

Importantly, parties fully expected that their commitments would evolve over time. At the time the Vienna Convention was being negotiated, the parties contemplated that a protocol with more specific obligations would be negotiated on the side. Due to a lack of consensus on the substance of such a protocol at the time the Vienna Convention was finally adopted, however, the parties left the drafting of a protocol to future development. As many observers have noted, the framework-protocol approach taken in the Ozone regime was particularly important because getting as many countries as possible on board was crucial to any long-term effort to address the environmental problem.

During the late 1980s and early 1990s, reduced uncertainty regarding the problem (particularly its magnitude)[93] and potential solutions, technological developments, and better alignment of developed countries' preferences combined to push States to agree to undertake significant, quantitative commitments to reduce specific ODSs. In 1987, the Montreal Protocol on Substances that Deplete the Ozone Layer was concluded,[94] and in early 1989 it entered into force. The Montreal Protocol imposes obligations on States to control the production and consumption of ODSs through specific measures and timetables. The 1987 version of the Protocol required that parties (other than developing countries) (1) freeze their production and consumption of CFCs at 1986 levels; (2) reduce their production and consumption of CFCs first by 20 percent and then by another 30 percent by 1999; and (3) freeze their consumption of halons at 1986 levels. Parties are also required to submit detailed annual reports containing statistical data on various controlled substances to the Secretariat for the Protocol in Nairobi, Kenya. The Secretariat compiles the various reports and makes them available to parties, other Ozone regime institutional bodies (such as expert panels, e.g., see *infra*), and the public for review. The primary substantive commitments are quantitative in nature, basically capping the amount of a particular chemical that a particular country can produce and

consume within a given year. Like the tariff bindings used in the GATT/WTO regime, these commitments are simple legal obligations (not subject to interpretation) and are a relatively easy obligation to police, provided, of course, that the information reported by parties is reliable.[95] As with the trading regime, governments, industry, and NGOs regularly monitor compliance with reporting requirements and, of course, substantive obligations.[96]

Numerous amendments to the 1987 Protocol expanded its reach in terms of chemicals covered (by adding phase-out schedules for newly identified ODSs) and ratcheted up (in terms of the phase-out rate) both the targets and timetables for those chemicals already covered, and added institutional mechanisms including for compliance purposes. As with the GATT 1947, amendments only bind ratifying countries, and there have been a series of amendments. As a result, "the pattern of obligations is complicated. Some states are still party only to the original Protocol, others to the Protocol as amended in London, and others to the Protocol as amended in London and in Copenhagen."[97] In addition to the formal amendment process, which only binds those countries that ratify a particular amendment, the parties adopted a dynamic adjustment process through which they could adjust commitments (targets and timetables) for covered chemicals in a less formal manner that does not require ratification. Through the adjustment process, parties receive notice of an adjustment, and it becomes binding six months thereafter. By contrast, the amendment procedures require a two-thirds vote and ratification and thus make adaptation quite cumbersome, which is particularly debilitating in a regime dependent upon rapid advancements in science and technology.

Given the nature of the interdependency problem, maximizing participation is essential to the long-term success of the Ozone regime and to prevent the entire process from unraveling. Basically, a single country with the potential to domestically produce and consume significant quantities of ODSs presents a risk to the Ozone layer, even if rest of the world were to halt production and consumption entirely. China and India in particular may pose such a risk in the future because of their size and development trajectory. Thus, the regime relies upon numerous flexible compliance mechanisms designed to encourage and sustain participation, including, for example, the grandfathering of plants being constructed in September 1987 into a 1986 baseline (to encourage Soviet participation), and various financial and technical assistance mechanisms. The Montreal Protocol provides for special treatment of developing countries to encourage their initial participation. Specifically, Article V of the Montreal Protocol makes special allowances for developing countries, including a 10-year delay for compliance with targets and timetables, a separate consumption limit of 0.3 kilograms per capita, Multilateral Fund access, and promotion of bilateral assistance programs.[98] Pursuant to the 1990 London Amendment, an interim $180 million Multilateral Fund was established to attract developing countries to the regime by funding the incremental costs incurred by developing countries in meeting their commitments and technology transfer activities. The fund was to expand to $240 million in the event that China and India join the Montreal Protocol, which occurred in 1991 and 1992, respectively.[99] These institutional mechanisms encourage participation by adjusting countries' payoff structures through positive incentives and sustain participation similarly. In addition, the Protocol prohibits trade in controlled substances between a party and any nonparty. This prohibition also adjusts countries' payoff structures, creating

a strong incentive for nonparties to join and a strong disincentive for parties to fall into noncompliance.

These innovative institutional mechanisms were designed to broaden participation and implement a Type I strategy, which, generally, has been quite successful. The ratification rate for the Montreal Protocol has been remarkable – 29 countries and the European Community had ratified the Montreal Protocol when it became effective in 1989, and 183 countries had ratified as of June 14, 2002.[100]

It is important to note, however, that participation has come at a cost. Developing countries have tested the value that developed countries give to protecting the Ozone layer and extracted considerable side-payments in exchange for promised participation. Existing ODS production capacity correlates directly with bargaining power because a country may threaten to utilize such capacity unless concessions are granted. This tactic is particularly successful in cases where the country has a credibly short time horizon. As developed countries have eliminated ODS production capacity, large developing countries that have maintained or even expanded production capacity have gained bargaining strength.[101] As David Downie has observed, large developing countries, such as India, substantially expanded their production capacity in the late 1980s and early 1990s at the same time as developed countries closed production facilities. The long-term success of the regime depends upon the participation of large developing countries.

In 1988, pursuant to Article 6 of the Montreal Protocol, expert panels – the Scientific Assessment Panel, Environmental Assessment Panel, and Technology and Economic Assessment Panel – were established to coordinate research and have since issued a series of important reports on issues relevant to negotiations and implementation. The Panels "involve several hundred scientists worldwide ... [and] have played a crucial role in causing parties to ratchet up targets and timetables, add chemicals to the list of controlled substances, [] address problems such as recycling[, and] ... provided credible risk assessments and evaluations of control options" (Brown Weiss, 1998, p. 146). Through the engagement of experts, the Ozone regime has significantly reduced the uncertainties and transaction costs that could have inhibited international cooperation.

In particular, the Technology and Economic Assessment Panel ("TEAP") has had an important role in facilitating compliance. In addition to its detailed reports on technical options, controls, and ODS substitutes, TEAP has established Technical Options Committees that perform a number of compliance-related functions including monitoring and reviewing implementation by particular ODS-using sectors. Furthermore, TEAP administers the "essential-use" procedure whereby a party may obtain an "essential-use" exemption to allow production or consumption in special cases but only after a detailed examination by TEAP of the party's application. As Owen Greene aptly points out, this procedure performs a key compliance function in that it tends to focus on implementation issues that pose the greatest noncompliance risk; TEAP reviews potential problems and examines possible "problem-solving responses" (Greene, 1998, p. 99). By requiring applicants to submit extremely detailed applications and then having TEAP experts thoroughly review and study the issues raised, potential noncompliance problems may be averted by "detailed and expert guidance and assistance on how to implement the phase-out" (Greene, 1998, p. 99). It is not surprising that this managerial or facilitative approach would work well in situations where implementation

problems relate to technical expertise. Although such situations tend to involve issues that pose the greatest risk of noncompliance, the risk does not tend to derive from opportunism but rather from well-intentioned parties that lack capacity. By engaging technical experts before a compliance problem arises, parties continue to participate and remain in compliance, and thereby avoid noncompliance procedures.

The Protocol created a specific Noncompliance Procedure and established an Implementation Committee comprised of 10 states as well as various procedures for monitoring and assessing compliance and dealing with noncompliance. The regime established potent noncompliance procedures that may be initiated by one party against another or, as practice has demonstrated, by a party that does not believe that it will be able to meet its commitments.[102] Importantly, the Implementation Committee was intentionally designed to be multilateral and nonconfrontational – and decidedly not judicial. As described by Kal Raustiala, the Committee "does not act in a judicial mode; discussions are not couched as legal arguments. Rather, the Committee relies upon facilitation and whatever political pressure emerges from open, transparent discussion of compliance difficulties. It employs what is essentially an administrative, rather than a judicial, approach to noncompliance" (Raustiala, 2000, pp. 418–19). Under the Protocol, it may review reports submitted by parties *ad hoc*, is empowered to make on-site inspections to assess compliance, and to use both positive and negative incentives to encourage compliance or bring a party back into compliance. Positive incentives may take the form of financial or technical assistance while negative incentives may include warnings or suspension of rights or privileges (including, e.g., the ability to trade in controlled substances with other parties). Importantly, positive incentives are generally made contingent upon participation of the party seeking assistance in the form of submission of detailed plans for national programs and/or annual reports on production and consumption of ODSs. As David Victor demonstrates, the noncompliance procedures established by the Montreal Protocol and implemented by the Implementation Committee blend the management-oriented and enforcement-oriented approaches to compliance (Victor, 1998).

The Ozone regime also envisions the possibility of dispute settlement. Specifically, Article 11 of the Vienna Convention suggests that parties should resolve their disputes through negotiation and if that is not successful, through a conciliation commission. The article also provides that parties may elect to settle their disputes through arbitration or the International Court of Justice and that parties may accept the compulsory jurisdiction of one of these mechanisms at the time of their ratification of the treaty. In the end, however, dispute settlement has not been invoked by any of the parties. The noncompliance procedures noted above have instead been universally invoked.

The Ozone regime employs all three compliance strategies. With respect to Type I, the focus is on positive incentives that are aimed at modifying the payoff structure to encourage participation and compliance. In addition, the trade restriction prohibiting trade with nonparties acts as an inducement to join and a deterrent from withdrawal. While there is some prospect for negative sanctions through dispute settlement or suspension of rights by the Implementation Committee, such sanctions have not yet been applied. With respect to Type II, there are a number of institutional mechanisms aimed at reducing uncertainty regarding the problem itself and identifying additional ODSs, and aimed at reducing the transaction or adjustment costs of compliance through facilitative

technology transfer and administrative and technical assistance. Notably, in contrast with the GATT/WTO regime, the need for legalistic mechanisms for interpreting, clarifying, and applying rules and obligations has not been an essential function of the compliance system in the Ozone regime. This is illustrated by the fact that compliance issues thus far have been resolved through the Noncompliance Procedure and the Technology and Economic Assessment Panel (TEAP) process rather than dispute settlement. With respect to Type III, the readjustment of commitments is an integral part of the Ozone regime: first, the dynamic ratcheting accomplished through adjustment and amendment as well as the addition of newly identified ODSs to the regime demonstrates a repeated cycling through the cooperative process stages. Second, and more directly related to compliance, both the Noncompliance Procedure and the TEAP process facilitate the readjustment of commitments as necessary to respond to capacity-related problems. As illustrated by high rates of ratification and compliance, the Type I and Type III strategies employed in the Ozone regime have been quite successful (so far) in getting and keeping parties on board.

Despite the effectiveness of the regime thus far and its well-designed institutional framework, the task of maintaining participation and compliance may become more difficult as developing countries are asked to reduce domestic production and consumption of various ODSs. Given the nature of the problem and use of positive incentives as a means to sustain participation, there is some risk that large developing countries may act opportunistically by refusing to participate or threatening to withdraw from the regime in the future. As David Leonard Downie explains, the political objective underlying the Ozone regime boils down to one of resource allocation, that is, allocating the use of the Ozone layer as a sink for ODSs. Given the nonexcludable nature of the Ozone layer, the "exploitable power to destroy" it becomes a significant source of bargaining power (Downie, 1999, pp. 103–4). As he noted, "[i]n recent years, the relative bargaining strength of the United States, Japan, Australia, and Europe appears to have weakened somewhat (as they eliminated CFC productions) while that of large developing countries has increased (along with their production capacity)" (Downie, 1999, p. 105). It remains to be seen whether the continued flow of positive incentives (e.g., financial assistance and technology transfer), the threat of sanctions, or appreciation of the long-term consequences of destroying the Ozone layer provide sufficient incentives to large developing countries, such as China and India in particular, to encourage long-term cooperation.[103] As Downie aptly points out, the nature of the underlying problem is that of a common pool resource, which means that initial regime success does not guarantee continued success because "[b]argaining dynamics and interests in [the common pool resource] . . . change over time, presenting new obstacles to cooperation as actors face new incentives and opportunities to seek distributive gains."

4 Conclusion

International trade law and international environmental law reflect two different evolutions of both international law and attendant institutional arrangements. International trade law currently reflects a strengthened version of public international law, strengthened in the sense that the traditional weaknesses inherent in public international law

do not prevent States from coordinating their behavior through binding international commitments, effective dispute resolution institutions, and (somewhat) effective means for imposing sanctions. The commitments undertaken by members of the GATT/WTO and other trade agreements generally are binding legal obligations between States, such that a violation of an obligation is enforceable only by a State.[104] International trade law has evolved into a strong version of public international law, in large part, because of the institutions employed by States to resolve disputes and the possibility of using trade sanctions post-dispute-resolution without collapsing the system into trade wars. While the WTO dispute settlement mechanism may be seen as an enforcement-oriented institution that adjusts the payoff structure through the threat of sanctions, in a manner that deters intentional noncompliance, the analysis set forth in this chapter suggests that the more important function served by the dispute settlement institution is to resolve disputes where noncompliance is unintentional either because (1) there is an actual dispute over the rules or obligations, or (2) the noncompliant State lacks the capacity to comply. With respect to the former cause of unintentional noncompliance, the WTO dispute settlement mechanism has thus far proven very effective. With respect to the latter cause, dispute settlement first establishes the existence of noncompliance and then provides the parties with an opportunity to renegotiate. The WTO has recently begun to focus on the capacity problems that developing countries may face and has been developing capacity building and technical assistance programs to enable developing countries that lack the resources to fully participate in the WTO. For example, in November 2001, at the Fourth Ministerial Conference in Doha, Qatar, a broad mandate was declared concerning negotiations on, *inter alia*, numerous implementation issues, capacity building and technical assistance programs, technology transfer, and special and differential treatment provisions.

International environmental law reflects a shift away from the hierarchical structure of international law toward more complex, multilayered legal frameworks that blur lines between public international law, private international law, and domestic law.[105] To the extent that States undertake binding commitments in international environmental agreements,[106] the obligations tend to be positively prescriptive, requiring regulatory action on the part of States that gives rise to subsidiary legal relationships between States and their citizens and among private actors themselves. While some suggest that a broadreaching WTO-like compliance system might improve compliance with international environmental commitments, such a system has not arisen. Instead, States (and other interested entities) often rely on complex, multifaceted, management-oriented compliance systems that utilize, for example, positive incentives to both induce and facilitate compliance as well as institutions that dynamically adjust commitments over time. The compliance system of the Ozone regime is illustrative. The flexible legal regime has permitted the commitments undertaken by States to be adjusted and expanded to new ODSs in accord with scientific and technological progress. Due to the nature of the underlying interdependency problem, maximizing participation is essential to the long-term success of the Ozone regime and to preventing the entire process from unraveling. As a result, parties rely on positive incentives, such as financial assistance, and negative incentives, such as the threat of a trade ban in controlled substances, to adjust the payoff structure and induce participation and continued compliance. For similar reasons, the focus of the compliance institutions

in the Ozone regime is management-oriented and facilitative. Unlike the GATT/WTO regime, dispute settlement has not been utilized in the Ozone regime; instead, the Noncompliance Procedure and the TEAP process substitute a nonadversarial approach to dealing with noncompliance problems.

This chapter provides an avenue for addressing a number of important questions that have surfaced recently in scholarship regarding international cooperation and compliance. Why do States choose to utilize enforcement-oriented compliance institutions and management-oriented compliance institutions? What are the conditions under which States choose one approach over the other or choose to rely on elements of both? Would a WTO-like institution make sense in the international environmental area? If so, why hasn't one arisen yet? Should the "jurisdiction" of the WTO itself be expanded into international environmental regimes? Will the WTO dispute settlement mechanism be sufficient to address compliance dilemmas in the future? Although these questions have not been directly addressed in this chapter, the analytical framework developed in this chapter provides a useful tool for approaching these questions.

Acknowledgments

The author thanks Steve Charnovitz, Jim Hartigan, Margaret Moses, Eugenia Phipps, and Spencer Waller for their very helpful suggestions and comments. In addition, the author thanks Ethan Zelizer for his valuable research assistance. The analysis in this chapter has been developed more fully in "A Dynamic Institutional Theory of International Law," *Buffalo Law Review*, 51 (2003), 679–809, available on ssrn.com.

Notes

1 Raustiala (2000) (distinguishing compliance with rules and the effectiveness of those rules in addressing the underlying problem justifying their creation).
2 Raustiala (2000) (describing "Systems for implementation review" or "SIRs"); Jacobson and Brown Weiss (1998) (adopting a broad definition of compliance that includes conformance with both procedural and substantive obligations as well as with the "spirit of the treaty").
3 Of course, even at the domestic level, where adjudicative and enforcement bodies exist and arguably perform well, compliance is imperfect and never guaranteed. See, e.g., Farber (1999). Some "slippage" is unavoidable and arguably desirable. Efficient slippage may be facilitated through a liability rule. For example, domestic contract law scholars have argued that legal institutions ought to allow a party to breach a contract and sustain damages attributable to its breach under certain circumstances to promote efficiency. See, e.g., Posner (1992). Efficient slippage also may be facilitated through flexible interpretation, modification, or application of the rules. Farber (1999) (discussing how the courts have created "slack" for regulated parties under environmental regulatory regimes). Both approaches, however, also open the door for inefficient strategic behavior.
4 For example, while the United States did not push the level of commitments set forth in the Kyoto Protocol higher and thus was probably not acting opportunistically during those negotiations, if the Kyoto Protocol enters into force, is implemented effectively, and parties comply with their obligations (granted, many unlikely "ifs"), the United States would realize both types of gains by dropping out of the picture. First, the United States would realize the environmental benefits (however marginal) of the parties' greenhouse gas emissions

reductions, and second, it would realize relative gains in the form of competitive advantages in affected industries. See, e.g., Frischmann (2001a) ("Opportunistic States may feign strict commitments, relying on a weak, politicized enforcement process and weak sanctions, such as delayed reductions. These States may wish to appease domestic constituencies on both sides of the issue, or they may wish to exploit the system in order to gain economic benefits – competitive advantages and cash inflow. In either case, when push comes to shove, the strict commitments would fall to the wayside and the climate change regime itself would likely suffer.").

5 Krasner (1983) defined a regime as the "implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations." He further stated that

> Principles are beliefs of fact, causation and rectitude. Norms are standards of behavior defined in terms of rights and obligations. Rules are specific prescriptions or proscriptions for action. Decision-making procedures are prevailing practices for making and implementing collective choice.

See also Koh (1996) (regime is a set of governing arrangements developed by governmental and private parties in a given issue area).

6 At the Fourth Ministerial Conference of the WTO on November 14, 2001, WTO members agreed to launch new trade negotiations pursuant to the Doha Development Agenda, which specifically puts the concerns of developing countries, such as market access and development, at the forefront. See World Trade Organization, Fourth Ministerial Declaration, WT/MIN(01)/DEC/1 (November 20, 2001) (memorializing the agreement made between the 142 nations that attended the talks at Doha), available at http://www.wto.org/english/thewto_e/minist_e/min01_e/mindecl_e.htm (last visited July 14, 2002); Haque (2002).

7 There is a rich literature on the application of game theory to international cooperation. See generally Schelling (1960); Keohane (1984); Axelrod (1984); Rasmusen (1994); Hovi (1998). Game theory encompasses both cooperative games in which promises are assumed to be binding, and noncooperative games in which promises are not necessarily binding. Given the absence of a supranational governance regime, international games are analyzed as noncooperative games.

8 There are various game theoretic approaches to modeling the interactions between domestic and international politics. The seminal paper setting forth the theory of two-level games was written by Putnam (1988).

9 For example, if States commit themselves to curb domestic greenhouse gas emissions and agree to implement their commitments through an international emissions trading system, the effectiveness of the trading system would ultimately depend on private trading and the effectiveness of the overall regime would ultimately depend on State regulation of private parties, i.e., domestic enforcement of emissions limitations.

10 See, e.g., Spero and Hart (1997) ["Congress tends to link trade policy with particular domestic interests (while) the US executive branch often links trade policy with larger foreign policy and foreign economic goals."].

11 See, e.g., Sykes (1999) (discussing the public choice perspective on international trade agreements); Macey (1996) (discussing public choice theory and its application to international cooperation).

12 Acknowledging the differences between (1) realist/rationalist/utilitarian, (2) norm-driven/process-oriented/sociological, and (3) liberal/institutional theories of international relations, this chapter contends that a flexible game theoretic framework is a useful

organization construct that may encapsulate the key considerations of each theory; a full exploration of this idea, however, is beyond the scope of this chapter.

13 *Traditional illustration*: Suppose two co-defendants are being interrogated in separate cells and that the police have sufficient evidence to convict both defendants of a robbery but insufficient evidence of specific factors that would enhance their sentences – say, e.g., violence. The police attempt to get each prisoner to "rat" on the other (defect) regarding his or her violent conduct in exchange for leniency. The payoffs may be the number of years of imprisonment less than the 10-year maximum, for example. If both prisoners refuse to cooperate with the police (and thereby cooperate with each other), they both receive a sentence of 7 years; if one prisoner implicates the other (defects) while the other refuses to do so (cooperates), the defector receives a sentence of 6 years while the cooperator receives a sentence of 9 years; if both prisoners defect, they each receive a sentence of 8 years. If the prisoners understand the payoff structure, are unable to communicate, and have no other assurance mechanism, they will follow their dominant strategies and end up with 8-year sentences. If, however, the prisoners care about their reputations (in or out of jail), expect to partner up in the future, have a preexisting agreement or understanding regarding future dealings (in or out of jail), or have the ability to retaliate against defection, cooperation is more likely.

14 This key assumption illustrates the importance of information and its role in strategic decision making.

15 *Illustration continued*: Suppose that the police in our previous example misrepresented the payoff structure to the co-defendants by exaggerating the possibility of leniency and understating the cumulative effect of additional evidence concerning violence on a jury. In reality, the payoff structure might be that of a Harmony game – cooperating with the police can only lead to more jail time. Unless the prisoners learn of this, however, they may proceed to defect because they view the payoff structure as a Prisoners' Dilemma.

16 Deadlock is a game where the parties each have the dominant strategy to defect and no incentive to cooperate because the joint payoff from cooperation is the minimum joint payoff.

Deadlock

		Party j Cooperate	**Party j** Defect
Party i	Cooperate	2, 2	1, 4
Party i	Defect	4, 1	3, 3

17 ***Chicken***

		Party j Cooperate	**Party j** Defect
Party i	Cooperate	3, 3	2, 4
Party i	Defect	4, 2	1, 1

18 For example, the United States has threatened to use, and at times, has used unilateral sanctions to promote compliance.

19 See discussion below on ozone depletion.

20 For a discussion of the disabling effects of increased numbers on compliance, see, e.g., Vogel and Kessler (1998) (attributing "[m]uch of the variance in national patterns of compliance with the treaties examined ... to the number of sources that require

monitoring"). But see Dannenmaier and Cohen (2000) (discussing the importance of broad participation for compliance purposes), available at http://www.pewclimate.org/projects/compliance.cfm (last visited October 3, 2002).

21 United Nations Conference on Environment and Development: Rio Declaration on Environment and Development, Principle 7, June 14, 1992, UN Doc.A/CONF.151/26/Rev. 1 (vol. I) at 3, reprinted in 31 I.L.M. 874, 878 (1992); see also Harris (1999).

22 Tit for tat is not a penance strategy, because the defector will have to cooperate in a period, even though the other players did not in the previous period, in order to restore cooperation by all.

23 The difficulty with such penalties is that they may not be deemed credible, and hence not deter defection because they are so costly to invoke. In particular, they may hurt the punisher as much as the punishee. Hence, they may be deemed not "renegotiation proof," as players expect that any instance of defection will be followed by immediate renegotiation to cooperation. However, this cooperation will be broken because the punishment is not credible.

24 The difficulty with trigger strategies is making the penalty phase not so onerous that it is neither credible, nor so short that it is not a deterrent.

25 As noted earlier, international relations scholars refer to the "implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations" as regimes. The regime formation process has similarly been divided. See, e.g., Young (1998) (dividing the cycle of regime formation into three stages – agenda formation, negotiation, and operationalization); see also Young (1994).

26 Nonbinding resolutions may also serve the same purpose. Nonbinding agreements may embody commitments with which States intend to comply while at the same time, explicitly reserving flexibility to adjust commitments as necessary. In a sense, nonbinding agreements incorporate the managerial-approach to compliance without institutional support.

27 It may also affect the payoffs for nonparticipants, as in the case of the nonparty trade ban in the ozone regime, discussed later.

28 For an interesting analysis of bargaining to create international governance systems, see Young (1994).

29 States may overestimate the risks of noncompliance. This may inhibit cooperation in the first place if, e.g., the costs of creating an "acceptable" compliance institution exceed the benefits of cooperation, or such overestimation may lead to the creation of unnecessary compliance institutions and thereby raise the costs of cooperation. For example, imagine that two States perceive their joint regulation of a shared problem to be a Prisoners' Dilemma game due to uncertain payoffs where in fact it is a Harmony game. The expected costs of creating and maintaining a compliance institution may exceed the expected benefits of cooperation and thus lead both States to opt out of the potential agreement. Alternatively, the States may enter into an agreement and create an unnecessary compliance institution.

30 Monitoring institutions frequently collect and analyze information relating to the underlying problem itself and the means by which the threat has been addressed; even where full compliance exists, the coordinated effort may be poorly targeted and in need of redirection. See e.g., Chayes and Chayes (1995); Susskind (1994).

31 Brown Weiss (2000b) ("The greater the capacity of the political unit to implement the accord, the more likely it is that it will comply. Administrative and bureaucratic capacity depends upon economic resources, but it also involves education, technical training and skills, and attitudes."); see also Drumbl (1999); Vogel and Kessler (1998).

32 Hovi (1998) (discussing modifications to the basic Prisoners' Dilemma model, including variable discount factors, imperfect detection of violations, incomplete information, multiple options, and package deals).
33 See, e.g., Jacobson and Brown Weiss (1998b) (three strategies are positive incentives, negative incentives, and sunshine); Knox (2001) (discussing adjudicatory model and managerial model).
34 Although generally associated with a negative incentive strategy (i.e., enforcement-oriented approach), detecting noncompliance is essential under a positive incentive compliance strategy as well. A stream of positive incentives (e.g., financial support and technology transfer) over the course of many rounds of play may be contingent upon compliance, such that detected noncompliance would stop the flow of benefits or at least initiate a facilitative response aimed at bringing the party into compliance.
35 See, e.g., Chayes et al. (1998) [discussing self-reporting and noting the Soviet Union's deliberate misreporting of kills to the International Whaling Commission and the increased use of independent reporting and verification as a means of avoiding the "'self-incrimination' problems inherent in (self-reporting) systems"].
36 Both the GATT/WTO system and the Ozone regime work toward their respective goals of liberalized trade and significantly limited consumption of ODSs by successive ratcheting of commitments.
37 Understanding on Rules and Procedures Governing the Settlement of Disputes, April 15, 1994, Final Act Embodying the Results of the Uruguay Round of Multinational Trade Negotiations, Annex 2, arts. 16(4), 17(14), reprinted in The Results of the Uruguay Round of Multilateral Trade Negotiations: The Legal Texts 417, 419 (1994) [hereinafter DSU].
38 Of course, this analysis breaks down at the same points that traditional arguments in favor of "the market" or against "government intervention into the market" break down. As the analysis generally goes for government intervention into domestic markets, the presence of positive or negative externalities may justify government action aimed at encouraging or discouraging production or consumption.
39 Yarbrough and Yarbrough (1992) (explaining that in the mid-nineteenth century, Britain adopted a unilateral trade policy of liberalization but suggesting that such a policy is not viable "[o]nce relation-specific investments enter the picture.").
40 Various theories have been developed to explain the trade liberalization–protectionism paradigm and the conditions under which States adopt certain policies and institutional arrangements. The neoclassical economic explanation of the liberalization–protectionism paradigm supposes that States balance efficiency gains from trade against distributional concerns. Business cycle theories suggest that States will liberalize trade when unemployment is low (or when the economy is booming), but will revert to protectionism when unemployment is high (or in times of contraction). Political power theory focuses on the political power of gainers and losers from liberalized trade. Hegemony theory focuses on the power of a single powerful nation (a hegemon) to coerce or entice liberalization. And there are others. Each theory has its advantages and disadvantages, but none fully explains the institutional form that accompanies trade liberalization. Yarbrough and Yarbrough (1992) apply a strategic organizational approach to this problem and argue that the degree of potential opportunism between trading parties (i.e., the risk of intentional noncompliance) and the effectiveness of third-party mechanisms for enforcing commitments influences the institutional form.
41 Compare Suasion game, discussed *supra*.
42 A race to subsidize would involve competition between governments in "investments" made in a particular industry (or company) to secure a dominant market position in an international market. Government "investment" may take many forms, e.g., direct

subsidies, tax subsidies, and procurement advantages. To the extent that a domestic market constitutes a major portion of the international market, a race to subsidize may involve protectionist measures, e.g., tariff and/or nontariff barriers. A trade war (narrowly defined) would involve retaliation (rather than "subsidy competition"), which may be accomplished through trade sanctions in a separate, unrelated market (e.g., automobiles rather than airplanes).

43 The simple strategic trade policy example discussed above can be expanded into a more complex, theoretical argument for government promotion of particular industries that generate positive externalities and for government measures aimed at ensuring that the surplus thereby created is captured domestically. Research and development ("R&D") intensive industries are a prime example. Grossman and Helpman (1990); Romer (1987); see also Frischmann (2000) (discussing Bayh-Dole Act as potential US response to foreign free-riding on US government research investments). In a sense, this strategic trade policy argument looks very similar to the basic market failure justification discussed below.

44 Elsewhere, I have argued that the traditional "government intervention into the market" analysis is incomplete and perhaps biased toward market-oriented solutions to public goods, governance, and other social problems. Frischmann (2001b). I do not develop that analysis here, however. I will note that international trade is a market-based discipline that concerns how to maximize the functionality of the market mechanism to best provide private goods to society and more generally improve economic welfare, and that international environmental problems generally concern how to best provide (or protect) public goods and common pool resources. While there is no doubt that improved economic welfare permits increased investment in environmental goods and development (technological, social, economic, etc.) that facilitates environmental protection, it is important to remember that the market mechanism, even when functioning very well (i.e., a perfectly competitive market), does not "efficiently" provide or protect public goods.

45 For example, national security is an important public good that may justify particular border measures aimed at inspection and verification of imports and may justify particular explicitly protectionist measures aimed at particular defense-related markets.

46 Hereinafter, "liberalization" will refer to the removal of illegitimate restrictions on international trade.

47 In some cases, "trade liberalization" very well may be a "disguised" Harmony game. Like the police officers that misrepresent the payoff structure (i.e., the consequences of their actions) to the prisoners, domestic lobbyists and special interests may distort decision-makers' perceptions of the payoff structure. In other words, if politicians sought to maximize social welfare rather than the sum of social welfare and campaign contributions (or other special interest-related payoffs), liberalization might be a dominant strategy, at least for some States.

48 Interestingly, it seems that trade negotiations proceed on the basis of exchanging liberalization concessions (e.g., reductions in a particular tariff) where the benefits are seen as increased access to a foreign market for particular domestic industries. Thus, looking at the payoff structure, these concessions may be seen as political benefits that offset the political gains from defection. Bagwell et al. (2002); Sykes (1999). The more diffuse national welfare benefits of cooperating almost seem incidental.

49 General Agreement on Tariffs and Trade, October 30, 1947, 61 Stat. A11, 55 U.N.T.S. 194 (entered into force January 12, 1948) [hereinafter GATT 1947]. "Parties" or "contracting parties" refers to those nations that have acceded to (1) the GATT 1947 under the Protocol of Provisional Application, (2) a Special Protocol (Chile, September 1948), or (3) respective Protocols of Accession.

50 For background on the evolution of the GATT/WTO regime, see Jackson (1997); Jackson, Davey, and Sykes (2002).

51 The provisions described later create substantive legal obligations but also constitute broad principles.

52 "GATT system" or "GATT regime" refers to the "implicit or explicit principles, norms, rules, and decision-making procedures around which (parties' and their citizens') expectations converge" (Krasner, 1983).

53 Paragraph 1(a) of Article II specifically provides: "Each contracting party shall accord to the commerce of the other contracting parties treatment no less favourable than that provided for in the appropriate Part of the appropriate Schedule annexed to this Agreement."

54 The bindings themselves were subject to renegotiation, but the core obligation not to charge a tariff above the relevant binding was not.

55 Dunoff (1994) (noting that MFN clauses have been included in trade agreements since the twelfth century) [citing Staff of Senate Comm. on Finance, 93rd Cong., 2nd Sess., The Most-Favored-Nation Clause Provision, Executive Branch GATT Study No. 9 (Comm. Print 1974)].

56 Notably, this principle is subject to significant exceptions; for example, the use of quotas to maintain agricultural price support schemes is permissible. Article IX, ¶ 2. While the exceptions are an extremely important aspect of the GATT and its evolution as sources of friction, bargaining power, and side-agreements, analysis of the exceptions is beyond the scope of this chapter.

57 In fact, Article XXII is a general consultation provision, specifically providing that

1. Each contracting party shall accord sympathetic consideration to, and shall afford adequate opportunity for consultation regarding, such representations as may be made by another contracting party with respect to any matter affecting the operation of this Agreement.
2. The contracting parties may, at the request of a contracting party, consult with any contracting party or parties in respect of any matter for which it has not been possible to find a satisfactory solution through consultation under paragraph 1.

In 1958, the parties agreed Article XXII could be used to initiate multilateral consultations. An Article XXII request for consultations would trigger notification to the Secretariat and subsequently to other parties who could then join the consultation process if the initial target of the request agreed that the joining country had a "substantial trade interest." *Procedures under Article XXII on Questions Affecting the Interests of a Number of Contracting Parties*, November 10, 1958, GATT B.I.S.D. (7th Supp.) at 24 (1959).

58 Although the language regarding the conditions under which the dispute settlement process may be initiated is obviously quite vague and seemingly quite broad, most disputes have involved (1) an alleged nullification or impairment rather than an alleged impediment to a GATT objective, and (2) a violation of the GATT, rather than "application of some measure" or "any other situation." There have been a few "nonviolation" disputes that rely on the mere existence of a nullification or impairment. See, e.g., EEC-Payments and Subsidies Paid to Processors and Producers of Oilseeds and Related Animal-Feed Proteins, Panel Report, adopted January 25, 1990, 37th Supp. B.I.S.D. (1991).

59 See Understanding Regarding Notification, Consultation, Dispute Settlement and Surveillance, 26th Supp. B.I.S.D. 210 (1980); 1982 Ministerial Declaration on Dispute Settlement Procedures, 29th Supp. B.I.S.D. 13 (1983); 1984 Action on Dispute Settlement Procedures, 31st Supp. B.I.S.D. 9 (1985); 1989 Improvements to the GATT Dispute Settlement Rules and Procedures, 36th Supp. B.I.S.D. 61 (1989); see also Hudec (1993).

60 Professor Hudec (1993) conducted a study of the actions of "losing" parties from 1947 to 1992 and concluded that the losing party eventually accepted the results of an adverse panel report in approximately 90 percent of the cases.

61 There certainly were other deficiencies with the dispute settlement system created by and developed under the GATT. For example, there was no procedure for appellate review of panel reports. Of course, given the fact that panel reports have no precedential value and the existence of the consensus rule, it is not clear that an appellate body would have been at all useful. Still, an appellate body has been a crucial element in the successful development of the WTO dispute settlement system. See later.

62 For the purposes of this chapter, "compensation" generally is used in the ordinary economic sense to refer to the transfer of payoffs from the noncompliant party to the injured party in an amount equivalent to the costs imposed upon the injured party. As will be seen later in the WTO context, there is a technical, legal distinction between "compensation" and "retaliation." A simplified explanation of this distinction is as follows: compensation refers to trade measures implemented by the noncompliant party, such as a reduction in existing tariffs, essentially aimed at offsetting the externalized costs associated with its unlawful action; retaliation refers to trade measures implemented by the injured party, such as the suspension of existing concessions, essentially aimed at offsetting the externalized costs associated with its unlawful action. Significant complications arise when one analyzes the appropriate aim of such measures under the rules of the dispute settlement system. For example, it is not clear whether compensation or retaliation is limited to some measure of externalized cost (e.g., trade effects) or may be broadened to offset more completely the benefits gained by the noncompliant party. The first measure would appear to permit efficient breaches while the latter would not. This issue is beyond the scope of this chapter, however, and accordingly is left for future work.

63 Grane (2001) (thorough analysis of remedies available under GATT and WTO); Jackson, Davey, and Sykes (2002) (emphasizing that the remedies available under the WTO DSU are determined on a prospective basis). In a handful of subsidy and countervailing duties cases, however, retroactive compensation has been recommended by a panel. Grane (2001) (analyzing cases). For example, in *New Zealand – Imports of Electrical Transformers from Finland*, the parties recommended a refund of antidumping duties that were inappropriately collected, but the United States blocked adoption of the report. See *New Zealand – Imports of Electrical Transformers from Finland*, 32nd Supp. B.I.S.D. 55 (1986). The United States later argued that the recommendation was "of an extraordinary retroactive and specific nature." GATT Activities 1992, at 40, quoted in Jackson, Davey, and Sykes (2002). More recently, in *Australia – Subsidies Provided to Producers and Exporters of Automotive Leather*, a panel recommended that an export subsidy that was deemed WTO-inconsistent could only be "withdrawn" by "full repayment." See *Australia – Subsidies Provided to Producers and Exporters of Automotive Leather*, WT/DS126/RW, adopted by the DSB on February 11, 2000, discussed in Jackson, Davey, and Sykes (2002) [noting that (1) the panel report was "quite controversial and heavily criticized at the DSB meeting and (2) there was no appeal because the parties had previously agreed that neither side would appeal].

64 Notably, the text of Article XXIII does not expressly preclude retroactive compensation or punitive sanctions, see *id*. at 310 n. 2, but GATT practice certainly does not evince such measures. Pauwelyn (2000).

65 Pauwelyn (2000).

66 By contrast, certain unilateral actions of the United States may have provided a sufficient deterrent to intentional noncompliance. See discussion later.

67 Linkage to "nontrade" areas, such as the environment or labor, is a more complex dynamic that is not considered here. On the linkage issue, see generally *Symposium: The Boundaries of the WTO*, 96 Am. J. Int'l L. 118 (January 2002).
68 General Agreement on Trade in Services, April 15, 1994, WTO Agreement, Annex 1B, Legal Instruments – Results of the Uruguay Round Vol. 31, 33 I.L.M. 1168 (1994) [hereinafter GATS].
69 Agreement on Trade-Related Aspects of Intellectual Property Rights, April 15, 1994, WTO Agreement, Annex 1C, Legal Instruments – Results of the Uruguay Round Vol. 31, 33 I.L.M. 1197 (1994) [hereinafter TRIPS].
70 See Agreement on Trade in Civil Aircraft, April 15, 1994, WTO Agreement, Annex 4(a), at http://www.wto.org/english/docs_e/legal_e/final_e.htm (last visited October 3, 2002); Agreement on Government Procurement, April 15, 1994, WTO Agreement, Annex 4(b), at http://www.wto.org/english/docs_e/legal_e/final_e.htm (last visited October 3, 2002).
71 It is important to note, however, the positively prescriptive nature of some of the obligations (e.g., TRIPS), which reach significantly deeper into domestic regulatory affairs than the GATT 1947. Another interesting point is that the Agreement on Agriculture, which is one of the Multilateral Agreements in Annex IA, prescribes the "tariffication"of agricultural NTBs such that the GATT 1947 ratcheting down process that has been so successful with tariffs on most goods may be applied to agricultural products.
72 See Moore (2000) (referring to dispute settlement as the backbone of the multilateral trading system); Dispute Settlements in the WTO: Hearing Before the Subcomm. on Int'l Trade of the Senate Comm. on Finance, 106th Cong. 2 (2000) (statement of Sen. Charles Grassley, Chairman, Int'l Trade Subcomm. of the Senate Comm. on Finance) (same).
73 The DSB was established to, *inter alia*, "administer these rules and procedures and, except as otherwise provided in a covered agreement, the consultation and dispute settlement provisions of the covered agreements. Accordingly, the DSB shall have the authority to establish panels, adopt panel and Appellate Body reports, maintain surveillance of implementation of rulings and recommendations, and authorize suspension of concessions and other obligations under the covered agreements." DSU, art. 2, para. 1.
74 See generally *id*.
75 The process itself was greatly improved by a series of reforms, including the creation of established time frames for certain actions such that a default procedural rule would be triggered at the end of the period, thus minimizing the delays endemic to the GATT dispute settlement process. Furthermore, the DSB is vested with the authority to establish panels and to select panel members in the event that the parties cannot agree on the selection.
76 Although the doctrine of *stare decisis* does not formally apply to panel reports and Appellate Body reports, there is a strong argument that, based on WTO dispute settlement practice, the doctrine applies de facto. Bhala (1999a); see also Bhala (1999b); Bhala (2001).
77 DSU, art. 22.
78 Grane (2001) (thorough analysis of remedies available under GATT and WTO); Jackson, Davey, and Sykes (2002) [excerpt from Davey (2001)]; Pauwelyn (2000). See also Jaffe (1996) [noting that (1) the issue of whether compensation may cover past harms is controversial, (2) the United States consistently maintained in GATT practice and during the Uruguay Round negotiations that dispute settlement remedies must be prospective, and (3) the DSU reflects this position].
79 DSU, art. 22, at ¶ 1.
80 *Id*. at ¶ 4.
81 See *supra* note 79.

82 *Decision by the Arbitrators, European Communities – Regime for the Importation, Sale and Distribution of Bananas – Recourse to Arbitration by the European Communities under Article 22.6 of the DSU*, WT/DS27/ARB, April 9, 1999, DSR 1999:II, 725, ¶ 6.3.
83 *Id*. at ¶ 4.3.
84 DSU, art. 22, at ¶¶ 3 and 4.
85 Bishop (2001/2002) (arguing that the WTO March 24, 2000 arbitration ruling granting Ecuador permission to retaliate through cross-sector retaliation under DSU Article 22 ¶ 3 by suspending EU intellectual property rights in Ecuador is "the second revolution in WTO dispute settlement because it effectively levels the playing field between the weaker developing WTO members and the stronger, industrial WTO members").
86 Trade Policy Review Mechanism, April 15, 1994, WTO Agreement, Annex 3, at http://www.wto.org/english/docs_e/legal_e/final_e.htm (last visited October 3, 2002).
87 For an excellent discussion, see Downie (1999).
88 Both public good and common pool resource problems can be "translated" in game theoretic terms into a Prisoners' Dilemma.
89 For recent assessments, see, e.g., Jacobson and Brown Weiss (1998a); Victor et al. (1998); Wettestad (2002).
90 Vienna Convention for the Protection of the Ozone Layer, UNEP Doc. IG.53/5; 26 I.L.M. 1529 (1987) (signed March 22, 1985, entered into force September 22, 1988).
91 The Conference of the Parties ("COP") to the Vienna Convention met every 2 years until 1993, and now meets every 3 years. The Meeting of the Parties ("MOP") to the Montreal Protocol meets every year.
92 See, e.g., Raustiala (2000) ("[N]on-binding instruments provide flexibility in the face of uncertain means and costs.").
93 Soon after the adoption of the Vienna Convention, the first empirical findings regarding the Antarctic Ozone hole were published. A few years later, shortly after the adoption of the Montreal Protocol, more extensive empirical findings strongly indicated that man-made chemicals were responsible for the observed ozone depletion (Wettestad, 2002, p. 157).
94 Montreal Protocol on Substances that Deplete the Ozone Layer, 26 I.L.M. 1550 (1987) (signed September 16, 1987, entered into force January 1, 1989).
95 Although in the "early days" of the Ozone regime many countries failed to submit complete reports and there was concern over the reliability of the data submitted, there have been signs of improvement. See, e.g., Brown Weiss (1998).
96 See, e.g., Brown Weiss (1998) (noting that the "handful of large companies that produce ozone-depleting substances have an important financial stake in ensuring that their competitors abide by the treaty, as well as the resources to monitor compliance, albeit quietly.").
97 Brown Weiss (1998).
98 Notably, parties must be classified as Article V countries to be accorded special treatment, and such classification is not permanent. Rather, an open-ended working group is empowered to classify and declassify countries for Article V status. To become classified and maintain classification, countries must submit detailed reports. Thus, the positive incentive of special treatment is made contingent upon compliance with reporting and data submission requirements. See, e.g., Brown Weiss (1998).
99 The Multilateral Fund and its implementing agencies form an important part of the Ozone regime. Greene (1998). Developing countries must submit detailed applications to obtain funding for particular projects. The applications are reviewed by various implementing agencies. Funding Eligibility Guidelines provide that a project proposing to create a new facility to produce ODS substitutes may not be funded unless the original ODS production

facility is shut down. Greene (1998) (providing example). If granted, the projects are monitored and reviewed to ensure that implementation proceeds as expected. Greene (1998) (describing the evolving implementation review system).

100 Status of Ratification/Accession/Acceptance/Approval of the agreements on the protection of the stratospheric ozone layer, at http://www.unep.org/ozone/ratif.shtml (last visited October 3, 2002).

101 For a discussion of this dynamic, see Downie (1999); see *id*. at 108–9.

102 In 1995, Belarus, Bulgaria, Poland, Russia, and Ukraine made noncompliance submissions regarding themselves that were interpreted as formal applications under the Procedure. For a discussion, see Victor (1998).

103 Of course, even if China, for example, fully intends to take on significant commitments and to comply with its commitments, the "exploitable power to destroy" may nonetheless be leveraged strategically to obtain concessions from developed countries.

104 See, e.g., Charnovitz (2001) ("The WTO agreements are a code of obligations and rights for member governments. None of these obligations apply directly to individual actors.") (footnote omitted). There are, of course, exceptions, such as the investor-State provisions in NAFTA Chapter 11, which permit private party challenges to State action through arbitration. See sources cited in the previous note.

105 Compare Brown Weiss (2000a) (suggesting that the lines are blurring for international law generally).

106 States frequently rely on nonbinding instruments rather than binding legal obligations to advance international environmental objectives. See, e.g., Shelton (2000); Brown Weiss (1997); Raustiala (2000).

References

Alvarez, José E. and John H. Jackson 2002: Afterword: The linkage problem – comments on five texts. Symposium: The Boundaries of the WTO, in *American Journal of International Law*, 96, 118.

Axelrod, Robert 1984: *The Evolution of Cooperation*. New York: Basic Books, chs. 12, 13.

Bagwell, Kyle et al. 2002: It's a question of market access. *American Journal of International Law*, 96, 56.

Bhala, Raj 1999a: The precedent setters: de facto stare decisis in WTO adjudication (Part Two of a Trilogy). *Florida State University Journal of Transnational Law and Policy*, 9, 1.

Bhala, Raj 1999b: The myth about stare decisis and International Trade Law (Part One of a Trilogy). *American University International Law Review*, 14, 845.

Bhala, Raj 2001: The power of the past: towards de jure stare decisis in WTO adjudication (Part Three of a Trilogy). *George Washington International Law Review*, 33, 873.

Bishop, Andrew S. 2001/2002: The second legal revolution in International Trade Law: Ecuador goes ape in banana trade war with European Union. *International Legal Perspectives*, 12, 3.

Brown Weiss, Edith (ed.) 1997: *International Compliance with Nonbinding Accords*. Washington, DC: American Society of International Law.

Brown Weiss, Edith 1998: The five international treaties: a living history. In Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, Mass.: MIT Press, 135–57, 144–5, 146, 153.

Brown Weiss, Edith 1999: Understanding compliance with international environmental agreements: the baker's dozen myths. *University of Richmond Law Review*, 32, 1555.

Brown Weiss, Edith 2000a: The rise or the fall of International Law? *Fordham Law Review*, 69, 345, 351.

Brown Weiss, Edith 2000b: Strengthening national compliance with trade law: insights from environment. In Marco Bronckers and Reinhard Quick (eds.), *New Directions in International Economic Law: Essays in Honour of John H. Jackson*. Great Britain: Kluwer Law International, 458.

Carlson, Lisa 2000: Game theory: international trade, conflict and cooperation. In Ronen Palan (ed.), *Global Political Economy: Contemporary Theories*. London: Routledge.

Charnovitz, Steve 2001: Economic and social actors in the World Trade Organization. *ILSA Journal of International and Comparative Law*, 7, 259.

Chayes, Abram and Antonia Chayes 1995: *The New Sovereignty: Compliance with International Regulatory Agreements*. Cambridge, Mass.: Harvard University Press.

Chayes, Abram et al. 1998: Managing compliance: a comparative perspective. In Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, Mass.: MIT Press.

Cicerone, Ralph et al. 1974: Stratospheric ozone destruction by man-made chloro-fluorocarbons. *Science*, 185, 1165–8.

Cooter, Robert and Thomas Ulen 1997: *Law and Economics*. Reading, Massachusetts: Addison Wesley Longman, 2nd edn., 40–1.

Dannenmaier, Eric and Isaac Cohen 2000: *Promoting Meaningful Compliance with Climate Change Commitments*. http://www.pewclimate.org/projects/, 31–7.

Downie, David L. 1999: The power to destroy: understanding stratospheric ozone politics as a common-pool resource problem. In J. Samuel Barkin and George E. Shambaugh (eds.), *Anarchy and the Environment: The International Relations of Common Pool Resources*. Albany, New York: State University of New York Press, 97–121.

Drumbl, Mark A. 1999: Does sharing know its limits? Thoughts on implementing international environmental agreements: a review of national environmental policies, a comparative study of capacity-building. *Virginia Environmental Law Journal*, 18, 281.

Dunoff, Jeffrey L. 1994: Institutional misfits: the GATT, the ICJ and trade-environment disputes. *Michigan Journal of International Law*, 15, 1043, 1049, and n. 15.

Farber, Daniel A. 1999: Taking slippage seriously: noncompliance and creative compliance in environmental law. *Harvard Environmental Law Review*, 23, 297.

Frischmann, Brett 2000: Innovation and institutions: rethinking the economics of US science and technology policy. *Vermont Law Review*, 24, 347, 407–11.

Frischmann, Brett 2001a: Using the multi-layered nature of international emissions trading and of international-domestic legal systems to escape a multi-state compliance dilemma. *Georgetown International Environmental Law Review*, 13, 463, 500–1.

Frischmann, Brett 2001b: Privatization and commercialization of the internet infrastructure: rethinking market intervention into government and government intervention into the market. *Columbia Science and Technology Law Review*, 2, http://www.stlr.org/cite.cgi?volume=2andarticle=1, 1.

Grane, Patricio 2001: Remedies under WTO law. *Journal of International Economic Law*, 4, 755–72.

Grassley, Sen. Charles A. 2000: Statement in “Dispute settlements in the WTO,” *Hearing Before the Subcomm. on International Trade of the Senate Committee on Finance*, 106th Cong. 2 (Statement of Sen. Charles A. Grassley, Chairman, Int’l Trade Subcomm.).

Greene, Owen 1998: The system for implementation review in the Ozone regime. In David G. Victor et al. (eds.), *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, Mass.: MIT Press, 99, 101–5.

Grossman, Gene M. and Elhanan Helpman 1990: Trade, innovation and growth. *American Economic Review*, 80: 2, 86–91.
Haque, Inaamul 2002: Doha development agenda: recapturing the momentum of multilateralism and developing countries. *American University International Law Review*, 17, 1097.
Harris, Paul G. 1999: Common but differentiated responsibility: the Kyoto Protocol and United States policy. *New York University Environmental Law Review*, 7, 27, 30.
Hovi, Jon 1998: *Games, Threats & Treaties: Understanding Commitments in International Relations*. London: Pinter Pub. Ltd., 96–100.
Hudec, Robert E. 1993: *Enforcing International Trade Law: The Evolution of the Modern GATT Legal System*. Salem, NH: Butterworth.
Jackson, John H. 1997: *The World Trading System: Law and Policy of International Economic Relations*. Cambridge, Mass.: MIT Press, 2nd edn.
Jackson, John H., William J. Davey, and Alan O. Sykes, Jr. 2002: *Legal Problems of International Economic Relations: Cases, Materials and Text on the National and International Regulation of Transnational Economic Relations*. St. Paul, Minn.: West, 4th edn., 6–7, 214–15, 227, 311–12, and ch. 6.
Jacobson, Harold K. and Edith Brown Weiss 1998a: A framework for analysis. In Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, Mass.: MIT Press.
Jacobson, Harold K. and Edith Brown Weiss 1998b: Assessing the record and designing strategies to engage countries. In Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, Mass.: MIT Press.
Jaffe, Matthew P. 1996: Are international institutions doing their job? World Trade Organization. *American Society of International Law Proceedings*, 90, 412, 424–5.
Keohane, Robert O. 1984: *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, N.J.: Princeton University Press.
Knox, John H. 2001: A new approach to compliance with International Environmental Law: the submissions procedure of the NAFTA environmental commission. *Ecology Law Quarterly*, 28, 1.
Koh, Harold Hongju 1996: Transnational legal process. *Nebraska Law Review*, 75, 181, 200.
Krasner, Stephen D. 1983: Structural causes and regime consequences: regimes as intervening variables. In Stephen D. Krasner (ed.), *International Regimes*. Ithaca, NY: Cornell University Press, 1, 2.
Krugman, Paul R. 1987: Is free trade passe? *Journal of Economic Perspectives*, 1, 131–44, reprinted in John H. Jackson et al. 2002: *Legal Problems of International Economic Relations*, St. Paul, Minn.: West, 4th edn. 24–9.
Macey, Jonathan R. 1996: A public choice model of international economic cooperation and the decline of the nation state. *Cardozo Law Review*, 18, 925.
Molina, Mario and F. Sherwood Rowland 1994: Stratospheric sink for chlorofluoromethanes: chlorine atomic catalyzed destruction of ozone. *Nature*, 249, 810–12.
Moore, Mike 2000: WTO's Unique System of Settling Disputes Nears 200 Cases in 2000. World Trade Organization Press Release, http://www.wto.org/english/news_ e/pres00_e/pr180_e.htm, last visited October 3, 2002.
Oakland, W. H. 1987: Theory of public goods. In Alan J. Auerbach and Martin Feldstein (eds.), *Handbook of Public Economics* 2, Oxford: Elsevier Science Ltd., 485–6.
Pauwelyn, Joost 2000: Enforcement and countermeasures in the WTO: rules are rules – toward a more collective approach. *American Journal of International Law*, 94, 335, 337–8.
Posner, Richard A. 1992: *Economic Analysis of the Law*, Boston: Little, Brown and Co., 4th edn. 95–6.

Putnam, Robert 1998: Diplomacy and domestic politics: the logic of two-level games. *International Organization*, 42, 427–60.

Rasmusen, Eric 1994: *Games and Information: An Introduction to Game Theory*, Oxford: Blackwell Publishers, 2nd edn. 16–19.

Raustiala, Kal 2000: Compliance and effectiveness in international regulatory cooperation. *Case Western Reserve International Law Review*, 32, 387, 392–9, 400, 415–16, 418–19, 425–6.

Romer, Paul M. 1987: Growth based on increasing returns due to specialization. *American Economic Review*, 77: 2, 56–62.

Schelling, Thomas 1960: *The Strategy of Conflict*, Cambridge, Mass.: Harvard University Press, 2nd edn.

Schwartz, Warren F. and Alan O. Sykes 2002: The economic structure of renegotiation and dispute resolution in the WTO/GATT system. *John M. Olin Law and Economics Working Paper No. 143* (2nd series), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=302823, last visited October 4, 2002, 12.

Shelton, Dinah (ed.) 2000: *Commitment and Compliance: The Role of Non-Binding Norms in the International Legal System*. New York: Oxford University Press.

Spero, Joan E. and Jeffrey A. Hart 1997: *The Politics of International Economic Relations*, London: Routledge, 5th edn. 49.

Stolarski, Richard and Ralph Cicerone 1974: Stratospheric chlorine: a possible sink for ozone. *Canadian Journal of Chemistry*, 52, 1610–15.

Susskind, Lawrence E. 1994: *Environmental Diplomacy: Negotiating More Effective Global Agreements*. Oxford: Oxford University Press, ch. 6.

Sykes, Alan O. 1999: Regulatory protectionism and the law of international trade. *University of Chicago Law Review*, 66, 24–5.

Tiefer, Charles 2001: SINO 301: how Congress can effectively review relations with China after WTO accession. *Cornell International Law Journal*, 34, 63–4.

Victor, David G. 1998: The Montreal Protocol's non-compliance procedure. In David G. Victor, Eugene B. Skolnikoff, and Kal Raustiala (eds.), *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, Mass.: MIT Press, ch. 4.

Victor, David G., Eugene B. Skolnikoff, and Kal Raustiala (eds.) 1998: *The Implementation and Effectiveness of International Environmental Commitments: Theory and Practice*. Cambridge, Mass.: MIT Press.

Vogel, David and Timothy Kessler 1998: How compliance happens and doesn't happen domestically. In Edith Brown Weiss and Harold K. Jacobson (eds.), *Engaging Countries: Strengthening Compliance with International Environmental Accords*. Cambridge, Mass.: MIT Press, 19, 20–3, 24–5, ch. 4.

Wettestad, Jorgen 2002: The Vienna Convention and Montreal Protocol on ozone-layer depletion. In Edward L. Miles et al. (eds.), *Environmental Regime Effectiveness: Confronting Evidence with Theory*. Cambridge, Mass: MIT Press, 149–70.

Yarbrough, Beth V. and Robert M. Yarbrough 1992: *Cooperation and Governance in International Trade: The Strategic Organizational Approach* 7–13, Princeton, NJ: Princeton University Press, 56–61.

Young, O. R. 1994: *International Governance: Protecting the Environment in a Stateless Society*. Ithaca, NY: Cornell University Press.

Young, O. R. 1998: *Creating Regimes: Arctic Accords and International Governance*. Ithaca, NY: Cornell University Press.

6

The National Treatment Principle in International Trade Law

Michael J. Trebilcock and Shiva K. Giri

Chapter Outline

The application of national treatment obligations, as stated in Article III of the General Agreement on Tariffs and Trade (GATT), in the context of facially non discriminatory national tax and regulatory measures with a disparate impact on imported products, has been one of the most difficult conceptual problems under the World Trade Organization (WTO). In this context, this chapter first reviews the case law on Article III with respect to such measures. GATT and WTO dispute settlement bodies have not been consistent in their interpretation of the key elements of Article III and decisions have been insufficiently informed by the anti protectionist purpose identified in Article I:1. The chapter then suggests a purposive interpretation of Article III by applying an economically oriented approach to define the key concepts of likeness of products and less favorable tax and regulatory treatment of imported products in Articles III:2 and III:4. The chapter argues that the inquiry as to likeness should be based on an existing or potential competitive relationship between imported and domestic products. It is further argued that the inquiry as to less favorable tax or regulatory treatment should focus on whether challenged measures disturb or undermine, actually or potentially, effective equality of competitive opportunities between imports and domestic products. These inquiries could usefully borrow from the case law and practice in developed antitrust jurisdictions around the world in defining relevant product markets and from the antitrust literature on raising rivals' costs.

1 INTRODUCTION

The National Treatment principle, along with the Most-Favored-Nation (MFN) principle, constitute the two pillars of the non discrimination principle that is widely seen as the foundation of the General Agreement on Tariffs and Trade (GATT)/World Trade Organization (WTO) multilateral trading regime.

The National Treatment principle has an ancient genesis in international trade law, arguably dating back to ancient Hebrew Law[1] and then appearing in agreements between Italian city states in the eleventh century,[2] in commercial treaties concluded during the twelfth century between England and continental powers and cities,[3] and in agreements among German city states constituting the Hanseatic League from the twelfth century onwards.[4] The principle was also adopted in various shipping treaties entered into between European powers in the seventeenth and eighteenth centuries,[5] and became commonplace in the trade treaties drawn up in large numbers in the latter part of the nineteenth century,[6] as well as appearing in the Paris and Berne Conventions governing intellectual property rights entered into late in the nineteenth century.[7]

While the principle was heavily undermined in the protectionist policies that characterized international trading relations between the two world wars,[8] bilateral trade agreements negotiated by the United States with various trading partners pursuant to the Reciprocal Trade Agreements Act of 1934 typically included some form of the National Treatment principle,[9] and the United States insisted on its incorporation in the GATT as one of its fundamental principles.[10] The principal initial rationale for the principle was to protect concessions reflected in tariff bindings from being undermined by internal taxes or other regulatory measures that replicated the protectionist effect of the previous tariffs.[11] However, on the insistence of the United States, the principle of National Treatment was applied not only to cases of imports that were subject to tariff bindings but extended to internal taxes and other regulatory measures that had a protectionist or discriminatory impact on imports,[12] even in the absence of such bindings, apparently on the assumption that protectionist policies should be channelled into border measures, especially tariffs, that could then be subject to subsequent negotiated reductions and bindings.

During the early years of the GATT, the principal impediment to imports was high tariffs, and the preoccupation of the GATT members was negotiating reductions in these tariffs on an MFN basis,[13] leaving a relatively minor role for the National Treatment principle in disciplining protectionism or discrimination in international trade. However, with the success of the GATT in reducing tariffs to very low levels by the 1980s,[14] the National Treatment principle began to emerge as an important source of discipline on residual forms of protectionism or discrimination that lay beyond or within each member country's borders.

The principle of National Treatment as embodied in Article III of GATT prohibits discrimination between domestic and foreign goods in the application of internal taxation and government regulations after the foreign goods satisfy customs measures at the border. Article III:1 prohibits the application of internal taxes and other internal charges as well as the laws, regulations, and requirements affecting the internal sale,

offering for sale, purchase, transportation, distribution, or use of products, and internal quantitative regulations requiring the mixture, processing or use of products in specified amounts or proportions, to imported or domestic products so as to afford protection to domestic production. Article III:2, first sentence, prohibits the direct or indirect application of internal taxes or other internal charges of any kind to imported products in excess of those applied, directly or indirectly, to like domestic products. Article III:2, second sentence, prohibits the application of internal taxes or other internal charges to imported or domestic products in a manner contrary to the principles set forth in Article III:1. The explanatory note added to Article III:2 states that a tax conforming to the requirements of Article III:2, first sentence, would be considered to be inconsistent with the provisions of the second sentence only in cases where competition was involved between, on the one hand, the taxed products and, on the other hand, directly competitive or substitutable products that were not similarly taxed. Article III:4 prohibits the accordance of less favorable treatment to imported products than that accorded to like products of national origin in respect of all laws, regulations, and requirements affecting their internal sale, offering for sale, purchase, transportation, distribution, or use.

According to Professor John H. Jackson, "one of the more difficult conceptual problems of GATT rules is the application of the National Treatment obligation in the context of a national regulation or tax which on its face appears to be nondiscriminatory, but because of various circumstances of the market place or otherwise has the effect of tilting the scales against imported products."[15] He claims that because of the language found in GATT Article III paragraph 1 prohibiting taxes or other regulations arranged "so as to afford protection," it could be strongly argued under the GATT that even though a tax (or regulation) appears on its face to be nondiscriminatory, if it has the effect of affording protection, and this effect is not essential to a valid regulatory purpose (as suggested in Article XX), then such tax or regulation is inconsistent with GATT obligations.[16] However, in a situation where the discrimination is made not on the basis of origin of products but on the basis of some other characteristics, it is not easy to distinguish between necessary and legitimate discrimination and illegitimate and trade-restrictive discrimination. Aaditya Mattoo and Arvind Subramanian argue that a difficulty lies in distinguishing between two types of situations – one, a non-protectionist government cannot prevent certain domestic policies from incidentally discriminating against foreign competitors; and two, a protectionist government uses a legitimate objective as an excuse to design domestic policies which inhibit foreign competition.[17] They claim that the challenge is to devise international rules that are sensitive to the difference between these two situations, exonerating the former while preventing the latter.[18]

Under the GATT and WTO dispute settlement systems, the issues of both explicit discrimination, where internal tax and regulatory measures provide explicitly different standards for foreign products as opposed to the standards applicable to domestic products, and implicit or origin neutral discrimination, where an internal tax or regulatory measure makes no distinction as to the origin of products but such a measure has a disparate or disproportionate impact on imported products, have been challenged before GATT panels as well as WTO panels and the Appellate Body.[19] According to Hudec,

the GATT was more preoccupied with explicit or *de jure* discriminatory measures than implicit or *de facto* discrimination.[20] He claims that of the first 207 legal complaints filed with the GATT between 1948 and 1990, only a small number of complaints involved claims of *de facto* discrimination by internal regulatory measures.[21] According to him, the first affirmative ruling sustaining a claim of *de facto* discrimination with regard to an internal regulatory measure was the 1987 panel decision in *Japan – Customs Duties, Taxes and Labeling Practices on Imported Wines and Alcoholic Beverages* (hereinafter 1987 *Japan Alcohol*).[22] However, as Maruyama argues,[23] this trend has changed since 1990 and the WTO dispute settlement system has been more concerned with facially neutral rather than explicitly discriminatory internal tax or regulatory measures.

This chapter first reviews the GATT panel case law on facially non discriminatory internal tax and regulatory measures in Section 2 and then provides a similar review of more recent WTO panel and Appellate Body case law in Section 3. Section 4 provides a critique of this case law, arguing that it has inconsistently adopted literalist, regulatory purpose, and economic approaches to the interpretation of Article III that have been insufficiently informed by a purposive interpretation of the provisions of Article III, reflecting the anti protectionist purpose identified in Article III:1. The chapter argues for an economically oriented test of "like products" in Article III:2 and III:4 that turns on an existing or potential competitive relationship between imported and domestic products. Similarly, it argues for an economically oriented test of less favorable treatment of imported products in Article III:2 and III:4 that focuses on whether challenged measures disturb the competitive equilibrium between imported and domestic products by imposing competitive burdens on the former that are not borne by the latter. Finally, it acknowledges that there may be a need to accommodate incidentally adverse impacts on imported products produced by domestic measures primarily aimed at non-protectionist policy objectives and not at restricting imports but which incidentally and unavoidably have this effect. The chapter does not explore, other than incidentally, the relationship between Article III and Article XX (the Exceptions provision), Article III and Article XI (the prohibition on quantitative restrictions), or Article III and the provisions of the WTO TBT, SPS, or GATS Agreements.

2 Facially Neutral Tax or Regulatory Measures and the Principle of National Treatment under the GATT Dispute Settlement System

As noted above, the question of the legitimacy of a regulatory measure that does not explicitly distinguish between foreign and domestic products but distinguishes on the basis of some characteristics or set of characteristics of the products arises when such a measure imposes burdens or has a disparate impact on foreign products. The central issue with regard to such a question is the criteria according to which the burdens or disparate impact on foreign products are determined to be illegitimate or contrary to the principle of National Treatment.[24] According to Hudec, the central finding required in this regard is the conclusion that imports are being treated less favorably than domestic

products, and the primary sources of differential impact in facially neutral regulatory measures are the distinctions these measures make between one class of products and another.[25] He claims that the finding of discrimination ultimately rests on a finding that the product distinction is illegitimate.[26]

Mattoo and Subramanian also accept that "a determination under Article III hinges on determining whether or not the imported product and its domestic comparator are 'like' each other."[27] They argue that GATT panels lurched between two different doctrinal approaches, which they describe as the "textual" and "contextual" approaches, to interpreting "like products."[28] They cite the example of the Panel Report in 1987 *Japan Alcohol* case as exemplifying the "textual approach" in its sharpest form, the example of the Panel Report in *United States – Measures Affecting Alcoholic and Malt Beverages* (hereinafter *US-Malt Beverages*) as having introduced, and the unadopted Panel Report in *United States – Taxes on Automobiles* as having fully expressed the "contextual approach."[29] According to them, these approaches have the following features:[30]

> The textual approach has the following features: first, it defines likeness *a priori* in terms of one or a combination of product characteristics, its end-use and its tariff classification; second it makes a distinction between "like" products and "directly competitive or substitutable" products in a manner faithful to the two sentences in Article III:2, and applies different standards of discrimination to the two cases; third, it preserves a distinct role for Article XX and other exceptions provisions in that they could come into play once (and only after) a measure is deemed to transgress Articles III.
>
> The contextual approach has the following features: first, it does not attempt to define likeness *a priori*; rather it allows *any* distinction to be made between products on regulatory grounds; and second, the standard for determining whether an infraction of Article III has occurred is to ensure that no protectionist intent underlies the distinction nor that any protectionist effect follows from it. In effect, this gives governments the freedom to define likeness, thereby permitting a larger set of measures to be deemed origin-neutral, and *prima facie*, consistent with Article III.

1987 *Japan Alcohol* case[31] was the first significant case brought before the GATT that involved the issue of facially neutral measures and that led to an affirmative ruling sustaining a claim that such measures were contrary to the principle of National Treatment set out in Article III of GATT. The issue in this case was an internal tax measure that classified alcoholic beverages into different categories, subcategories, and grades, based on alcohol content and other qualities, and set different tax rates on each category of alcoholic beverages. The European Communities complained that the Japanese liquor tax system violated the first sentence of Article III:2, by taxing imports at higher rates than "like" domestic products, and the second sentence of Article III:2 by affording protection to "directly competitive or substitutable" domestic products. Japan responded by arguing that each contracting party to the GATT was free to classify products for tax purposes as it chose and that the "likeness" or "directly competitive or substitutable" relationship of imported and domestic products were legally irrelevant to the interpretation of Article III if both of these products were taxed in a non-discriminatory manner, regardless of their origin.

The panel concluded that the ordinary meaning of Article III:2 in the light of its object and purpose[32] supported the practice of examining the conformity of internal

taxes with Article III:2 by determining, firstly, whether the taxed imported and domestic products were "like" or "directly competitive or substitutable," and, secondly, whether the taxation was discriminatory (first sentence) or protective (second sentence). The panel began its examination of the "likeness" of products by noting that GATT contracting parties had never developed a general definition of the term "like products." However, it found the prior GATT decisions on this question were made on a case-by-case basis after examining a number of relevant factors. It cited the Working Party Report on "Border Tax Adjustments" adopted in 1970 (BISD 18S/102) which concluded that problems arising from the interpretation of the terms "like" or "similar" products should be examined on a case-by-case basis using the following criteria: (i) the product's end-uses in a given market; (ii) consumers' tastes and habits which change from country to country; and (iii) the product's properties, nature, and quality. It applied the above criteria and other criteria recognized in previous GATT practice, such as Customs Co-operation Council nomenclature for the classification of goods in customs tariffs, to determine whether the alcoholic beverages classified by Japanese law into different categories, subcategories, and grades were "like" products. The panel concluded, in view of their similar properties, end-uses, and usually uniform classification in tariff nomenclatures, that imported and Japanese-made gin, vodka, whisky, grape brandy, other fruit brandy, certain classic liquors, unsweetened still wine, and sparkling wines should be considered as "like" products in terms of Article III:2, first sentence, because such "likeness" of these alcoholic beverages was recognized not only by governments for the purposes of tariff and statistical nomenclature, but also by consumers to constitute "each in its end-use a well defined and single product intended for drinking" and that minor differences in taste, color, and other properties did not prevent products from qualifying as "like products." The panel did not rule out the possibility of considering other alcoholic beverages as "like products" and it was of the view that the "likeness" of the products must be examined taking into account not only objective criteria, such as manufacturing and composition processes of products, but also subjective consumer viewpoints, such as consumption and use by consumers. However, the panel cautioned that consumer habits were variable in time and space and differential taxes could be used to crystallize consumer preferences for traditional domestic products. It argued that "like" products do not become "unlike" merely because of differences in local consumer traditions within a country or differences in their prices, which were often influenced by government measures (e.g., customs duties) and market conditions (e.g., supply and demand, sales margins).

The panel further concluded that even if imported alcoholic beverages, for example, vodka, were not considered to be "like" Japanese alcoholic beverages, for example, shochu, flexibility in the use of alcoholic drinks and their common characteristics often offered an alternative choice for consumers leading to a competitive relationship. In the view of the panel, under Article III:2 second sentence, there was direct competition or substitutability[33] between imported and Japan-made distilled liquors including all grades of whiskies/brandies, vodka, and shochu, among each other; imported and Japan-made liquors among each other; imported and Japan-made sweetened and unsweetened wines among each other; and imported and Japan-made sparkling wines among each other.

After having compared the imported and domestic alcoholic beverages to determine their "likeness" or "directly competitive or substitutable relationship," the panel next proceeded to a comparison of the fiscal burdens on the products at issue in the dispute. The panel noted that Article III:2 first sentence prohibited the direct or indirect imposition of "internal taxes or other internal charges of any kind in excess of those applied, directly or indirectly, to like domestic products." Thus, a prohibition of tax discrimination was strict. Even very small tax differentials were prohibited, and a *de minimis* argument based on allegedly minimal trade effects was not relevant. In assessing whether there was tax discrimination, account was to be taken not only of the rate of the applicable internal tax but also of taxation methods (e.g., different kinds of internal taxes, direct taxation of the finished product, or indirect taxation by taxing the raw materials used in the product during the various stages of its production) and of the rules for tax collection (e.g., basis of assessment). After having noted that Japanese specific tax rates on imported and Japanese special grade whiskies/brandies were considerably higher than the tax rates on first and second grade whiskies/brandies, the panel found that these tax differentials did not correspond to objective differences between the various distilled liquors, for instance, nondiscriminatory taxation of their respective alcohol contents. In the opinion of the panel, as a result of this differential taxation of "like products," almost all whiskies/brandies imported from the EEC were subject to the higher rates of taxes whereas more than half of whiskies/brandies produced in Japan benefited from considerably lower rates of taxes, and thus, the whiskies/ brandies imported from the EEC were subject to internal Japanese taxes in excess of those applied to like domestic products in the sense of Article III:2, first sentence.

With regard to the mixed system of specific and *ad valorem* taxes adopted by Japan, the panel was of the view that such a mixed system was not as such inconsistent with Article III:2 because it prohibited only discriminatory or protective taxation of imported products but not the use of differentiated taxation methods, provided the differentiated taxation methods did not result in discriminatory or protective taxation. Since the *ad valorem* taxes were not applied to all liquor categories such as the traditional Japanese products shochu, mirin, and saké, the panel found that the differences as to the applicability and non taxable thresholds of the *ad valorem* taxes were not based on corresponding objective product differences, such as alcohol content, nor formed part of a general system of internal taxation equally applied in a trade-neutral manner to all "like" or "directly competitive" liquors. For this reason and for the reason that liquors above the non taxable thresholds were subjected to *ad valorem* taxes in excess of the specific taxes on "like" liquors below the threshold, the panel concluded that the imposition of *ad valorem* taxes on wines, spirits, and liquors imported from EEC was inconsistent with Article III:2, first sentence. Regarding the different methods of calculating *ad valorem* taxes on imported and domestic liquors, the panel agreed that Article III:2 did not prescribe the use of any specific method or system of taxation. There could be objective reasons proper to the tax in question, which could justify or necessitate differences in the system of taxation for imported and for domestic products. It could also be compatible with Article III:2 to allow two different methods of calculation of price for tax purposes. What mattered was whether the application of

the different taxation methods actually had a discriminatory or protective effect against imported products.

Under the first sentence of Article III:2, the tax on the imported product and the tax on the like domestic product had to be equal in effect, but Article III:2, second sentence, prohibited only the application of internal taxes to imported or domestic products in a manner "so as to afford protection to domestic production." Small tax differentials could influence the competitive relationship between directly competing products, but the existence of protective taxation could be established "only in light of the particular circumstances of each case" and "there could be a *de minimis* level below which a tax difference ceases to have the protective effect" prohibited by Article III:2, second sentence.

The panel found that the Japanese tax system was applied "so as to afford protection to domestic production" because of considerably lower specific tax rates on domestic products, and the imposition of high *ad valorem* taxes on most imported products but the absence of *ad valorem* taxes on most domestic products. Similarly, the product taxed at lower rates was almost exclusively produced in Japan, and the mutual substitutability of domestic products with imported products was illustrated by increasing imports of like products and consumer use. According to the panel, Article III:2 protects expectations on the competitive relationship between imported and domestic products rather than expectations on trade volumes. Therefore, it was not necessary to examine the quantitative trade effects of these tax differentials for its conclusion that the application of considerably lower internal taxes by Japan on exclusively domestic products than on directly competitive or substitutable imported products had trade-distorting effects affording protection to domestic production contrary to Article III:2, second sentence.

The 1987 *Japan Alcohol* case was related to internal tax measures and to the issues required to be examined in determining the consistency or inconsistency of such a measure with Article III:2. Since the language of Article III:4 which is related to non tax regulatory measures is different from that of Article III:2, particularly in regard to the treatment required to be provided to the imported products compared to the domestic products, it is necessary to examine separately how the GATT panels interpreted Article III:4 in the context of determining the consistency of a facially neutral regulatory measure with the National Treatment principle.[34]

Although the regulatory measures in dispute were based on the country of origin of products, *United States – Section 337 of the Tariff Act of 1930* (hereinafter *US – Section 337*)[35] was an important GATT case with regard to the issues required to be examined in determining the consistency of a non tax regulatory measure with Article III:4. In this case, the panel had to determine whether US patent enforcement procedures, which were formally different for imported and for domestic products, violated Article III:4.

Since there was no dispute on the "likeness" of domestic and imported products affected by the measure, the panel mainly examined the meaning of the terms "laws, regulations, and requirements" and "no less favorable treatment" as provided in Article III:4, and how an assessment should be made as to whether the regulatory measure in dispute does or does not accord imported products less favorable treatment than that accorded to "like" domestic products. With regard to the meaning of the terms "laws, regulations, and requirements," the panel concluded that not only substantive laws, regulations, and requirements but also procedural laws, regulations, and requirements are

covered by Article III:4. According to the panel, Article III:4 is intended to cover not only the laws and regulations which directly govern the conditions of sale or purchase but also any laws or regulations which might adversely modify the conditions of competition between the domestic and imported products on the internal market.

With regard to the "no less favorable treatment" standard of Article III:4, the panel stated that the "no less favorable treatment" requirement set out in Article III:4 is unqualified as an expression of the underlying principle of equality of treatment of imported products as compared to the treatment given to the domestic products. According to the panel, the words "treatment no less favorable" call for effective equality of opportunities for imported products, as a minimum permissible standard, in respect of the application of laws, regulations, and requirements affecting the internal sale, purchase, transportation, distribution, or use of products. The panel said:

> On the one hand, contracting parties may apply to imported products different formal legal requirements if doing so would accord imported products more favorable treatment. On the other hand, it also has to be recognized that there may be cases where the application of formally identical legal provisions would in practice accord less favorable treatment to imported products and a contracting party might thus have to apply different legal provisions to imported products to ensure that the treatment accorded them is in fact no less favorable.

Therefore, according to the panel, the mere fact that imported products are subject to legal provisions that are different from those applying to domestic products is in itself not conclusive in establishing inconsistency with Article III:4. With regard to the issue of how an assessment should be made as to whether the regulatory measure in dispute accords imported products less favorable treatment than that accorded to "like" domestic products, the panel rejected the respondent's claim that this determination could only be made on the basis of an examination of the actual effects of the regulatory measure. Relying on the previous panel decision in United States – Taxes on Petroleum and Certain Imported substances (GATT, BISD 34S/136, 138, Report of the Panel adopted on June 17, 1987) *Japan Alcohol Beverages* that the purpose of Article III is to protect expectations on the competitive relationship between imported and domestic products, the panel concluded that in order to establish whether the "no less favorable" treatment standard of Article III:4 is met, it had to assess whether or not the contested regulatory measure in itself may lead to the application to imported products of treatment less favorable than that accorded to domestic products. Any decision in this regard should be based on the distinctions made by the contested measure itself and on its potential impact rather than on the actual consequences for specific imported products or actual trade effects.

United States – Restrictions on Imports of Tuna[36] was a significant, but controversial, GATT case involving the issue of "like products" within the meaning of Article III:4. Although the main contested issue in this case was whether the measures prohibiting certain yellowfin tuna and tuna products from Mexico on the ground that the tunas were caught by a dolphin-unfriendly process were internal quantitative restrictions on imports under Article XI or internal regulations under Article III:4, the panel concluding that the measures did not constitute internal regulations covered by Article III:4, the panel made an alternate ruling on the issue of the US measures' consistency with Article III:4 and concluded that even if the contested measures were regarded as internal regulations under Article III:4, they would still not meet the requirement of Article III. Giving the

reasons for such a conclusion, the panel said:

Article III:4 calls for a comparison of the treatment of imported tuna as a product with that of domestic tuna as a product. Regulations governing the taking of dolphins incidental to the taking of tuna could not possibly affect tuna as a product. Article III:4 therefore obliges the United States to accord treatment to Mexican tuna no less favorable than that accorded to United States tuna, whether or not the incidental taking of dolphins by Mexican vessels corresponds to that of United States vessels.

The panel in this case implied that that the difference in fishing methods do not make the two tuna products unlike products within the meaning of Article III:4, but the product-process distinction drawn in this case has been the subject of intense subsequent controversy[37], as we discuss further below.

US Malt Beverages case[38] was another significant GATT case involving facially neutral measures. The test applied to determine the consistency or inconsistency of such measures with Article III was significantly different from that applied in the earlier cases. In this case, Canada had complained, among other things, that a lower tax rate applied by the state of Mississippi to wines made from a certain variety of grape discriminated against "like" Canadian products and was therefore inconsistent with Article III:1 and Article III:2, and that restrictions on points of sale, distribution and labeling based on the alcohol content of beer above 3.2 percent by weight maintained by some US states were inconsistent with Article III:4 since all beers, whether containing an alcohol content of above or below the said level, were "like" products and an alcohol level of 3.2 percent was entirely arbitrary. The panel in this case considered that Canada's claim depended upon whether wine imported from Canada was "like" the domestic wine in Mississippi made from the specified variety of grape that qualified for special tax treatment, and noted that past decisions on the question of "likeness" had been made on a case-by-case basis after examining a number of relevant criteria, such as the product's end-uses in a given market, consumers' tastes and habits, and the product's properties, nature, and quality. However, it considered that the "like" product determination under Article III:2 should have regard to the purpose of the Article, which was not to prevent contracting parties from using their fiscal and regulatory powers for purposes other than to afford protection to domestic production. The panel concluded that the purpose of Article III was not to prevent contracting parties from differentiating between different product categories for policy purposes unrelated to the protection of domestic production. Consequently, in determining whether two products subject to different treatment were like products, it was necessary to consider whether such product differentiation was being made "so as to afford protection to domestic production." Unlike 1987 *Japan Alcohol* case, the panel began its examination by looking into the rationality of the product differentiation made by the Mississippi wine tax law. The panel found that the special treatment accorded to wine produced from a particular type of grape grown only in the South-eastern United States and Mediterranean region was a rather exceptional basis for a tax distinction, and that this particular tax treatment implied a geographical distinction which afforded protection to local production of wine to the disadvantage of wine produced where the type of grape could not be grown. Since tariff nomenclatures and tax laws, including those at the US federal and state level, did not generally make such a distinction between still wines on the basis of the

variety of grape used in their production, and the United States also did not claim any public policy purpose for the tax provision other than to subsidize small local producers, the panel concluded that unsweetened still wines were "like" products and that the particular distinction in the Mississippi law in favor of still wine of a local variety must be presumed to afford protection to Mississippi vintners. Therefore, according to the panel, the lower rate of excise tax applied by Mississippi to wine produced from the specified variety of grape was inconsistent with Article III:2, first sentence.[39]

On the issue of whether the restrictions on points of sale, distribution, and labeling based on the alcohol content of beer were inconsistent with Article III:4, the panel again examined, first, the rationality of the regulatory measure in making a distinction between low alcohol beer and high alcohol beer and then, the competitive effects of such regulations. It stated that the purpose of Article III was not to harmonize the internal taxes and regulations of contracting parties. In the view of the panel, it was imperative that the "like" product determination in the context of Article III be made in such a way that it does not unnecessarily infringe upon the regulatory authority and domestic policy options of contracting parties. Therefore, even if low alcohol beer and high alcohol beer were similar on the basis of their physical characteristics, they need not be considered as "like" products in terms of Article III:4 if the differentiation in the treatment of low alcohol beer and high alcohol beer was not such "as to afford protection to domestic production."

In determining the validity of the regulatory distinction based on the alcohol content of beer, the panel examined the issue of whether the aims and effects of such a regulatory measure showed that it was applied so as to favor domestic producers over foreign producers. From the legislative history of relevant laws, the panel found that the policy background of the laws distinguishing alcohol content of beer was the protection of human health and public morals or the promotion of a new source of government revenue, and the alcohol content of beer had not been singled out as a means of favoring domestic producers over foreign producers. With respect to the effects of the regulatory measure, the panel found that Canadian and US beer manufacturers produced both high and low alcohol content beer, and that the regulatory measure did not differentiate between imported and domestic beer as such, so that where a state law limited the points of sale of high alcohol content beer or maintained different labelling requirements for such beer, that law applied to all high alcohol content beer regardless of its origin. Similarly, the burdens resulting from the measures did not fall more heavily on Canadian than US producers and despite the physical similarities and overlapping in the market for the two types of beer, there was a certain degree of market differentiation or specialization.[40] Therefore, according to the panel, the regulatory measures were consistent with Article III.4.

The "aims-and-effects" approach to determining "likeness" that was applied for the first time in *US Malt Beverages* was also applied and elaborated on in the unadopted GATT panel decision in *United States – Taxes on Automobiles*.[41] In this case, the EEC had complained against US regulations that imposed a luxury excise tax and gas-guzzler tax on domestic and imported automobiles on the basis of their value and gasoline consumption per mile. The threshold value of automobiles for the luxury excise tax was $30,000 and the threshold gasoline consumption for the gas-guzzler tax was 22.5 mpg. Automobiles that were above the stated thresholds were subject to

higher levels of tax. Most of the automobiles imported to the United States from the EEC were more expensive and subject to a higher rate of taxes.[42]

The panel proceeded to determine the "likeness" of the automobiles in question by examining the protective aim and effect of these tax measures. Although there was evidence that the protective effects of these measures had not been ignored during the formulation of regulations providing for one of the taxes, the panel found that these tax measures served a bona fide regulatory purpose and the competitive effects of these measures were neither clear nor inherent enough to be considered as protective. Applying the inherence criterion, the panel attempted to evaluate whether the regulations inherently divided products into those of domestic or foreign origin. Using this criterion, the panel found that the threshold set for the gas-guzzler tax did not discriminate between automobiles of domestic and foreign origin because the technology to manufacture high fuel-economy automobiles – above the 22.5 mpg threshold – was not "inherent" to the United States, nor were low fuel-economy automobiles inherently of foreign origin. Such an advantage would not, therefore, alter the conditions of competition in favor of domestic automobiles, and thereby have the effect of affording protection to domestic production. The panel applied the same "inherence" test to conclude that the threshold set for the luxury excise tax also did not discriminate between automobiles of domestic and foreign origin because no evidence had been advanced that foreign automobile manufacturers did not in general have the design, production, and marketing capabilities to sell automobiles below the stipulated threshold, or that they did not in general produce such models for other markets.

3 Facially Neutral Tax or Regulatory Measures and the Principle of National Treatment under the WTO Dispute Settlement System

After the establishment of the WTO, the panels and Appellate Body under the WTO, as under the GATT dispute settlement system, have also addressed various internal tax and regulatory measures which were facially neutral but were claimed to violate the principle of National Treatment as set out in Paragraphs 1, 2, and 4 of Article III of GATT. Although the GATT panels had taken two different approaches; that is, a textual or "like" product approach as applied in the 1987 *Japan Alcohol* case and a contextual or "aim-and-effect" approach as applied in *US – Malt Beverages* and *US – Taxes on Automobiles*, in examining the validity of a facially neutral regulatory measure, the WTO panels and Appellate Body have rejected the "aims-and-effects" approach to test the validity of any measures which are claimed to violate the provisions of Article III and have accepted that the "like product" approach taken in the 1987 *Japan Alcohol* case is the proper approach.[43]

As there are differences in the National Treatment obligations set forth in Article III:2 with respect to internal tax measures and the National Treatment obligations set forth in Article III:4 with respect to other regulatory measures, it is appropriate to examine separately the interpretations adopted by WTO panels and the Appellate Body of Article III:2 and Article III:4.

3.1 Internal Tax Measures and National Treatment[44]

The first WTO case under Article III involving a facially neutral internal tax measure is the second *Japan – Taxes on Alcoholic Beverages* case. The requirements set out in this case in order to prove that such a tax measure violates Article III of GATT have been consistently followed by other WTO panels and the Appellate Body in other cases involving internal tax measures, such as *Canada – Certain Measures Concerning Periodicals*, *Korea – Taxes on Alcoholic Beverages*, *Chile – Taxes on Alcoholic Beverages*, and *Indonesia – Certain Measures Affecting the Automobile Industry*. According to the Appellate Body's decision in *Japan – Taxes on Alcoholic Beverages,*[45] when an issue is raised that an internal tax measure violates the National Treatment obligation set out in Article III:2, first sentence, the words of the first sentence require an examination of the conformity of an internal tax measure with Article III by determining, first, whether the taxed imported and domestic products are "like" and, second, whether the taxes applied to the imported products are "in excess of" those applied to the like domestic products. If the imported and domestic products are "like products," and if the taxes applied to the imported products are "in excess of" those applied to the like domestic products, then the measure is inconsistent with Article III:2, first sentence. The Appellate Body claimed that this approach to an examination of Article III:2, first sentence, was consistent with the object and purpose of Article III:2 and with past practice under the GATT 1947.

According to the Appellate Body, if the imported and domestic products are not "like" products for the purposes of Article III:2, first sentence, then they are not subject to the strictures of Article III:2, first sentence, and there is no inconsistency with the requirements of that sentence. However, depending on their nature, and on the competitive conditions in the relevant market, those products may well be among the broader category of "directly competitive or substitutable products" that fall within the domain of Article III:2, second sentence. In such a case, a separate examination is required to determine the consistency of an internal tax measure with Article III:2, second sentence. In the view of the Appellate Body, three issues[46] must be established separately in this examination in order to find that a tax measure imposed is inconsistent with Article III:2, second sentence. These three issues are: (i) whether the imported products and domestic products "are directly competitive or substitutable products"; (ii) whether the directly competitive or substitutable imported and domestic products are "not similarly taxed"; and (iii) whether the dissimilar taxation of the directly competitive or substitutable imported and domestic products is "applied … so as to afford protection to domestic production."

According to the Appellate Body, Article III of GATT obliges Members of the WTO to provide equality of competitive conditions for imported products in relation to domestic products. The Appellate Body said that it is irrelevant that "the trade effects" of tax differentials between imported and domestic products, as reflected in the volumes of imports, are insignificant or even nonexistent, as Article III protects expectations not of any particular trade volume but rather of the equal competitive relationship between imported and domestic products.

With regard to the difference in the tests for "like" products and "directly competitive or substitutable," the Appellate Body claimed that this is due to the difference in wording of the first and second sentences of Article III:2. Article III:2, first sentence, does not refer specifically to the general principle of National Treatment articulated in Article III:1 which requires that internal tax and other regulatory measures should not be applied so as to afford protection to domestic production, whereas the language of Article III:2, second sentence, which contains a general prohibition against internal taxes or other internal charges applied to imported or domestic products in a manner contrary to the principles set forth in Article III:1, specifically invokes Article III:1. The Appellate Body argued that the omission of any reference to Article III:1 in Article III:2, first sentence, and the specific invocation of Article III:1 in Article III:2, second sentence, must have some meaning, and the meaning is simply that the presence of a protective application need not be established separately from the specific requirements that are included in the first sentence in order to show that a tax measure is inconsistent with the general principle set out in the first sentence. In the view of the Appellate Body, this does not mean that the general principle of Article III:1 does not apply to the first sentence. The first sentence of Article III:2 is, in effect, an application of the general principle set forth in Article III:1.

By establishing the above-mentioned standards for examination of the conformity of an internal tax measure with Article III:2, the Appellate Body seems to have accepted the panel's rejection of an "aims-and-effects" test to determine the validity of an internal tax measure. Although they reached opposite results by applying essentially the same test, both the complainant, the United States, and the respondent, Japan, had argued at the panel level,[47] as well as before the Appellate Body,[48] that the contested internal tax measure including the product distinction made for tax purposes should be examined in the light of its aims-and-effects in order to determine whether or not it is consistent with Article III:2, and where the aim and effect of the contested tax measure do not operate so as to afford protection to domestic production, no inconsistency with Article III:2 can be established. Such arguments by Japan and United States were based upon rulings and findings by the GATT panels in *US Malt Beverages* and *US – Taxes on Automobiles* cases. The panel simply rejected the "aims-and-effects" test applied in these GATT cases, stating that the panel was not in a position to detect how the 1992 *US Malt Beverages* panel weighed the different criteria that it took into account in order to determine whether the products in dispute were like, that the panel report in *US – Taxes on Automobiles* remained unadopted, and that even if a panel could find useful guidance in the reasoning of an unadopted panel report that it considered to be relevant, unadopted panel reports have no legal status in the GATT or WTO system since they have not been endorsed by the Contracting Parties to the GATT or WTO Members.[49]

The panel gave the following reasons for rejecting the "aims-and-effects" test: first, such a test is not consistent with the wording of Article III:2, first sentence, as the basis of this test is the words "so as to afford protection" contained in Article III:1, and Article III:2, first sentence, contains no reference to these words; second, the adoption of such a test would have important implications for the burden of proof imposed on the complainant because according to this test, the complainant would have the burden of showing not only the effect of a particular measure, which is, in principle, discernible, but also its aim, which sometimes can be indiscernible; third, very often

there is a multiplicity of aims that are sought through enactment of legislation and it would be a difficult exercise to determine which aim or aims should be determinative for the "aims-and-effects" test; fourth, access to the complete legislative history, which is argued by proponents of this test to be relevant to detect the protective aims, could be difficult or even impossible for a complainant to obtain, and even if the complete legislative history is available, it would be difficult to assess which kinds of legislative history (statements in legislation, in official legislative reports, by individual legislators, or in hearings involving interested parties) should be primarily determinative of the aims of the legislation; and fifth, the list of exceptions contained in Article XX of GATT could become redundant or useless because the aims-and-effects test does not contain a definitive list of grounds justifying departure from the National Treatment obligations incorporated in Article III.[50]

With regard to the definition of "like products" in Article III:2, first sentence, the Appellate Body agreed with the panel's conclusion that this term should be construed narrowly so as not to condemn measures that its strict terms are not meant to condemn, because the second sentence of Article III:2 provides for a separate and distinctive consideration of the protective aspect of a measure in examining its application to a broader category of products that are not "like products" as contemplated by the first sentence. According to the Appellate Body, how narrowly is a matter that should be determined separately for each tax measure in each case. The Appellate Body agreed with the practice under the GATT 1947 of determining whether imported and domestic products are "like" on a case-by-case basis in accordance with the criteria, including the product's properties, nature, and quality, the product's end-uses in a given market, and consumer tastes and habits, which change from country to country, set out in the 1970 adopted Report of the GATT Working Party on Border Tax Adjustments. However, the Appellate Body cautioned that in applying the criteria cited in the Border Tax Adjustments Report to the facts of any particular case, and in considering other criteria that may also be relevant in certain cases (such as tariff classifications), panels can only apply their best judgment in determining whether in fact products are "like." Although the Appellate Body did not agree with the panel's observation that distinguishing between "like products" and "directly competitive or substitutable products" under Article III:2 is an arbitrary exercise, it acknowledged that this would always involve an unavoidable element of individual, discretionary judgment, which must be made in considering the various characteristics of products in individual cases. The Appellate Body said:

> No one approach to exercising judgment will be appropriate for all cases. The criteria in Border Tax Adjustments should be examined, but there can be no one precise and absolute definition of what is "like." The concept of "likeness" is a relative one that evokes the image of an accordion. The accordion of "likeness" stretches and squeezes in different places as different provisions of the WTO Agreement are applied. The width of the accordion in any one of those places must be determined by the particular provision in which the term "like" is encountered as well as by the context and the circumstances that prevail in any given case to which that provision may apply. We believe that, in Article III:2, first sentence of the GATT 1994, the accordion of "likeness" is meant to be narrowly squeezed.

Regarding the relevance of a uniform tariff classification of products in determining "like products," the Appellate Body said that a sufficiently detailed tariff classification

could be a helpful sign of product similarity. However, the Appellate Body cautioned that tariff bindings that include a wide range of products may not be a reliable criterion for confirming or determining product "likeness" under Article III:2, and, therefore, the determinations on which tariff bindings provide significant guidance as to the identification of "like products" need to be made on a case-by-case basis.[51] In all other respects, the Appellate Body affirmed the findings and the legal conclusions of the panel with respect to "like products."

According to the panel, the appropriate test to define whether two products are like or directly competitive or substitutable is the marketplace. In the panel's view, although the decisive criterion in determining whether two products are directly competitive or substitutable is whether they have common end-uses, *inter alia*, as shown by the elasticity of substitution in a market where competition exists, commonality of end-uses is a necessary but not sufficient criterion to define "likeness." According to the panel, the term "like products" suggests that for two products to fall under this category they must share, apart from commonality of end-uses, essentially the same physical characteristics. By applying the above-mentioned criteria for examination of the products at issue, the panel concluded that vodka and shochu were like products because both vodka and shochu shared most physical characteristics and except for the media used for filtration there was virtual identity in the definition of the two products.[52] The panel, however, did not conclude that shochu and other alcoholic beverages in dispute were "like products" because substantial noticeable differences in physical characteristics existed between the remaining alcoholic beverages in dispute and shochu that would disqualify them from being regarded as like products.[53]

According to the Appellate Body, after the determination of the "likeness" of the products at issue, the only remaining step to determine the conformity of an internal tax measure with Article III:2, first sentence, is the examination of whether the taxes on imported products are "in excess of" those on like domestic products. If so, then the Member that has imposed the tax is not in compliance with Article III. In the view of the Appellate Body, even the smallest amount of "excess" is too much because the prohibition of discriminatory taxes in Article III:2, first sentence, is not conditional on a "trade effects test" nor is it qualified by a *de minimis* standard. Accordingly, the Appellate Body agreed with the panel's legal reasoning and with its conclusions[54] on this aspect of the interpretation and application of Article III:2, first sentence.

As noted earlier, even if the imported and domestic products are not "like products," they may still be "directly competitive or substitutable products." In such a case a three-step test is required to determine the validity of an internal tax measure under the principle of National Treatment. The first step is the determination of "directly competitive or substitutable products." In the Appellate Body's view, as with "like products," the determination of the appropriate range of "directly competitive or substitutable products" under Article III:2, second sentence, must be made on a case-by-case basis, taking into account all the relevant facts. The Appellate Body agreed with the panel's approach in this regard. The panel had emphasized the need to look not only at such matters as physical characteristics, common end-uses, and tariff classifications, but also at the "market place" because the important issues in this regard were factors like market strategies and the responsiveness of consumers to the various products offered in the market. In the view of the Appellate Body, it was not inappropriate to look at

competition in the relevant markets as one among a number of means of identifying the broader category of products that might be described as "directly competitive or substitutable." The Appellate Body also agreed with the panel's view that the decisive criterion in order to determine whether two products are directly competitive or substitutable is whether they have common end-uses, *inter alia*, as shown by elasticity of substitution in the relevant markets.[55] It thus found the panel's legal analysis of whether the products are "directly competitive or substitutable products" to be correct.

According to the Appellate Body, after the determination of directly competitive or substitutable products, the next step in the test is whether these products are similarly taxed. In its view, the phrase "not similarly taxed" does not mean the same thing as the phrase "in excess of" in Article III:2, first sentence, because if "in excess of" and "not similarly taxed" were construed to mean one and the same thing, then "like products" and "directly competitive or substitutable products" would also mean one and the same thing.[56] According to the Appellate Body, there may be an amount of excess taxation that may well be more of a burden on imported products than on domestic "directly competitive or substitutable products" but may not be enough to justify a conclusion that such products are "not similarly taxed" for the purposes of Article III:2, second sentence. It agreed with the panel that the amount of differential taxation must be more than *de minimis* to be deemed "not similarly taxed"; and whether any particular differential amount of taxation is *de minimis* or not must be determined on a case-by-case basis. Thus, to be "not similarly taxed," the tax burden on imported products must be heavier than on "directly competitive or substitutable" domestic products, and that burden must be more than *de minimis* in any given case. The Appellate Body also agreed with the legal reasoning applied by the panel in determining whether "directly competitive or substitutable" imported and domestic products were "not similarly taxed." However, the Appellate Body also found that the panel erred in blurring the distinction between that issue and the issue of whether the tax measure in question was applied "so as to afford protection," which, in the Appellate Body's view, were entirely different issues that must be addressed separately. The panel had concluded that the following indicators, *inter alia*, were relevant in determining whether the products in dispute were similarly taxed in Japan: tax per liter of product, tax per degree of alcohol, *ad valorem* taxation, and the tax/price ratio.[57]

According to the Appellate Body, if "directly competitive or substitutable products" are "similarly taxed," then there is neither need nor justification under Article III:2, second sentence, for inquiring further as to whether the tax has been applied "so as to afford protection." However, if such products are "not similarly taxed," a further inquiry must necessarily be made. In its view, this third inquiry must determine whether "directly competitive or substitutable products" are "not similarly taxed" in a way that affords protection. The Appellate Body argued that this was not an issue of intent and that it was not necessary for a panel to sort through the reasons given by legislators and regulators in imposing the measure in dispute. In its view, if the measure is applied to imported or domestic products so as to afford protection to domestic production, then it is irrelevant that protectionism was not an intended purpose. What is relevant is how the particular tax measure in question is applied. In this respect, the Appellate Body found the approach followed in the 1987 *Japan Alcohol* case in the examination of the issue of "so as to afford protection" persuasive and concluded that an examination

of whether dissimilar taxation has been applied so as to afford protection requires a comprehensive and objective analysis of the structure and application of the measure in question on domestic as compared to imported products. In its view, it is possible to examine objectively the underlying criteria used in a particular tax measure, its structure, and its overall application to ascertain whether it is applied in a way that affords protection to domestic products. The Appellate Body argued that even if the aim of a measure may not be easily ascertained, its protective application can most often be discerned from "the design, the architecture, and the revealing structure of a measure," and the very magnitude of the dissimilar taxation in a particular case may be evidence of such a protective application. However, there may be other factors to be considered as well. Therefore, full consideration should be given to all the relevant facts and circumstances in any given case, and in every case, a careful, objective analysis, must be undertaken of each and all such facts and circumstances in order to determine "the existence of protective taxation."

Despite arguing for a separate inquiry on the issue of "so as to afford protection to domestic production" and the rejection of the panel's conclusion of equating the determination of dissimilar taxation with the separate requirement of demonstrating that the tax measure affords protection to domestic production, the Appellate Body, however, agreed with the panel's conclusion that the very fact that the substantially dissimilar taxation was applied to directly competitive or substitutable imported and domestic products was enough in this case to conclude that the tax measure in dispute was applied "so as to afford protection."[58]

The tests outlined by the Appellate Body in *Japan – Taxes on Alcoholic Beverages* have been followed by the panels and the Appellate Body in other cases involving internal taxes as well as other regulatory measures. The practical difficulties in applying these tests[59] were evident in *Canada – Certain Measures Concerning Periodicals*[60] case, where the panel found that imported split-run periodicals and domestic non-split-run periodicals were "like" products under Article III:2, first sentence, whereas the Appellate Body found that such periodicals were not "like" products, but were "directly competitive or substitutable" products under Article III:2, second sentence. In this case, one of the issues in dispute was Part V.1 of the Canadian Excise Tax Act which imposed an 80 percent excise tax on advertising in each split-run edition of a periodical.[61] The United States claimed that these provisions of the Excise Tax Act were in violation of the National Treatment obligation enshrined in Article III:2 of GATT because they discriminated between two "like" products, domestic non-split-run periodicals and imported split-run periodicals. The panel concluded that Part V.1 of the Canadian Excise Tax Act was inconsistent with Article III:2, first sentence, of GATT 1994. Canada and the United States both appealed.[62] Although the Appellate Body agreed with the application by the panel of the two-step "like" products test established by the Appellate Body in *Japan – Taxes on Alcoholic Beverages* case in examining the consistency of a tax measure with Article III:2, first sentence, it did not agree with the panel's conclusion that imported split-run periodicals and domestic non-split-run periodicals were "like" products.[63] According to the Appellate Body, the panel did not base its findings on the exhibits and evidence before it[64] and that the panel's conclusions lacked proper legal reasoning based on adequate factual analysis. However, the Appellate Body did not determine whether the imported split-run periodicals and

domestic non-split-run periodicals were "like" products.[65] Instead, it proceeded to examine the consistency of the tax measure with Article III:2, second sentence.[66] It said that if the answer to the question of whether imported and domestic products are "like" products is negative, there is then a need to examine the consistency of the measure with the second sentence of Article III:2.

Applying the three-step test established by the Appellate Body in *Japan – Taxes on Alcoholic Beverages*, the Appellate Body found that the imported split-run periodicals and Canadian non-split-run periodicals were "directly competitive or substitutable" products in so far as they were part of the same segment of the Canadian market for periodicals. This conclusion was based on a study carried out by a Canadian economist, a Task Force Report submitted by Canada, and statements made by the Minister of Canadian Heritage and Canadian officials, all of which had acknowledged the substitutability of, and considerable competition between, imported split-run periodicals and domestic non-split-run periodicals in the Canadian market.[67] Similarly, the Appellate Body concluded that "directly competitive or substitutable" imported split-run periodicals and domestic non-split-run periodicals were "not similarly taxed" by the Canadian Excise Tax Act because it taxed split-run editions of periodicals in an amount equivalent to 80 percent of the value of all advertisements, whereas domestic non-split-run periodicals were not subject to the tax, and the amount of the taxation was far above the *de minimis* threshold specified by the Appellate Body in *Japan – Taxes on Alcoholic Beverages*. Finally, it concluded that the design and structure of Canadian excise tax was clearly "to afford protection to the production of Canadian periodicals." This conclusion was based on the magnitude of dissimilar taxation,[68] the evidence of protective purpose from several statements of the Government of Canada's explicit policy objectives in introducing the measure, and the demonstrated actual protective effect of the measure.[69] Thus, the Appellate Body concluded that Part V.1 of the Canadian Excise Tax Act was inconsistent with Canada's obligations under Article III:2, second sentence, of the GATT 1994.

The *Canada – Certain Measures Concerning Periodicals* case suggests[70] that it is difficult to prove the "likeness" of products under Article III:2, first sentence, unless there is a substantial identity in the physical characteristics and perfect substitutability of the products in question. However, this difficulty has not affected the outcome of the examination of whether a tax measure is inconsistent with the principle of National Treatment because of the availability of a further examination under Article III:2, second sentence, which covers "directly competitive or substitutable" products, and there is not a single decided case under the WTO where a tax measure has been determined to be consistent with Article III:2, second sentence, once the products in question have been found to be "directly competitive or substitutable." This is evident from the Appellate Body decisions in *Korea – Taxes on Alcoholic Beverages* and *Chile – Taxes on Alcoholic Beverages*. In both of these cases, the Appellate Body affirmed the findings of the respective panels which had found both the Korean and Chilean alcohol taxation systems to be inconsistent with the National Treatment principle set forth in Article III:2.

In *Korea – Taxes on Alcoholic Beverages*,[71] the Appellate Body upheld the findings of the panel that soju (diluted and distilled), whiskies, brandies, cognac, rum, gin, tequila, liquors, and admixtures were directly competitive or substitutable products. It also

upheld the panel's conclusion that Korea had taxed the imported products in a dissimilar manner and that the dissimilar taxation was applied so as to afford protection to domestic production. Both the panel and Appellate Body applied the three-step test established by the Appellate Body in *Japan – Taxes on Alcoholic Beverages*. In the panel's view, an assessment of whether there is a direct competitive relationship between two products or groups of products requires evidence that consumers consider or could consider the two products or groups of products as alternative ways of satisfying a particular need or taste, and the determination of whether domestic and imported products are directly competitive or substitutable requires evidence of a direct competitive relationship between the products, including comparisons of their physical characteristics, end-uses, channels of distribution, and prices. According to the panel the focus should not be exclusively on the quantitative extent of the competitive overlap. Quantitative analyses and studies of cross-price elasticity of demand are helpful and relevant, but should not be considered necessary and are not exclusive or even decisive in nature because protectionist government policies can distort the competitive relationship between products, causing the quantitative extent of the competitive relationship to be understated. According to the panel, the assessment of competition has a temporal dimension. Therefore, panels should examine evidence of trends and changes in consumption patterns and make an assessment as to whether such trends and patterns lead to the conclusion that the products in question are either directly competitive now or can reasonably be expected to become directly competitive in the near future.[72] According to the Appellate Body, the context of the competitive relationship between imported and domestic products is necessarily the marketplace since this is the forum where consumers choose between different products. In its view, the word "substitutable" indicates that the requisite relationship may exist between products that are not, at a given moment, considered by consumers to be substitutes but which are, nonetheless, capable of being substituted for one another. Products are competitive or substitutable when they are interchangeable or if they offer alternative ways of satisfying a particular need or taste.

With regard to the issue of whether or not the Korean liquor taxes were applied so as to afford protection to domestic products, the panel found that the Korean tax law had very large differences in levels of taxation,[73] and that the very magnitude of dissimilar taxation itself was sufficient to conclude that the taxes at issue were applied so as to accord protection to Korean domestic liquors. In addition to the very large levels of tax differentials, the panel also found the structure of the Liquor Tax Law itself to be discriminatory.[74] The Appellate Body upheld the panel's conclusions and rejected the arguments of Korea that there were no such protective effects in the market because of the large pre tax price difference between diluted soju and imported alcoholic beverages. According to the Appellate Body, this argument did not change the pattern of application of the contested measures because Article III is not concerned with trade volumes and therefore it was not incumbent on the complainant to prove that tax measures were capable of producing any particular trade effect.

The panel and Appellate Body in *Chile – Taxes on Alcoholic Beverages*[75] followed the same approach as followed in *Japan – Taxes on Alcoholic Beverages* and in *Korea – Taxes on Alcoholic Beverages* in determining the issues of whether or not pisco, whisky, and other spirits are directly competitive or substitutable, whether or not the domestic alcoholic beverages and directly competitive or substitutable imported alcoholic

beverages were similarly taxed and, if there were dissimilar taxes above the *de minimis* level, whether or not dissimilar taxes were applied so as to afford protection to domestic products. With regard to the first issue, the panel looked at evidence of the relationship between the products, including comparisons of their end-uses, physical characteristics, channels of distribution, and prices, and found that pisco and other spirits were directly competitive or substitutable products.[76] According to the panel, products do not have to be substitutable for all purposes at all times to be considered competitive and it is sufficient that they may be substituted for some purposes at some times by some consumers.

In evaluating substitutability in end-uses, the panel also found it useful to consider consumer theory, which, according to the panel, holds that "goods are, in the eyes of consumers, never really perceived as commodities that are in themselves direct objects of utility; rather, it is the properties or characteristics of the goods from which utility is derived that are the relevant considerations. It is these characteristics or attributes that yield satisfaction and not the goods as such. Goods may share a common characteristic but may have other characteristics that are qualitatively different, or they may have the same characteristics but in quantitatively different combinations. Substitution possibilities arise because of these shared characteristics." According to the panel, one hypothetical example in this regard is that of butter, milk, and margarine. "Butter and milk are both dairy products and they share important characteristics that margarine does not have. However, butter and margarine each have combinations of characteristics that make them good substitutes as complements for bread, which is not the case with milk. The characteristics of butter and margarine can be expressed as physical properties such as spreadability, taste, color, and consistency. These physical characteristics combine to render both products good substitutes as bread complements. The latter represents the end-use of the commodities as determined by their combination of characteristics derived from certain physical characteristics." In the panel's view, the same type of reasoning can be applied to the substitutability of pisco and other spirits such as whisky, brandy, cognac, etc.[77]

Similarly, the panel also found that its conclusion on competition or substitutability between pisco and other spirits was consistent with the production and marketing decisions of the pisco producers who desired to convey an image of pisco as a drink that competes with the best imported distilled spirits. According to the panel, when a product is being marketed in ways that suggest that it is in competition with up market imported distilled spirits, this is evidence of at least potential competition with those imports. Likewise, the panel also found that the Chilean Central Preventive Commission, in deciding on a merger between two major pisco producers, had stated that pisco faced major competition from other alcoholic beverages, such as wine, beer, and whisky, and that these were alternative products which consumers of alcoholic beverages could choose to drink in the market for alcoholic beverages. Thus, the panel concluded that the totality of the evidence presented supported a finding that the imported distilled spirits and pisco were directly competitive or substitutable.

With respect to the issue of whether or not the imported distilled spirits and directly competitive or substitutable pisco were similarly taxed, the panel found that both the Transitional and New Systems applied dissimilar taxes to these alcoholic beverages. According to the panel, the level of difference in taxation between whisky and pisco under the Old System was greater than *de minimis* because whisky was taxed at more

than twice the rate of pisco and even if the Transitional System would make the difference in taxation somewhat narrower in the following years, the tax difference would still remain more than *de minimis*, and even with respect to other spirits, the tax difference of five percentage points *ad valorem* was greater than *de minimis*. The New System, which assessed taxes on an *ad valorem* basis that varied according to alcohol content, also applied dissimilar taxes greater than *de minimis* to directly competitive or substitutable imported and domestic products because the difference in taxation between the top (47 percent) and bottom (27 percent) levels of *ad valorem* rates of taxation of distilled alcoholic beverages was clearly more than *de minimis* and was so by a very large margin. Similarly, the difference of four percentage points between the various levels of alcohol content also constituted a greater than *de minimis* level of dissimilar taxation. According to the panel, the question of dissimilar taxation does not involve judgments about the objectives of the laws or regulations involved, nor does it involve an assessment of who benefits from the tax system. It is sufficient for this step of the analysis to find that some of the imports are being taxed dissimilarly from some of the domestic production and the difference is more than *de minimis*. In the view of the panel, a tax system based on taxing value is generally considered not to be applying dissimilar taxation if done on a purely *ad valorem* basis (i.e., a single *ad valorem* rate applied uniformly to all products). However, the New Chilean System was not strictly an *ad valorem* system because it applied *ad valorem* rates that varied not just by value but also by alcohol content.

On the issue of whether or not the Chilean alcohol taxes were applied so as to afford protection to domestic products, the panel concluded that both the Transitional and New Systems applied dissimilar taxes to domestic products and directly competitive or substitutable imported products so as to afford protection to Chilean domestic products. According to the panel, the central issue in this regard is the design, architecture, and revealing structure of the tax measure and an important question in the determination of protective application is who receives the benefit of the dissimilar taxation. Since the Transitional System assessed tax rates by type of spirits and the lowest tax rate was on pisco, which under Chilean law was exclusively a domestic product, it was clear that the beneficiary of the tax structure was the domestic industry. Similarly, the largest category of imports was whisky, which was taxed at a rate of 53 percent (at its least discriminatory level) compared to pisco's 25 percent, and pisco accounted for almost 75 percent of domestic production of distilled spirits. The panel rejected the argument of Chile that the Transitional System did not have any protective application as it actually reduced the tax rate on whisky. The panel held that the fact that the Transitional System lessened the protective effect did not vitiate the conclusion that, even at its least discriminatory, it was a system that did and would afford protection to domestic production.

The New System also afforded protection to domestic production because the structure of the New System applied its lowest rate at the level of alcohol content of the large majority of domestic production and its highest rate at the level of the overwhelming majority of imports;[78] the large magnitude of the differentials were applied over a short range of physical difference (27 percent for 35 degrees versus 47 percent for 39 degrees of alcohol content); the interaction of the New System with the Chilean regulation which required most of the imports to remain at the highest tax level without losing

their generic name and changing their physical characteristics;[79] and the lack of any connection between the stated objectives and the results of the measures.[80] The panel rejected the arguments made by Chile to support the non protective application of the tax measure that any producer, whether foreign or domestic, could produce spirits at lower levels and benefit from the tax structure; that there was a great deal of spirits produced in the EEC at 35 degrees of alcohol or less which could easily be exported to Chile and enjoy a lower level of taxation; that there was more absolute production of domestic spirits in Chile at the higher levels of taxation than there were imports; that there was not even *de facto* discrimination because the imported product could easily be diluted to take advantage of the lower available tax rates; and that if protection was the goal, Chile could have raised tariffs which were currently at 11 percent, but bound at 25 percent. The panel found these factors either irrelevant or as demonstrating that there would not be equal competitive conditions unless the foreign producers make certain important changes in their products, changes not justified by any exception or rule of the WTO Agreements. The Appellate Body upheld the findings of the panel in *Chile – Taxes on Alcoholic Beverages*.[81]

In *Indonesia – Certain Measures Affecting the Automobile Industry*,[82] Japan, the United States, and the European Communities complained that the sales tax benefits provided under the February 1996, 1993, and June 1996 Indonesian car programs violated Article III:2 of GATT. Indonesia argued that the sales tax and luxury tax benefits provided to its national car companies were subsidies and were consistent with the Agreement on Subsidies and Countervailing Measures (SCM Agreement) even if such tax benefits were inconsistent with Article III:2. It argued that there was a conflict between Article III:2 and the SCM Agreement in that the obligations contained in Article III:2 and the SCM Agreement were mutually exclusive because the SCM Agreement "explicitly authorized" Members to provide subsidies that were prohibited by Article III:2. However, the panel rejected the arguments made by Indonesia and concluded that whether or not the SCM Agreement was considered generally to authorize Members to provide actionable subsidies so long as they did not cause adverse effects to the interests of another Member, the SCM Agreement clearly did not authorize Members to impose discriminatory product taxes. The SCM Agreement and Article III:2 were not mutually exclusive because it was possible for Indonesia to respect its obligations under the SCM Agreement without violating Article III:2 since Article III:2 was concerned with discriminatory product taxation, rather than the provision of subsidies as such.

Once the panel concluded that Article III:2 applied in regard to the Indonesian tax benefit scheme for national car producers, it followed the approach adopted by the Appellate Body in *Japan – Taxes on Alcoholic Beverages*, to test the validity of the Indonesian tax benefit scheme under Article III:2 of GATT. The panel concluded:

> Under the Indonesian car programmes the distinction between the products for tax purposes is based on such factors as the nationality of the producer or the origin of the parts and components contained in the product. An imported vehicle alike in all aspects relevant to a likeness determination would be taxed at a higher rate simply because of its origin or lack of sufficient local content. Such an origin-based distinction in respect of internal taxes suffices in itself to violate Article III:2 without the need to demonstrate the existence of actually traded like products.

3.2 Regulatory Measures and National Treatment

Article III:4 of GATT, along with the general principle in Article III:1, sets out the National Treatment obligations with regard to various internal regulations other than internal tax measures. The significant difference between the National Treatment obligations set forth in Article III:4 and Article III:2 is that Article III:4 in its wording only applies to "like" products and not to "directly competitive or substitutable" products. Similarly, the required treatment of imported products is "no less favorable than that accorded to 'like' domestic products" and there is no reference to Article III:1 in Article III:4. This means, according to the interpretation of Article III adopted by the Appellate Body in *Japan – Taxes on Alcoholic Beverages* and followed by panels and the Appellate Body in other cases, such as *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef* (hereinafter *Korea – Measures on Beef*), and *European Communities – Regime for the Importation, Sale, and Distribution of Bananas* (hereinafter *EEC – Bananas*), that no separate inquiry as to whether a regulatory measure has been applied "so as to afford protection to domestic production" is required to determine the consistency of a regulatory measure with National Treatment obligations set out in Article III:4. A determination that the imported and domestic products in question are "like" and that the regulatory measure in dispute provides less favorable treatment to imported products than that accorded to like domestic products, is sufficient to establish a violation of Article III of the GATT.

The first case under the WTO dispute settlement system where an issue of the violation of Article III:4 was raised is *United States – Standards for Reformulated and Conventional Gasoline* (hereinafter US–Gasoline).[83] The regulatory measure in question in this case was explicitly discriminatory[84] and not facially neutral because the gasoline product standard at issue in the case set a different and potentially more onerous standard for foreign suppliers, and the United States' main defense of the gasoline standard was the exceptions to general GATT obligations set out in Article XX. However, the panel[85] in this case made rulings with regard to the steps in the inquiry required to determine whether a non tax regulatory measure is consistent with the National Treatment obligations set out in Article III:4. According to the panel, complainants under Article III:4 are required to show the existence of: (a) a law, regulation, or requirement affecting the internal sale, offering for sale, purchase, transportation, distribution, or use of an imported product; and (b) treatment accorded in respect of the law, regulation, or requirement that is less favorable to the imported product than to the "like" product of national origin. The panel concluded that the establishment of these two issues was sufficient to determine the inconsistency of a regulatory measure with Article III:4, and there is no need to establish the issue of "so as to afford protection to domestic production" as set forth in Article III:1 because the provision of Article III:1 is a general one and the provision of Article III:4 is more specific.

The panel began its examination in this regard by the determination of 'like' products. To determine the likeness of products, the panel followed the criteria suggested by the 1970 GATT Working Party Report on Border Tax Adjustments and considered that the criteria applied in the 1987 *Japan Alcohol* case in the examination under Article III:2, first sentence of internal tax measures were also applicable to the examination of like

products under Article III:4. The panel found that the domestic and imported gasoline were "like" products because the chemically identical imported and domestic gasoline by definition had exactly the same physical characteristics, end-uses, tariff classification, and were perfectly substitutable.

In order to determine whether the treatment provided to the imported products was less favorable than that accorded to like domestic products, the panel followed the conclusions of the GATT panel in *US – Section 337*, which had said that the words "treatment no less favorable" in Article III:4 call for effective equality of opportunities for imported products in respect of laws, regulations, and requirements affecting the internal sale, offering for sale, purchase, transportation, distribution, or use of products. The panel concluded that the US gasoline regulations treated the imported gasoline less favorably than the domestic gasoline because, under the baseline establishment methods provided in the regulations, the imported gasoline was effectively prevented from benefiting from as favorable sales conditions as were afforded to domestic gasoline. Relying on the conclusions in *US – Section 337*, the panel also concluded that, under Article III:4, less favorable treatment of particular imported products in some instances could not be balanced by more favorable treatment of other imported products in other instances.

The approach taken by the panel in *US – Gasoline* in determining the inconsistency of a non tax regulatory measure with Article III:4 was not fully followed by the panel in *EEC – Bananas*.[86] In this case, the panel, citing the Appellate Body's decision in *Japan – Taxes on Alcoholic Beverages* and relying on the GATT panel decision in *US – Section 337*, also examined the issue of whether the regulatory measure in question was applied so as to afford protection to domestic production, in addition to the two issues examined by the panel in *US – Gasoline* case. However, the Appellate Body in *EEC – Bananas* rejected this part of the panel's approach, stating that the panel misinterpreted its conclusion in *Japan – Taxes on Alcoholic Beverages* and that "a determination of whether there has been a violation of Article III:4 does not require a separate consideration of whether a measure 'affords protection to domestic production.'"

The first WTO case on the National Treatment principle involving the issue of facially neutral non tax regulatory measures was the *Japan – Measures Affecting Consumer Photographic Film and Paper*[87] (hereinafter *Japan – Film*). The panel in this case followed the same approach as that established by the panel in US – *Gasoline* to determine whether the various Japanese distribution measures violated the National Treatment principle contained in Article III:4. In this case, the United States complained that eight different decisions, reports, guidelines, etc., of various Japanese authorities accorded less favorable treatment to imported film and paper than to like domestic film and paper in the Japanese market. In response, Japan argued that the United States failed to show how the alleged measures applied less favorable treatment to imported film and paper. The panel concluded that none of the alleged Japanese distribution measures violated Article III:4. Relying on the Appellate Body's decision in *Japan – Taxes on Alcoholic Beverages*, the panel held that the standard of effective equality of competitive conditions on the internal market for imported products in relation to domestic products is the standard of national treatment that is required, not only with regard to Article III generally, but also more particularly with regard to the no less favorable treatment standard in Article III:4. According to the panel, the United States failed to show that any of the measures cited by the United States discriminated against imported

products either in terms of *de jure* discrimination or in terms of *de facto* discrimination. The United States had argued that the measures in question were directed at promoting vertical integration in the photographic materials distribution system with a view to impeding market access for foreign products. However, the panel rejected the US arguments, stating that the Japanese measures were formally neutral as to the origin of products and their application did not have a disparate impact on imported film or paper. The basis of the US claim was the existence of a single brand wholesale distribution system in the Japanese market for film and photographic papers, which according to the United States, impeded market access for foreign products. The panel found that the United States could not establish a causal link or a meaningful nexus between the challenged measures and the market structure because the contested market structure existed even prior to the introduction of the measures in question. It also found that a single brand wholesale distribution system was the common market structure – indeed the norm – in most major national film markets, including the US market. The panel argued that it was unclear why the same economic forces acting to promote single brand wholesale distribution in the United States would not also exist in Japan.

Thus, the panel in *Japan – Film* established that a causal link or meaningful nexus between the challenged measures and the competitive conditions in the market must be shown by the complainant in order to prove a violation of Article III:4. However, what constitutes a regulatory measure (i.e., a law, regulation, or requirement affecting the internal sale, offering for sale, purchase, transportation, distribution, or use of an imported product) subject to the purview of Article III:4 may itself be a contentious issue. In *Japan – Film*, the panel's interpretation of the terms "laws, regulations or requirements" in Article III:4 was not entirely clear. Although it argued that a literal reading of the words "all laws, regulations or requirements" in Article III:4 could suggest that they may have a narrower scope than the word "measure" in Article XXIII:1(b) in the context of nullification and impairment, the panel assumed for the purposes of this case that the terms "laws, regulations or requirements" in Article III:4 should be interpreted as having a meaning similar to the term "measures" in Article XXIII:1(b), and found that only three measures met the definition of "laws, regulations or requirements" within the meaning of Article III:4. However, the panel also assumed that the remaining five contested measures were also "laws, regulations or requirements" for the sake of completeness of its analysis in examining whether less favorable treatment was accorded to imported products.

The issue as to the meaning of "laws, regulations or requirements" in Article III:4 also arose in *Canada – Certain Measures Affecting the Automotive Industry* (hereinafter Canada – Automotive).[88] The issues in dispute relating to Article III:4 in this case were Canadian measures, which accorded to certain motor vehicle manufacturers established in Canada, the right to import motor vehicles with an exemption from the generally applicable customs duty. In order to qualify for the exemption, an eligible manufacturer's local production of motor vehicles (including in certain cases the production of parts) must have achieved a minimum amount of Canadian value added (CVA) and its local production must have maintained a minimum production-to-sales ratio with respect to its sales of motor vehicles in Canada. Japan and the European Communities claimed that the CVA and production-to-sales ratio contained in various government Orders as well as the commitment with regard to the CVA expressed by

certain manufacturers in Letters of Undertaking to the government were "requirements" affecting the internal sale, offering for sale, purchase, transportation, distribution, or use of an imported product within the meaning of Article III:4 and these requirements accorded less favorable treatment to imported parts, materials, and non permanent equipment for use in the production of motor vehicles. Canada argued that these measures did not affect the "internal sale, …, or use" of imported products because they did not in law or in fact require the use of domestic products and therefore played no role in the parts sourcing decisions of manufacturers.

The panel concluded that Article III:4 applies not only to mandatory measures, but also to conditions that an enterprise accepts to receive an advantage, including cases where the advantage is in the form of a benefit with respect to the conditions of importation of a product. The fact that compliance with the CVA requirements is not mandatory but a condition that must be met in order to obtain an advantage consisting of the right to import certain products duty-free does not preclude application of Article III:4. Similarly, the panel found that the word "affecting" in Article III:4 of the GATT has been interpreted to cover not only laws and regulations which directly govern the conditions of sale or purchase but also any law or regulation that might adversely modify the conditions of competition between domestic and imported products. The panel concluded that the CVA requirements in government Orders must be regarded as measures which "affect" the "internal sale, …, or use" of imported products because a measure which provides that an advantage can be obtained by using domestic products, but not by using imported products, has an impact on the conditions of competition between domestic and imported products and thus affects the "internal sale, …, or use" of imported products, even if the measure allows for other means to obtain the advantage, such as the use of domestic services rather than products. Similarly, the panel claimed that neither legal enforceability nor the existence of a link between a private action and an advantage conferred by a government was a necessary condition in order for an action by a private party to constitute a "requirement." According to the panel, a determination of whether a private action amounts to a "requirement" under Article III:4 must necessarily rest on a finding that there is a nexus between that action and the action of a government such that the government must be held responsible for that action. The panel concluded that the commitments expressed in the Letters of Undertakings were "requirements" within the meaning of Article III:4.[89]

On the issue of whether the CVA requirements accorded less favorable treatment to imported products, the panel rejected the argument of Canada that these requirements did not in practice accord less favorable treatment to imported products as the CVA levels were so low that they could easily be met on the basis of labor alone. The panel found that the CVA requirements accorded less favorable treatment within the meaning of Article III:4 to imported parts, materials, and non permanent equipment than to like domestic products because, by conferring an advantage on the use of domestic products, they adversely affected the equality of competitive opportunities of imported products in relation to like domestic products. For the same reasons, the panel concluded that the commitments contained in the Letters of Undertaking also accorded less favorable treatment to imported products.

Despite distinctions noted in some cases between *de jure* discrimination caused by explicitly discriminatory regulatory measures and *de facto* discrimination caused

by facially neutral regulatory measures, the WTO jurisprudence has not developed separate tests to determine the validity of such measures under Article III of the GATT. Although in the context of Article III:2, first sentence, WTO panels and Appellate Body have declared any internal tax measure that imposes even slightly different tax rates on imported products compared to like domestic products to be inconsistent with the National Treatment principle on the very basis of such origin-specific differentiation, origin-specific regulatory measures are not *per se* inconsistent with the National Treatment principle. The Appellate Body in *Korea – Measures on Beef* rejected the panel's conclusion that "any regulatory distinction that is based exclusively on criteria relating to the nationality or the origin of the products is incompatible with Article III and this conclusion can be reached even in the absence of any imports, confirming that there is no need to demonstrate the actual and specific trade effects of a measure for it to be found in violation of Article III."[90] The Appellate Body stated that a formal difference in treatment between imported and like domestic products is neither necessary nor sufficient to show a violation of Article III:4. In its view, whether or not imported products are treated "less favorably" than like domestic products should be assessed instead by examining whether a measure modifies the conditions of competition in the relevant market to the detriment of imported products.

In this case,[91] both the panel and Appellate Body concluded that Article III:4 is violated if the complainant demonstrates: (a) that imported and domestic products are "like;" (b) that the measure at issue is either a law, regulation, or requirement affecting their internal sale, offering for sale, purchase, transportation, distribution, or use; and (c) that the measure provides to imported products treatment less favorable than that accorded to domestic products. As there was no dispute at both the panel and Appellate Body levels on the "likeness" of domestic and imported beef and the measure at issue being a law or regulation within the meaning of Article III:4, both the panel and Appellate Body only examined whether or not the dual retail system for beef in the Korean market provided less favorable treatment to imported beef. Although both the panel and Appellate Body reached the same conclusion that the retail system for beef in the Korean market provided less favorable treatment to imported beef, they based their conclusion on different reasons.

Korea had appealed against the finding of the panel, which concluded that the dual retail system applied by Korea to imported and domestic beef accorded less favorable treatment to imported beef and thus was inconsistent with Article III:4. In addition to the above-mentioned reason based on origin of products that was rejected by the Appellate Body, the finding of the panel was also based on its assessment of how the dual retail system modified the conditions of competition between imported and like domestic beef in the Korean market. The panel gave several reasons for why it believed that the dual retail system altered the conditions of competition in the Korean market in favor of domestic beef: first, the dual retail system would "limit the possibility for consumers to compare imported and domestic products," and thereby "reduce opportunities for imported products to compete directly with domestic products"; second, under the dual retail system, "the only way an imported product can get on the shelves is if the retailer agrees to substitute it, not only for one but for all existing like domestic products," and this disadvantage would be more serious when the market share of imports (as is the case with imported beef) is small; third, the dual retail system,

by excluding imported beef from "the vast majority of sales outlets" limited the potential market opportunities for imported beef, and this would apply particularly to products "consumed on a daily basis," like beef, where consumers may not be willing to "shop around"; fourth, the dual retail system imposed more costs on the imported product, since the domestic product would tend to continue to be sold from existing retail stores, whereas imported beef would require new stores to be established; fifth, the dual retail system "encourages the perception that imported and domestic beef are different, when they are in fact like products belonging to the same market," which gave a competitive advantage to domestic beef "based on criteria not related to the products themselves"; and sixth, the dual retail system "facilitates the maintenance of a price differential" to the advantage of domestic beef. On appeal, Korea argued that dual retail system does not on its face violate Article III:4, since there was "perfect regulatory symmetry" in the separation of imported and domestic beef at the retail level, and there was "no regulatory barrier" which prevented traders from converting from one type of retail store to another. Korea also argued that the dual retail system did not deny consumers the possibility to make comparisons, and it neither added to the costs of, nor sheltered high prices for, domestic beef.

Relying on the GATT panel decision in *US – Section 337* and its decision in *Japan – Taxes on Alcoholic Beverages*, the Appellate Body stated that "treatment no less favorable" means according conditions of competition no less favorable to the imported product than to the like domestic product and it implies that a measure according formally different treatment to imported products does not *per se* violate Article III:4. The Appellate Body did not agree with the panel that the limitation on the ability of consumers to compare visually two products at the point of sale necessarily reduced the opportunity for the imported product to compete "directly" or on "an equal footing" with the domestic product, nor did it agree that the alleged encouragement provided by the dual retail system to the perception of consumers that imported and domestic beef were "different" necessarily implied a competitive advantage for domestic beef. In its view, although the Korean dual retail system formally separated the selling of imported beef and domestic beef by the requirement of two distinct retail distribution systems, such formal separation, in and of itself, did not necessarily compel the conclusion that the treatment thus accorded to imported beef was less favorable than that accorded to domestic beef. According to the Appellate Body, to determine whether the treatment accorded to imported beef was less favorable than that accorded to domestic beef, it was necessary to inquire into whether or not the Korean dual retail system for beef modified the conditions of competition in the Korean beef market to the disadvantage of the imported product. After examining the beef market structure in Korea, the Appellate Body concluded that the introduction of the dual retail system resulted in the imposition of a drastic reduction of commercial opportunities for imported beef to reach, and hence to generate sales to, the same consumers served by the traditional retail channels for domestic beef.[92] Although it agreed that the dramatic reduction in number of retail outlets for imported beef followed from the decisions of individual retailers who could choose freely to sell the domestic product or the imported product, it found that the legal necessity of making a choice was imposed by the government measure itself and the reduction of access to normal retail channels was, in legal contemplation, the effect of that measure. The Appellate Body concluded, therefore, that the Korean Government's

measure was responsible for the resulting establishment of competitive conditions less favorable for the imported product than for the domestic product, and the fact that the WTO-consistent quota for beef was fully utilized did not detract from the lack of equality of competitive conditions entailed by the dual retail system.[93]

The next significant case involving a facially neutral regulatory measure that was claimed to violate the National Treatment principle in Article III:4 was the *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products* (hereinafter *EEC – Asbestos*).[94] In this case, the issue was the general ban imposed by a Decree of the French Government on the manufacture, processing, sale, import, placing on the domestic market, and transfer under any title whatsoever of all varieties of asbestos fibers. However, on an exceptional and temporary basis, the ban was not to apply to certain existing materials, products, or devices containing chrysotile fiber when, to perform an equivalent function, no substitute for that fiber was available which posed a lesser health risk. Canada complained, *inter alia*, that the French Decree violated the National Treatment principle of Article III:4 of the GATT by banning the marketing of chrysotile fibers and chrysotile-cement products because chrysotile fibers and chrysotile-cement products were "like" polyvinyl alcohol (PVA), cellulose, and glass fibers within the meaning of Article III:4 and by prohibiting chrysotile fibers and chrysotile-cement products, the EEC was favoring its national industry of PVA, cellulose, and glass fibers (hereinafter "PCG fibers") and fibro-cement products containing these fibers.

The panel, following the steps established by WTO panels and the Appellate Body in past cases, began its inquiry by examining whether or not the chrysotile fibers were "like" PCG fibers, and whether or not cement-based products containing chrysotile asbestos fibers were "like" cement-based products containing one of the PCG fibers. To define the "likeness" of products, the panel followed the same approach as that taken by the panel in *US – Gasoline* which had applied the criteria suggested by the Appellate Body in *Japan – Taxes on Alcoholic Beverages* for the purposes of determining "like" products in the context of Article III:2, first sentence. The panel specifically noted the observations made by the Appellate Body in *Japan – Taxes on Alcoholic Beverages* that the term "like" products should be examined on a case-by-case basis, which would inevitably involve a degree of judgment. Despite the acknowledgment that the structure of chrysotile fibers is unique by nature and that none of the substitute fibers has the same structure, either in terms of its form, diameter, length, or potential to release particles that possess certain characteristics, and that they do not have the same chemical composition or in purely physical terms the same nature or quality, the panel still found that chrysotile fibers were "like" PCG fibers. The basis of the panel's finding was that, for many industrial uses, PCG fibers have the same applications as chrysotiles. The panel rejected the narrow definition of "like product" as applied in other WTO cases, arguing that consideration of only the physical structure, chemical composition, and properties of products in the examination of "likeness" of products would exclude many products from being "like" even if they had a similar use. The panel also claimed that the context for the application of Article III:4 is not a scientific classification exercise but is to provide market access for products, and in the context of market access, it is not necessary for domestic products to possess all the physical similarities and properties of the imported products in order to be "like" products. In the view of the panel, the fact

that chrysotile fibers and PCG fibers have certain identical or at least similar end-uses in cement products was sufficient to consider them as "like" products even if in other circumstances their end-uses may be different.

The panel also rejected as irrelevant the argument of the EEC that chrysotile fibers are a widely recognized carcinogen and pose serious threats to human health. The panel claimed that the risk of a product to human or animal health has never been used as a factor of comparison by panels entrusted with applying the concept of "likeness" within the meaning of Article III, and introducing a criterion as to the health risks of a product into the analysis of "likeness" within the meaning of Article III would largely nullify the effect of Article XX(b) which specifically covers the protection of human health and life (under which the panel went on to uphold the measures in question). The panel also did not consider the criterion of consumers' tastes and habits, stating that the products concerned were not everyday consumer goods. Similarly, the panel disregarded the difference in tariff classification of the products in dispute in the Harmonized System stating that the difference in tariff classification was not a decisive criterion in this case.

On the issue of whether or not the EEC measure provided less favorable treatment to imported products than that accorded to like domestic products, the panel concluded that the terms of the EEC measure themselves established less favorable treatment for asbestos and products containing asbestos as compared to PCG fibers and products containing PCG fibers because the measure imposed a ban on asbestos fibers, and did not place an identical ban on PCG fibers and fibro-cement products containing PCG fibers.[95] Thus, the panel found that the EEC measure in regard to asbestos products was inconsistent with Article III:4.

It is evident from the panel's decision in *EEC – Asbestos* that the determination of the issue of whether or not a regulatory measure is inconsistent with the principle of National Treatment depends very much on whether or not the imported product and its domestic comparator are "like" each other. As stated by the Appellate Body in *EEC – Asbestos*, the determination of the "likeness" of two products in the context of Article III:4 rests on how a panel decides three issues: first, which characteristics or qualities are important in assessing the "likeness" of products since most products have many qualities and characteristics, ranging from physical properties such as composition, size, shape, texture, and possibly taste and smell, to the end-uses and applications of the product; second, the degree or extent to which products must share qualities or characteristics in order to be "like" products since products may share only very few characteristics or qualities or they may share many; and third, from whose perspectives "likeness" should be judged because ultimate consumers may have a view about the likeness of two products which may be very different from that of the inventors, producers, or regulators of those products.[96] The Appellate Body attempted to resolve these issues.

The Appellate Body first noted that the appeal from the panel's decision provided it with its first occasion to examine the meaning of the term "like products" in Article III:4. Although it observed that the term "like product" appears in the first sentence of Article III:2 and in Article III:4 in the context of National Treatment principle, and both of these provisions constitute specific expressions of the overarching general principle of National Treatment set forth in Article III:1, it concluded that the term "like products" in Article III:4 should not be construed as narrowly as in the context of

Article III:2. The reason for a different approach to interpreting the same words in the context of the National Treatment principle is, according to the Appellate Body, that Article III:2 contains two separate obligations in two sentences covering "like" products and "directly competitive or substitutable" products respectively and there is a need to interpret these two sentences in a harmonious manner in order to give meaning to both sentences of Article III:2, whereas Article III:4 contains a single obligation that applies solely to "like" products and the harmony required to be attributed to the two sentences of Article III:2 need not and cannot be replicated in interpreting Article III:4. In the view of the Appellate Body, a determination of "likeness" under Article III:4 is fundamentally a determination about the nature and extent of a competitive relationship between and among products, even if there is a spectrum of degrees of competitiveness or substitutability of products in the market place and it is difficult, in the abstract, to indicate precisely where on this spectrum the word "like" in Article III:4 falls. The Appellate Body concluded that the product scope of Article III:4, although broader than the first sentence of Article III:2, is certainly not broader than the combined product scope of the two sentences of Article III:2.[97] After having so defined the scope of "like" products in Article III:4, the Appellate Body proceeded to outline a framework for analyzing the "likeness" of particular products in a particular case. It found that past GATT panels as well as WTO panels and the Appellate Body have developed and followed an approach consisting of four general criteria in order to determine the "likeness" of products. These four criteria are: (i) the properties, nature, and quality of the products; (ii) the end-uses of the products; (iii) consumers' tastes and habits; and (iv) the tariff classification of the products. However, the Appellate Body claimed that these criteria are neither a treaty-mandated nor a closed list of criteria that should determine the legal characterization of products, but are simply tools to assist in the task of sorting and examining the relevant evidence in a particular case. According to the Appellate Body, all the pertinent evidence needs to be examined in each case and the kind of evidence to be examined in assessing the "likeness" of products depends upon the particular products and the legal provision at issue.

The Appellate Body rejected the approach taken by the panel in *EEC – Asbestos* to determine the "likeness" of chrysotile fibers with PCG fibers, and reversed the determination that chrysotile fibers were "like" PCG fibers and cement-based products containing chrysotile asbestos fibers and cement-based products containing PCG fibers were "like products." It concluded that the panel should have examined the evidence relating to each of the four criteria and then weighed all of this evidence, along with any other relevant evidence, in making an overall determination of whether the products at issue could be characterized as "like," and that it was inappropriate for the panel to express a conclusion after examining only one of the four criteria (end-uses). According to the Appellate Body, physical properties of products deserve a separate examination which should not be confused with the examination of end-uses, and although not decisive, the extent to which products share common physical properties may be a useful indicator of "likeness" because the physical properties of a product may influence how the product can be used, consumer attitudes about the product, and tariff classification. The evidence relating to the health risks associated with a product may be pertinent to an examination of "likeness" under Article III:4, but need not be examined under

a separate criterion and can be evaluated under the criteria of physical properties and of consumers' tastes and habits.

After reversing the panel's conclusion in regard to the "likeness" of chrysotile fibers with PCG fibers and cement-based products containing chrysotile asbestos fibers with cement-based products containing PCG fibers, the Appellate Body proceeded to its own examination of "likeness" of the products at issue on the basis of the evidence available in the panel's report. It first examined the physical properties of chrysotile fibers and PCG fibers and noted the panel's conclusion that these fibers are physically very different. Then, it emphasized the fact, which was treated as irrelevant although acknowledged by the panel in examining "likeness" – that chrysotile fibers have been recognized internationally as a known carcinogen because of the particular combination of their molecular structure, chemical composition, and fibrillation capacity. The Appellate Body also noted the evidence that PCG fibers are not classified by the World Health Organization at the same level of risk as chrysotile and the experts consulted by the panel also confirmed that current scientific evidence indicates that PCG fibers do not present the same risk to health as chrysotile fibers. It then concluded that when the evidence relating to properties indicates that the products in question are physically different, then "in order to overcome the indication that products are not like, a high burden is imposed on a complainant to establish that, despite the pronounced physical differences, there is a competitive relationship between the products such that, all the evidence, taken together, demonstrates that the products are 'like' under Article III:4." The Appellate Body found that the complainant had not satisfied its burden because the end-uses of chrysotile fibers and PCG fibers were the same for only a small number of applications, no evidence was submitted on consumers' tastes and habits[98] and chrysotile fibers and PCG fibers have different tariff classifications.

Applying the same criteria as in the examination of the "likeness" of chrysotile fibers with PCG fibers, the Appellate Body also examined whether cement-based products containing chrysotile asbestos fibers are "like" cement-based products containing PCG fibers and found that these products were not "like" products. It specifically rejected the contention of Canada that evidence on consumers' tastes and habits concerning cement-based products was irrelevant. According to the Appellate Body, it was of particular importance under Article III to examine evidence relating to competitive relationships in the market place, and it was likely that the presence of a known carcinogen in one of the products would have an influence on both intermediate and final consumers' tastes and habits regarding that product. In the view of the Appellate Body, it might be that, although cement-based products containing chrysotile fibers were capable of performing the same functions as other cement-based products, consumers were, to a greater or lesser extent, unwilling to use products containing chrysotile fibers because of the health risks associated with them. However, the Appellate Body considered it as only speculation and did not make any determination on this issue because of lack of evidence. In its view, a determination on the "likeness" of the cement-based products could not be made, under Article III:4, in the absence of an examination of evidence on consumers' tastes and habits.

On the basis of these findings, the Appellate Body concluded that, as Canada had not demonstrated that chrysotile asbestos fibers were "like" PCG fibers or that

cement-based products containing chrysotile asbestos fibers were "like" cement-based products containing PCG fibers, it did not succeed in establishing that the EEC measure at issue was inconsistent with Article III:4 of the GATT. The Appellate Body, however, also observed that there is a second element that must be established before a regulatory measure can be held to be inconsistent with Article III:4. Even if two products are "like," the complainant must still establish that the measure accords to the group of "like" imported products "less favorable treatment" than it accords to the group of "like" domestic products. In the view of the Appellate Body, the term "less favorable treatment" expresses the general principle set out in Article III:1, that internal regulations should not be applied "so as to afford protection to domestic production." It said that if there is "less favorable treatment" of the group of "like" imported products, there is, conversely, "protection" of the group of "like" domestic products. Nevertheless, the Appellate Body also said that distinctions may be drawn between products which have been found to be "like," without, for this reason alone, according to the group of "like" imported products "less favorable treatment" than that accorded to the group of "like" domestic products.[99]

It is notable that one Member of the Appellate Body in *EEC – Asbestos* expressed a separate opinion about the approach to be taken in order to determine the "likeness" of two products. He took the view that, considering the nature and quantum of the scientific evidence showing the carcinogenicity of chrysotile asbestos fibers, there was ample basis for a definitive characterization of such fibers as not "like" PCG fibers, and that definitive characterization might and should be made even in the absence of evidence concerning the other two criteria of end-uses and consumers' tastes and habits.[100] He also cautioned that the necessity or appropriateness of adopting a "fundamentally economic" interpretation of the "likeness" of products under Article III:4 was not free from substantial doubt, and in future contexts, the line between a "fundamentally" and "exclusively" economic view of "like products" under Article III:4 might well prove very difficult, as a practical matter, to identify. However, he did not offer any suggestion as to the appropriate approach to the interpretation of the "likeness" of products under Article III:4, but rather he reserved his opinion on this matter.

After the *EEC – Asbestos* case, two other cases, which involve issues pertaining to Article III:4 of the GATT, have been decided by the Dispute Settlement Body (DSB) of the WTO. However, the tests applied by the panel and Appellate Body to examine the consistency or inconsistency of the measure in question with Article III:4 in *United States – Tax Treatment for "Foreign Sales Corporations" – Recourse to Article 21.5 of the DSU by the European Communities*[101] [hereinafter *US – FSC (Article 21.5)*] and the panel in *India – Measures Affecting the Automotive Sector*[102] (hereinafter *India – Automotive)* are similar to those followed by the panel and the Appellate Body in *Canada – Automotive, Korea – Beef*, and *EEC – Asbestos*.

In the *US – FSC (Article 21.5)*,[103] the panel cited the rulings of the panel and Appellate Body in *Canada – Automotive* and *EEC – Asbestos* in respect of the meaning of "like products" and "less favorable treatment," and viewed the principal purpose of the "like product" inquiry under Article III:4 as ascertaining whether any formal differentiation in treatment between an imported and a domestic product could be based upon the fact that the products are different (not like) rather than on the origin of the products involved. According to the panel, when a regulatory measure of general

application makes a distinction between imported and domestic products solely and explicitly on the basis of origin of such products, and applies horizontally to all possible products that can be used for the production of goods that might eventually be a recipient of the benefit accorded by the said regulatory measure, then there is no need to demonstrate the existence of actually traded like products in order to establish a violation of Article III:4. On the issue of when a regulatory measure at issue is considered as one "affecting" the internal sale or use of the products concerned, the panel said, relying on the rulings in *EEC – Bananas* and *Canada – Automotive*, that the ordinary meaning of the term "affecting" implies a measure that has "an effect on," thereby indicating a broad scope of application. The panel also noted that the term "affecting" in Article III:4 has been interpreted to cover not only laws and regulations that directly govern the conditions of sale or purchase but also any law or regulation that might adversely modify the conditions of competition between domestic and imported products. The panel then considered that a measure pursuant to which the use of domestic, but not imported, products contributes to obtaining an advantage has an impact on the conditions of competition between domestic and imported products and thus "affects" the internal "use" of imported products, even if the measure allows for other means to obtain the advantage, such as the use of domestic inputs other than products.

On the issue of "less favorable treatment," the panel recalled the previous rulings in *Canada – Automotive* and *Korea – Beef* that Article III:4 of the GATT is an obligation addressed to governments requiring that they ensure equality of competitive opportunities to domestic and like imported products, and it does not require a demonstration of trade effects, nor proof that the sourcing decisions of private firms have actually been impacted by the regulatory measure in question. The panel also stated that any distinction that is based exclusively on criteria relating to the nationality or origin of the product would not necessarily be incompatible with Article III. To be incompatible with the provisions of Article III:4, a measure must accord treatment to imported products that is "less favorable than" that accorded to like domestic products. According to the panel, when an advantage is conferred upon the use of domestic products that is not conferred upon the use of imported products, it constitutes a formal differentiation of treatment between imported and like domestic products, which, in the view of the panel, affords less favorable treatment to imported products than to like domestic products because by conferring an advantage upon the use of domestic products but not upon the use of imported products, it adversely affects the equality of competitive opportunities of imported products in relation to like domestic products. The Appellate Body upheld the rulings of the panel in this case.[104]

In *India – Automotive* case,[105] the issues were similar to those in *Canada – Automotive* and *US – FSC (Article 21.5)*. Therefore, the panel followed the same approach and gave similar reasons in determining the inconsistency of the measure in question with Article III:4. On the issue of the meaning of the term "requirement" under Article III:4, the panel concluded that a binding enforceable condition falls squarely within the ordinary meaning of the word "requirement," in particular as "a condition which must be complied with." According to the panel, the enforceability of the measure in itself, independently of the means actually used or not to enforce it, is a sufficient basis for a measure to constitute a requirement under Article III:4. Similarly, with respect to the meaning of the term "affecting," the panel said that this term goes

beyond laws and regulations which directly govern the conditions of sale or purchase to cover also any laws or regulations which might adversely modify the conditions of competition between domestic and imported products. On the issue of "less favorable treatment" to imported products, the panel said that in determining whether imported products are treated less favorably than domestic products, it (the panel) is obliged to examine whether the contested regulatory measure modifies the conditions of competition in the relevant market to the detriment of imported products. According to the panel, any requirement that provides an incentive to purchase and use domestic or local products and hence creates a disincentive to use like imported products modifies the conditions of competition between the domestic and imported products in the relevant market within the meaning of Article III:4 because such a requirement creates a situation where imported products cannot compete on an equal footing with domestic products.

4 Critique of the GATT/WTO Case Law on National Treatment: A Competition Policy Perspective

According to Article 31 of the Vienna Convention on the Law of Treaties,"a treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its objective purpose." As the case law amply demonstrates and as the Appellate Body has acknowledged in several cases, the interpretation and application of key terms in Article III involve an unavoidable element of judgment because the terms have no self-evident "ordinary meaning," thus suggesting the importance of an interpretation of these terms that is consonant with the purpose of the Article. Obviously, merely looking at, touching, feeling, smelling, or decomposing two products is unlikely to reveal whether they are like products in any legal or policy-relevant sense, suggesting the need for something more purposive than a "smell test." The purpose of Article III is set out in Article III.1 which provides that internal taxes and other internal charges, laws, regulations, and requirements affecting the internal sale, offering for sale, purchase, transportation, distribution, or use of good … should not be applied to imported or domestic products *so as to afford protection to domestic production*. Panels and the Appellate Body have not been consistent in their adoption of a purposive interpretation of key elements of Article III:2 and Article III:4. Decisions lurch inconsistently from a literalist, context-independent approach to a regulatory "aims-and-effect" (or regulatory purpose) approach to an economic approach. For example, with respect to Article III:2, first sentence, the Appellate Body in the *Japanese Alcohol* case held that the reference to "like domestic products" in Article III:2, first sentence, should be interpreted more narrowly than the reference to "directly competitive or substitutable products" in Article III:2, second sentence, by virtue of the Ad Note to Article III:2; and in interpreting and applying Article III:2, first sentence, "the purpose of Article III set out in Article III:1 "… so as to afford protection to domestic production" was inapplicable because Article III:2, second sentence, expressly incorporates the principles set forth in Article III:1 while Article III:2, first sentence, does not. Despite the convoluted efforts of the Appellate Body to maintain a distinction between the interpretation and application of Article III:2, first sentence, and Article III:2, second

sentence, the distinctions in practice seem of little significance in that internal taxes that cannot be successfully challenged under Article III:2, first sentence, because of its narrower scope, can almost always be successfully challenged under Article III:2, second sentence, where these taxes have a protectionist application or effect.

The rejection by the Appellate Body and panel in the *Japanese Alcohol* case of the "aims-and-effects" test in interpreting both sentences of Article III:2 has somewhat more ambiguous implications. With respect to Article III:2, second sentence, which expressly incorporates the principle set forth in Article III:1 "… so as to afford protection to domestic production," the Appellate Body in the *Japanese Alcohol* case seems to have developed an objective purpose test, or perhaps a potential effects test, for establishing protection to domestic production ("the design, the architecture, and the revealing structure of a measure"). What is less clear is whether, despite a potential protectionist effect from an internal tax, a respondent country is able to adduce evidence of a non protectionist policy purpose, for example, in order to reduce the social effects of excessive alcohol consumption, taxing high alcohol content beverages at proportionately higher rates than lower alcohol content beverages. The willingness of the panel in the *Chilean Alcohol* case to evaluate evidence pertaining to whether the structure of the measures in question could be rationally justified in these terms suggests that such an argument may be open under Article III:2, second sentence.

The interpretation of Article III:2, first and second sentences, has clearly infected and confused GATT/WTO interpretations of Article III:4 which refers only to "like products" and not "directly competitive or substitutable products," and does not explicitly incorporate any reference to the principle set forth in Article III:1. However, the Appellate Body in *Asbestos* adopted an essentially economic test of "like products" in Article III:4 that focuses on the *competitive relationship* between imported and domestic products, which largely subsumes the interpretation of like product in Article III:4 into the concept of "directly competitive or substitutable products" in Article III:2, second sentence. Moreover, in interpreting the phrase "treatment no less favorable" in Article III:4, the Appellate Body in *Asbestos* (as it had in the *Korean Beef* case) held that a mere finding of likeness between two products does not oblige the regulating country to treat them identically in regulation. The complainant must also demonstrate that the differences in regulation amount to "less favorable" treatment as between domestic and imported like products, each taken as a group. In so stating, the Appellate Body recalled the anti protectionist purpose of Article III and suggested that "less favorable treatment" is equivalent to protectionism, although this is in puzzling conflict with its statement in *EEC – Bananas*, overruling the panel in this respect, that "a determination of whether there has been a violation of Article III:4 does not require a separate consideration of whether a measure affords protection to domestic production." Assuming, following the Appellate Body's decisions in *Asbestos* and *Korean Beef*, that "treatment no less favorable" under Article III:4 means that the measure in question may not have an objective protective purpose or effect, a similar ambiguity remains to that under Article III:2, second sentence, as to whether despite such an effect, a respondent country is free to adduce arguments or evidence that the primary motivation or justification for the measure in question was some non protectionist policy.

Some commentators have been critical of GATT/WTO decisions, such as that of the panel in the first *Tuna–Dolphin* case, that Article III:4 in referring to like products refers

only to products as such, and not to differences in production or processing methods (PPM's) between domestic and imported products.[106] While they take some solace from the Appellate Body's decision in *Asbestos* that differences in the health characteristics of domestic and imported products may render these two classes of products "unlike products," the emphasis by the Appellate Body in *Asbestos* on an economic test of "like products" (i.e., whether they are viewed as directly competitive or substitutable by users) would seem largely to rule out differential treatment based on differences in PPMs that users, that is intermediate or final consumers, do not, for the most part, regard as salient in choosing between products in the market place, for example, perhaps, for many consumers, tuna caught by dolphin-friendly or dolphin-unfriendly fishing methods. Focusing on whether consumer preferences or choices in importing countries may be sensitive to differences in PPMs raises a number of problems. First, consumers may not be well informed about differences in PPMs and hence ignore them in market place choices, requiring a somewhat speculative inquiry as to whether these differences would be salient to consumers if they were well-informed of them. Second, even consumers who are well-informed of these differences may not, in many cases, alter their behavior significantly because of collective action problems – the perceived futility of forgoing purchase of an otherwise better or cheaper imported product if other consumers will seek to free ride on the self-sacrificing decisions of others, leading to a non cooperative Prisoner's Dilemma outcome. Third, how many consumers, even if well informed and uninfluenced by collective action problems, would need to change their consumption patterns in the light of differences in PPMs before one could conclude that imported and domestic products were "unlike products?"[107]

However, abandoning altogether an economically based test of "like products" and allowing panels and the Appellate Body to deem products "unlike," or allowing importing countries unilaterally to deem products to be "unlike," on account of production or process differences, raises several formidable problems. First, this would centrally contradict the whole theory of comparative advantage, because outside of the category of raw, fungible commodity exports, the comparative advantage that most imports will enjoy will turn on differences in PPMs, or inputs more generally. Second, while the exceptions listed in Article XX of the GATT are now over 50 years old and have not been revised since the inception of the GATT, and are arguably out of date in not incorporating exceptions for example, for core international labor standards, at least some subset of universal human rights, or more clearly identifying consumer protection and the environment as legitimate exceptions or objectives for trade restricting measures,[108] additions to or refinements of the exceptions listed in Article XX are clearly, in many respects, a matter of extremely high politics amongst many WTO members. To view Article III of the GATT as providing a mandate to panels and the Appellate Body to invent on their own a set of normative justifications for apparently less favorable treatment for imported products relative to directly competitive domestic products is likely to severely strain the internal legitimacy of the WTO dispute settlement process relative to its political institutions. An expansive reading by panels or the Appellate Body of the "public morals" exception in Article XX(a) and a less stringent interpretation of the "necessity" requirement in a number of the Article XX exceptions and of the non discrimination and non protectionist conditions in the chapeau to Article XX may be able to accommodate some of these concerns. However, similar issues will arise as to

the institutional legitimacy of panels and the Appellate Body engaging in expansive judicial lawmaking in these respects.

Arguments that actual regulatory purpose (the aims-and-effects test) should be the controlling determinant of the definition of "like products" – that domestic measures motivated by nonprotectionist rationales should be exempt from Article III[109] – raise these difficulties in the clearest form. Apart from difficulties in ascertaining the actual intent of legislators or regulators (given the frequency of mixed motivations, as reflected in domestic Baptist–Bootlegger coalitions favoring measures restrictive of imports), what regulatory purposes count as legitimate and what as protectionist? Leveling the competitive playing field, preventing a race to the bottom, unilateral sanctions against foreign countries' violations of international labor standards, or international human rights? These are profoundly normative and highly contested rationales for the invocation of trade sanctions and remitting to panels and the Appellate Body the responsibility for determining their legitimacy and scope would entail a gross usurpation of the political authority of the WTO membership. Moreover, it is likely to promote highly inconsistent decision making. For example, a violation of Article XI (quantitative restrictions) will require justification under the strictures of Article XX. But a presumptive violation of Article III will be excused if the measure in question has a non protectionist regulatory purpose that renders domestic and imported products "unlike" in the view of the panels and the Appellate Body in interpreting and applying Article III. Assuming that exculpatory, non protectionist regulatory purposes are interpreted more broadly under Article III than Article XX, why not convert an import ban vulnerable to challenge under Article XI into internal tax discrimination against imports challengeable only under Article III?

A refined version of the aims-and-effects approach in the form of a means–ends examination to determine the consistency or inconsistency of a facially neutral internal tax or regulation with the provisions of National Treatment under Article III has recently been advanced by Gaetan Verhoosel,[110] relying in part on the panel's willingness to examine the relationship between the means and ends of a regulation under Article III in *Chile – Taxes on Alcoholic Beverages* case. Verhoosel supports the domestic "regulatory autonomy" of a state, which encompasses the state's autonomy as regards the policy objectives it chooses to pursue and as regards the means by which it chooses to pursue such policy objectives, and argues that WTO law should not interfere with either aspect of this autonomy, except to the extent that the free choice of a policy objective amounts to overt protectionism and the free choice of regulatory means amounts to covert protectionism. The author advocates a so-called "integrated necessity test" which integrates the tests to be carried out under Article III and Article XX into a single test and argues that National Treatment should be understood to require a necessity test in the context of facially neutral domestic regulation to determine *de facto* discrimination under Article III. The author claims that *de facto* discrimination can only be revealed in an objective manner by "engaging in an analysis as to whether a particular regulatory instrument (1) specifically and adversely affects imported products as compared with their like domestic counterparts, and (2) is necessary to achieve a purported legitimate policy goal, or, alternatively, whether other, less restrictive, regulatory means are available."

However, this so-called "integrated necessity test" raises similar problems to those we identified above regarding the consideration of regulatory purposes in determining

Article III issues[111]. In particular, unconstrained expansion of the legitimate policy exceptions list of Article XX by adjudication and interpretation under Article III would entail a gross usurpation of the political authority of the WTO membership and risk serious challenges to the credibility and legitimacy of the dispute settlement mechanism.

In contrast, we favor a much more restricted and less normatively contestable approach to the interpretation and application of Article III that is more consistent with its original purpose. First, we favor an economic approach to the definition of "like products" that is motivated by the need to establish, as a positive threshold issue, a significant competitive relationship between imported and domestic products before any issue of protectionist intent or effect can factually arise under Article III. However, while we endorse the view of the Appellate Body in *Asbestos* that the "like product" requirement should focus on the competitive relationship between domestic and imported products, we believe that this inquiry could be much more tightly structured than is currently the case and could usefully borrow from decades of learning and case law in developed antitrust jurisdictions around the world in defining relevant product markets.[112] One such approach, reflected in the United States and Canadian Merger Enforcement Guidelines,[113] is the hypothetical monopolist test where one asks of the hypothesized sole producer of a given class of product whether that producer could sustain a significant (e.g. 5 percent) increase in the price of that product for a non transitory period (e.g. 1 year) without inducing a sufficient number of consumers to substitute to other products so as to render such a price increase unprofitable. If, from empirical evidence, such substitution is likely to occur then the products to which consumers substitute are deemed like products, and then a similar question is posed with respect to this expanded class of products and so on, until a class of products is defined which satisfies this test. A similar test is typically adopted in defining the scope of the relevant geographic market: could a hypothetical monopolist with respect to a relevant class of products raise price by a significant degree for a non transitory period without inducing sufficient consumers to switch to more distant suppliers of these products so as to render such price increase unprofitable? While these tests are by no means easy to operationalize in practice and can be supplemented and refined in various ways, if the key to the "like product" definition in Article III is the *competitive relationship* between imports and domestic products, as the Appellate Body stated in *Asbestos*, then it is inexplicable that international trade tribunals would not draw much more extensively on the framework of analysis and body of experience that has developed in competition law in many jurisdictions around the world in addressing precisely the same question.

On an economic approach to the definition of "like products," one would take revealed consumer preferences, even if informationally flawed, as given. For example, in *Asbestos*, if evidence suggests that intermediate and final consumers treat asbestos and other fibers as close substitutes in a significant number of end-uses, we would treat them as like products. Indeed, it is difficult to understand the French ban on the sale of asbestos or cement-based products containing asbestos, or the Canadian complaint about the ban, unless intermediate and final consumers in fact treated them as close substitutes in a significant number of end-uses. If the argument is that well-informed consumers would not do so, this may provide a justification for health and safety measures under Article XX(b), but this is where this issue should be resolved, not Article III.

Assuming that a competitive relationship is established between imports and domestic products, the question that then arises is how Article III:2, second sentence, with its reference to "… so as to afford protection to domestic production" in paragraph one, and the phrase "treatment no less favorable" in Article III:4, should be interpreted. Consistent with the economic approach that we favor in interpreting "like product," we favor the adoption of an economic approach to these two requirements similar to that espoused by the GATT panel in *US – Section 337* and the Appellate Body in several decisions reviewed above, which focuses on the preservation of effective equality of competitive opportunities. As to whether a challenged measure disturbs or undermines, actually or potentially, effective equality of competitive opportunities is, as with the like product inquiry, largely a positive empirical and predictive inquiry which, while not requiring identical treatment of imports and domestic products, as the GATT panel recognized in *US – Section 337*, and as the Appellate Body recognized in *Korean Beef* and *Asbestos*, nevertheless must entail treatment that does not impose on foreign exporters competitive burdens that are not imposed on domestic producers. Drawing on the antitrust literature on raising rivals' costs,[114] the inquiry would condemn domestic measures that raise the marginal costs (including opportunity costs) of foreign rivals relative to the marginal costs (including opportunity costs) of domestic rivals so as substantially to lessen competition in the domestic market by rendering it likely that domestic firms will be able profitably to raise prices significantly for a non transitory period without attracting sufficient entry to render such a strategy unprofitable. This strategy would violate the anti protectionist rationale of Article III.

Some critics of the product–process distinction have seized on references by the Appellate Body in *Asbestos* to the need to interpret Article III:4 in light of the anti protectionist purpose of the Article, and recognition by the Appellate Body that this does not require identical treatment of like products, as providing an opening for taking account of differences in PPMs not in defining "like products" but rather in determining under Article III:2, second sentence, whether an internal tax measure has been applied so as to afford protection to domestic production and under Article III:4 whether treatment no less favorable than that accorded to like products of national origin has been accorded to imports.[115] However, there is nothing to date in the Appellate Body's willingness to recognize that "no less favorable treatment" of imports does not require identical treatment to suggest that it had anything in mind other than an examination of the impact of the measure in question on effective equality of competitive opportunities of imports and domestic products – for example, in its analysis of the dual retail store system for beef in the *Korean Beef* case.[116] Its ruling on the issue of "less favorable treatment" under Article III:4 in the recent case-*US – FSC (Article 21.5)*, discussed above, which was decided almost 9 months after its decision on *EEC – Asbestos* suggests that the pattern of examination on the issue of "less favorable treatment" of imported products under Article III:4 has not changed since *Korean Beef.* Moreover, simply switching the focus of attempts to incorporate these profoundly normative considerations from the definition of "like product" to the interpretation of "treatment no less favorable" raises all the same objections to attempting such an exercise with respect to the interpretation of "like product."

This said, it is important to acknowledge that domestic measures that are adopted that simultaneously impose similar constraints on domestic and foreign like products for environmental, labor, or other reasons, including domestic measures that are enforced

at the border in the case of imported products (thus engaging the Ad Note to Article III) may indeed satisfy the test of preserving equality of competitive opportunities. Thus, for example, to take the *Tuna–Dolphin* case, if domestic measures banning the catching of tuna in dolphin-unfriendly ways are simultaneously imposed on domestic and foreign suppliers of tuna to the US market, and entail similar compliance costs for both (including opportunity costs), the competitive equilibrium between domestic and foreign like products is not disturbed. Similarly, to take the *Shrimp–Turtles* case[117], if measures banning the catching of shrimp in turtle-unfriendly ways are simultaneously applied to domestic and foreign suppliers of shrimp to the US market, and entail similar compliance and opportunity tests, foreign goods are not being treated less favorably than domestic like goods. Similarly, again in the case of a simultaneous ban on the sale of domestic and foreign products made with child labor if such a measure imposes similar costs on foreign and domestic producers of like products.

We acknowledge that even this interpretation of Article III may leave open the ambiguity noted above as to whether a measure which is adopted by an importing country for a non protectionist, non anticompetitive purpose or policy objective, but which has incidental and disproportionately adverse impacts on competing imports might nevertheless be viewed as not violating of Article III:2, second sentence, "… so as to afford protection to domestic production" or as not violating the "no less favorable treatment" requirement of Article III:4. Illustrative of the potential problem is the Canada–US FTA panel decision in *Lobsters*[118] where the application of domestic US minimum size requirements to Canadian lobsters was challenged by Canada. The purpose of the measure was to conserve lobster stock by ensuring that young lobster would not be taken before they could breed. However, because Canada has colder waters, its mature lobsters are generally of a smaller size. The United States argued that the application of its size requirement to Canadian lobster was necessary to the enforcement of the requirement with respect to American lobster. Since lobsters do not carry passports, it would be costly and impractical to determine whether a lobster was Canadian or American once it had entered the stream of commerce. This difficulty was obviated by applying the size requirement to all lobster in the market. Another example might be the imposition by the Canadian government of bilingual labeling requirements (English and French) on most products sold in Canada. Because of the pre existing bilingual capacity of many Canadian producers, this requirement may impose additional costs on many foreign producers of like products.

In our ideal world we would prefer to remit such measures for evaluation under the (perhaps revised and expanded) Article XX Exceptions List. However, even as Article XX is presently framed, most measures that have provoked controversy under Article III with respect to the scope and application of the "aims and effects" and "regulatory purpose" approach to its interpretation are potentially justifiable under Article XX. For example, both the panel and the Appellate Body in *Asbestos* agreed that the measures in question were justified under Article XX(b) (measures necessary to protect human, animal or plant life, or health). While the *Tuna–Dolphin* panels found violations of Article III and Article XX, the Appellate Body decision in *Shrimp–Turtles* (including its compliance decision) makes it clear that such measures may now be justifiable under Article XX(b) or (g) (conservation of exhaustible natural resources). Similarly, the measures in dispute in *Lobsters* would now likely be justifiable under

Article XX(g). Thus, in contrast to recent arguments by Donald Regan[119] for treating a competition test as a necessary but not sufficient test for violation of Article III, requiring in addition proof of actual protectionist regulatory purpose, we favor treating a competition test as a necessary and sufficient test for violations of Article III, remitting justifications for offending measures based on non protectionist regulatory purposes to Article XX, subject to its "necessity," "arbitrary and unjustifiable discrimination," and "disguised restriction on international trade" constraints. A major virtue of this division of labor between Articles III and XX, beyond preserving the integrity of the limited exceptions list in Article XX, is that it appropriately allocates the burden of proof between the complainant and respondent. The complainant must prove a disparate impact on foreign producers of a domestic measure adopted by the respondent in accordance with the requirements of the competition test that we have proposed. Assuming that this burden is discharged under Article III, the burden of proof then shifts to the respondent to justify a measure within the strictures of Article XX. This burden properly is assigned to the respondent because it turns on information uniquely within its possession, in contrast to placing the burden on the complainant of proving protectionist intent on the part of the respondent under Article III.

However, given the limited scope of Article XX, it may be the case that some domestic regulatory measures that are adopted for non protectionist reasons outside the scope of Article XX may have incidentally adverse impacts on foreign exporters. Hence, it may be argued that such measures will, for the foreseeable future, have to be evaluated under Article III. Thus, Hudec has argued that it may be impossible to suppress altogether some form of the "aims-and-effects" test, at least in a limited form[120] in interpreting Article III. We consider that there will be few such cases. However, in such cases, beyond the adoption of a *de minimis* requirement of the kind that the Appellate Body has already adopted with respect to Article III:2, second sentence, and the "substantial lessening of competition" test that we propose in interpreting "no less favorable treatment" in Article III:4, one might, adopting a more constrained variant of Verhoosel's proposals, address the problem of incidentally disproportionate impacts on imports by borrowing the test adopted by a GATT panel in the *Herring and Salmon Processing* case[121] in interpreting Article XX(g) of the GATT, and ask whether the measure in question was "primarily aimed at" a non protectionist, non anticompetitive domestic policy objective and not "primarily aimed at" restricting imports; whether the adverse impact on imports was necessarily incidental to attainment of this domestic policy objective; and whether any less trade restrictive means was reasonably available to achieve that objective (borrowing this test from the WTO/TBT and SPS Agreements).[122] However, in order to avoid compromising the internal and perhaps external political legitimacy of panels and the Appellate Body,[123] this exception would need to be narrowly defined; respondent countries should clearly bear the burden of proof of satisfying its requirements (in large part because of their superior access to the relevant information on intended policy objectives), and it should be viewed as an interim, second-best option to reforming Article XX of the GATT where all justificatory social rationales for trade-restricting domestic measures should ideally reside. In particular, measures primarily aimed at banning or restricting imports, for example, on account of PPMs such as labor or environmental standards in the country of origin could not be justified under these qualifications of Article III, but would require

justification (if at all) under Article XX. Even so, we remain to be convinced that the risks of recognizing such qualifications to a pure competition-oriented approach to the interpretation of Article III, even on an interim basis, justify the benefits.

Notes

1. See William Smith Culbertson, *International Economic Policies: A Survey of the Economics of Diplomacy* (D. Appleton Company, 1925), p. 24.
2. See Michael M. Hart, *The Mercantilist's Lament: National Treatment and Modern Trade Negotiations*, 21(6) *Journal of World Trade Law* 38 (1987).
3. See Georg Schwarzenberger, *The Most-Favored-Nation Treatment in British State Practice*, XXII *The British Yearbook of International Law* 97 (1945).
4. See G. Erler, *Grundprobleme des Internationalen Wirtschaftsrechts* (Gottingen, 1956), p. 47, cited by Pieter VerLoren van Themaat in *The Changing Structure of International Economic Law* (Hague: Martinus Nijhoff, 1981), p. 19.
5. See Georg Schwarzenberger, *The Principles and Standards of International Economic Law*, in *RECUEIL DES COURS*, vol. I, 1966, Academie De Droit International, Hague, p. 80; W. McClure, *German–American Commercial Relations*, 19 *American Journal of International Law* 692 (1925), and William Smith Culbertson, *supra* note 1.
6. Ibid.; see also Gerard Curzon, *Multilateral Commercial Diplomacy* (London: Michael Joseph, 1965), p. 15, and Pieter VerLoren van Themaat, *supra* note 4.
7. See Article 2 of the Paris Convention for the Protection of Industrial Property of March 20, 1883; and Article 5 of the Berne Convention for the Protection of Literary and Artistic Works of September 9, 1886.
8. Clair Wilcox, *A Charter for World Trade* (Macmillan Company, 1949), p. 3; see also Gerard Curzon, *supra* note 6; and Michael M. Hart, *supra* note 2, p. 42.
9. Henry Joseph Tasca, *The Reciprocal Trade Policy of the United States: A Study in Trade Philosophy* (University of Pennsylvania Press, 1938), p. 18; see also John H. Jackson, *World Trade and the Law of GATT* (The Bobbs-Merrill Company, 1969), p. 37.
10. See John H. Jackson; *supra* note 9, pp. 276–8.
11. Ibid.; see also Kenneth D. Dam, *The GATT Law and International Economic Organization* (University of Chicago Press, 1970), pp. 6–12.
12. Ibid.
13. See Michael M. Hart, *supra* note 2, pp. 44–6, and Michael J. Trebilcock and Robert Howse, *The Regulation of International Trade* (New York: Routledge, 1999), 2nd edn., chapter 5, pp. 112–34.
14. Ibid.
15. John H. Jackson, *National Treatment Obligations and Non-tariff Barriers*, 10(1) *Michigan Journal of International Law*, 212 (1989).
16. Ibid.
17. See Aaditya Mattoo and Arvind Subramanian, *Regulatory Autonomy and Multilateral Disciplines: The Dilemma and a Possible Resolution*, 1(2) *Journal of International Economic Law* 303.
18. Ibid.
19. See Robert Hudec, *GATT/WTO Constrains on National Regulation: Requiem for an 'Aim and Effects' Test* in Robert E. Hudec, *Essays on the Nature of International Trade Law* (Cameron May, 1999), p. 360.
20. Ibid., p. 363.
21. Ibid.
22. Ibid.

23 He claims that starting in the mid-1980s, the GATT, and subsequently the WTO, expanded the National Treatment obligation of the GATT to effectively address *de facto* discrimination, see Warren H. Maruyama, *A New Pillar of the WTO: Sound Science*, 32 *International Lawyer* 651 (1998).
24 See Hudec, *supra* note 19, p. 364.
25 Ibid.
26 Ibid.
27 See Aaditya Mattoo and Arvind Subramanian, *supra* note 17, pp. 303–4.
28 Ibid.
29 Ibid., p. 305.
30 Ibid.
31 Panel Report adopted on November 10, 1987, see GATT, B.I.S.D. 34S/83; see also Pierre Pescatore, William J. Davey, and Andreas F. Lowenfeld (hereinafter Pescatore, Davey, and Lowenfeld), *Handbook of WTO/GATT Dispute Settlement* (Transnational Publishers, 2000), Volume 2. For a brief commentary on this case see Robert E. Hudec, *Enforcing International Trade Law: The Evolution of the Modern GATT Legal System* (Butterworth Legal Publishers, 1993), pp. 212–14.
32 The panel stated the following in regard to the context, purpose, and object of Article III:2:

> Just as Article I:1 was generally construed, in order to protect the competitive benefits accruing from reciprocal tariff bindings, as prohibiting "tariff specialization" discriminating against "like" products, only the literal interpretation of Article III:2 as prohibiting "internal tax specialization" discriminating against "like" products could ensure that the reasonable expectation, protected under GATT Article XXIII, of competitive benefits accruing under tariff concessions would not be nullified or impaired by internal tax discrimination against like products.

33 According to the panel, the increasing imports of "Western-style" alcoholic beverages into Japan bore witness to this competitive relationship and to the potential product substitution through trade among various alcoholic beverages.
34 We are mindful of the fact that GATT and WTO panels and Appellate Body have not made any distinction as such between the explicit discrimination by regulatory measures based on nationality or country of origin of the products and implicit discrimination by facially neutral regulatory measures in the course of determining the consistency or inconsistency of a regulatory measure with Article III of GATT.
35 Report of the Panel adopted on November 7, 1989, see GATT, B.I.S.D. 36S/345; see also Pescatore, Davey, and Lowenfeld, *supra* note 31.
36 *United States – Restrictions on Imports of Tuna* (1991) 30 I.L.M. 1594. The Panel ruling in this case was not formally adopted by the GATT Council.
37 See, e.g., Robert E. Hudec, "The Product–Process Doctrine in GATT/WTO Jurisprudence," in Marco Bronckers and Reinhard Quick (eds.), *New Directions in International Economic Law: Essays in Honour of John. H. Jackson* (Kluwer Law International, 2000); Robert Howse and Donald Regan, "The Product–Process Distinction: An Illusory Basis for Disciplining Unilateralism in Trade Policy," 11 *European Journal of International Law* 249 (2000); and Henry L. Thaggert, "A Closer Look at the Tuna–Dolphin Case: 'Like Products' and 'Extrajuridictionality' in the Trade and Environment Context" in James Cameron, Paul Demaret and Damien Geradin (eds.), *Trade and the Environment: The Search for Balance* (Wm Gaunt & Sons, 1994), vol. I, p. 69.
38 Report of the panel adopted on June 19, 1992, GATT, B.I.S.D. 39S/206; see also Pescatore, Davey, and Lowenfeld, *supra* note 31, at DD88/1.

39 In the panel's view, even if the wine produced from the specified variety of grape were to be considered unlike other wine, the two kinds of wine would still have to be regarded as "directly competitive" products in terms of Article III:2, second sentence, and the imposition of a higher tax on directly competing imported wine so as to afford protection to domestic production would have been inconsistent with that provision.

40 In the panel's view, consumers who purchased low alcohol beer might be unlikely to purchase beer with high alcohol and *vice versa*, and the advertising and marketing by manufacturers showed such different market segments.

41 See GATT Doc. DS31/R, September 29, 1994. For an analysis and critique of this case, see Mattoo and Subramanian, *supra* note 17; and James H. Snelson, "Can GATT Article III Recover From Its Head-On Collision with *United States — Taxes on Automobiles*?", 5 *Minnesota Journal of Global Trade* 467 (1996).

42 The EEC claimed that all automobiles were "like" products and the distinction made on the basis of their value and gasoline consumption resulted in the imposition of internal taxes on imported products in excess of those applied to "like" domestic products. The United States claimed that the tax measures were applied equally to domestic and imported automobiles and the United States and EEC producers manufactured automobiles with both the low and high values as well as with high and low gasoline consumption.

43 See Robert E. Hudec, *supra* note 19.

44 For a recent review of the case-law under Article III pertaining to internal tax discrimination, see Elsa Horn and Petros Mavroidis, "Still Hazy After All These Years: The Interpretation of National Treatment in the GATT/WTO Case-Law on Tax Discrimination," December 3, 2002 (a copy of the manuscript is on file with the authors).

45 *Japan – Taxes on Alcoholic Beverages*, Report of the Panel, WT/DS8/R, WT/DS10/R and WT/DS11/R (July 11, 1996) (96-2651); and *Japan – Taxes on Alcoholic Beverages*, Report of the Appellate Body, WT/DS8/AB/R, WT/DS10/AB/R and WT/DS11/AB/R (October 4, 1996) (96-3951), AB-1996-2.

46 In the panel report, the panel stated that such an examination requires two determinations: (i) whether the products concerned are "directly competitive or substitutable," and (ii) if so, whether the treatment afforded to foreign products is contrary to the principles set forth in Article III:1.

47 The issue in this case was the Japanese Liquor Tax Law that divided all liquors into different categories and subcategories, and applied different tax rates to each of these categories and subcategories. The tax rates were expressed as a specific amount in Japanese Yen per liter of beverage, and for each category or subcategory, the Liquor Tax Law laid down a reference alcohol content per liter of beverage and the corresponding reference tax rate. The European Communities complained, *inter alia*, that Japan had acted inconsistently with Article III:2 of GATT by applying a higher tax rate on the categories of spirits, whisky/brandy and liquors than on each of the two subcategories of shochu. Canada and United States complained that the higher rates of taxation on imported alcoholic beverages including whiskies, brandies, and other distilled alcoholic beverages and liquors than on Japanese shochu imposed under the Liquor Tax Law were inconsistent with Article III:1 and Article III:2 of GATT.

48 The issues raised before the Appellate Body were the conclusions reached by the panel that shochu and vodka are like products and Japan, by taxing the latter in excess of the former, was in violation of its obligation under Article III:2, first sentence, of GATT 1994, and that shochu, whisky, brandy, rum, gin, genever, and liquors are "directly competitive or substitutable products" and Japan, by not taxing them similarly, was in violation of its obligation under Article III:2, second sentence, of GATT 1994. Japan and United States appealed against the panel's findings.

49 According to the panel, even if the adopted panel reports have any legal status, it does not necessarily have to follow their reasoning or results. Although the Appellate Body endorsed the panel's conclusion in regard to unadopted panel reports and did not agree with the conclusion on the legal status of adopted panel reports, it, however, agreed that adopted panel reports are not binding, except on the parties to the dispute, even if they create legitimate expectations among WTO Members and should be taken into account where they are relevant to any dispute.

50 According to the panel, if the "aim-and-effect" test was applied in regard to Article III, then in principle, a WTO Member could, for example, invoke protection of health in the context of invoking the "aim-and-effect" test, and if this were the case, then the standard of proof established in Article XX would effectively be circumvented and WTO Members would not have to prove that a health measure is necessary to achieve its health objective. For a response to the panel's criticism of the "aims and effects" test, see Serena B. Wille, *Recapturing a Lost Oppurtunity: Article III: 2 GATT 1994 Japan-Taxes on Alcoholic Beverages*, Jean Monnet Working Paper 11-97 (NYU School of Law, 1997).

51 According to the Appellate Body, many least-developed countries and developing countries have bindings in their schedules which include broad ranges of products that cut across several different HS tariff headings.

52 The panel noted that a difference in the physical characteristic of alcoholic strength of two products did not preclude a finding of "likeness" especially since alcoholic beverages are often drunk in diluted form. The panel also noted the similar findings in the 1987 *Japan Alcohol* case and that vodka and shochu were classified in the same heading in the Japanese tariffs bindings.

53 According to the panel, the use of additives would disqualify liquors, gin, and genever; the use of ingredients would disqualify rum; and appearance (arising from manufacturing processes) would disqualify whisky and brandy.

54 The panel concluded that the tax imposed on vodka was in excess of the tax imposed on shochu because vodka was taxed at 377,230 Yen per kiloliter – for an alcoholic strength below 38 degrees – 9,927 Yen per degree of alcohol – whereas shochu A was taxed at 155,700 Yen per kiloliter – for an alcoholic strength between 25 and 26 degrees – 6,228 Yen per degree of alcohol.

55 Applying the criterion of elasticity of substitution between products, the panel concluded that shochu, whisky, brandy, rum, gin, genever, and liquors were "directly competitive or substitutable products." To find the elasticity of substitution, the panel relied on the conclusions of the 1987 *Japan Alcohol* case that both white and brown spirits were directly competitive or substitutable products to shochu, the studies put forward by the complainants supporting such elasticity of substitutions, and the evidence submitted by the complainants concerning the 1989 Japanese tax reform which showed that the products in question were essentially competing for the same market.

56 In the view of the Appellate Body, this would eviscerate the distinctive meaning that must be respected due to the distinctions in the wordings of the text of Article III:2, first sentence, and Article III:2, second sentence.

57 The panel concluded that the products at dispute were not similarly taxed because the differences in the amounts of taxes were not *de minimis* and Japan's Liquor Tax Law did not specifically provide that tax/price ratio was the basis of taxation, as there were significantly different tax/price ratios even within the same product categories.

58 To support its conclusion, the Appellate Body noted the findings of the panel that the combination of customs duties and internal taxation in Japan had the impact of making it difficult for foreign-produced shochu to penetrate the Japanese market as well as the impact of not guaranteeing equality of competitive conditions between shochu and the rest

of "white" and "brown" spirits; and thus, through a combination of high import duties and differentiated internal taxes, Japan managed to "isolate" domestically produced shochu from foreign competition.

59 For various aspects of practical difficulties in applying the tests advocated by the Appellate Body in *Japan – Taxes on Alcoholic Beverages*, see Mattoo and Subramanian, *supra* note 17 (arguing that this case follows a strict textual interpretation of Article III:2 which is difficult to apply to a range of known situations); Sarah Hogg and Mahmud Nawaz, "Economic Considerations and the DSU" in James Cameron and Karen Campbell (eds.), *Dispute Resolution in the World Trade Organisation* (Cameron May, 1998) (arguing that the interpretation was focused on *supply side* factors and the key *demand side* question – whether the products concerned competed in the same market – was not considered as important); and Ramon R. Gupta, *Appellate Body Interpretation of the WTO Agreement: A Critique in Light of Japan – Taxes on Alcoholic Beverages*, *6 Pacific Rim Law and Policy Journal 683* (July 1997) (criticising the vague approach in defining "like" and "directly competitive or substitutable" products in light of the importance of predictability and clarity in developing credible dispute settlement procedure).

60 *Canada – Certain Measures Concerning Periodicals*, Report of the Panel, WT/DS31/R (March 14, 1997) (97-0939); and *Canada – Certain Measures Concerning Periodicals*, Report of the Appellate Body, WT/DS31/AB/R (June 30, 1997) (97-2653), AB-1997-2.

61 A split-run edition was one that was distributed in Canada, had more than 20 percent of editorial material substantially the same as the editorial material that appeared in one or more excluded editions of one or more issues of one or more periodicals, and contained an advertisement that did not appear in identical form in all excluded editions.

62 Canada claimed, *inter alia*, that the panel erred in law in finding that imported United States' split-run periodicals and Canadian non-split-run periodicals were like products. The US appeal related to some other issues.

63 Based on a single hypothetical example constructed using a Canadian-owned magazine *Harrowsmith Country Life*, which was previously a split-run periodical but stopped its US edition as a result of the tax, the panel compared two editions, before and after the discontinuation of the US edition, of the same magazine, and concluded that imported split-run periodicals and domestic non-split-run periodicals were "like" products because the two editions of the said magazine would have common end uses, very similar physical properties, nature and qualities as well as they would have been designed for the same readership with the same tastes and habits.

64 The Appellate Body particularly noted the facts that the panel based its findings on a single, incorrect, hypothetical example that involved a comparison between two editions of the same magazine, both imported products, which could not have been in the Canadian market at the same time, but the panel did not examine the evidence of likeness of TIME, TIME Canada and Maclean's magazines, presented by Canada, and the magazines, Pulp & Paper and Pulp & Paper Canada, presented by the United States, or the Report of the Task Force on the Canadian Magazine Industry.

65 In its view, the determination of "likeness" was a delicate process by which legal rules had to be applied to facts, and due to the absence of adequate analysis of facts in the Panel Report in that respect, it was not possible for the Appellate Body to proceed to a determination of "like" products.

66 The Appellate Body rejected the argument of Canada that it did not have the jurisdiction to examine a claim under Article III:2, second sentence, as no party had appealed the findings of the panel on that provision.

67 The Appellate Body rejected the argument of Canada that the Task Force Report's description of the relationship as one of "imperfect substitutability" characterized the absence of perfect substitutability that was required to prove the direct competitiveness or

substitutability of products. In its view, a case of perfect substitutability makes products the "like" products. It also cautioned that the conclusion that imported split-run periodicals and domestic non-split-run periodicals were "directly competitive or substitutable" did not mean that all periodicals belong to the same relevant market, whatever their editorial content. In its view, a periodical containing mainly current news is not directly competitive or substitutable with a periodical dedicated to gardening, chess, sports, music, or cuisine, but news magazines, like TIME and Maclean's, are directly competitive or substitutable.

68 The Appellate Body claimed that the magnitude of the dissimilar taxation was prohibitive.

69 The effects cited were the moving of the production of a split-run magazine of United States for the Canadian market from Canada to the United States and the cessation of production of the US edition by a Canadian split-run periodical after the imposition of the tax.

70 For the analysis of different aspects of the Appellate Body decision in the *Canada Periodicals* case, see Stephen de Boer, "Trading Culture: The Canada–US Magazine Dispute," in James Cameron and Karen Campbell (eds.), *Dispute Resolution in the World Trade Organisation* (Cameron May, 1998); Richard L. Matheny III, "In the Wake of the Flood: 'Like Products' and Cultural Products after the World Trade Organization's Decision in Canada Certain Measures Concerning Periodicals," in 147 *University of Pennsylvania Law Review*, November 1998, p. 245; Sydney M. Cone III, "The Appellate Body and *Harrowsmith Country Life*" in 32(2) *Journal of World Trade* 117 (1998); Chi Carmody, "When "Cultural Identity Was Not At Issue": Thinking About *Canada — Certain Measures Concerning Periodicals*" in 30 *Law and Policy in International Business* 231 (1999); and Trevor Knight, "The Dual Nature of Cultural Products: An Analysis of the World Trade Organization's Decisions Regarding Canadian Periodicals", 57(2) *University of Toronto Faculty of Law Review* 165 (1999).

71 *Korea – Taxes on Alcoholic Beverages*, Report of the Panel, WT/DS75/R and WT/DS84/R (September 17, 1998) (98-3471); and *Korea – Taxes on Alcoholic Beverages*, Report of the Appellate Body, WT/DS75/AB/R and WT/DS84/AB/R (January 18, 1999) (99-0100), AB-1998-7. In this case, the United States and EEC complained against the Korean taxes under the Korean Liquor and Education Tax Laws, as being inconsistent with Article III:2 because they accorded preferential tax treatment to soju, a traditional Korean alcoholic beverage, as compared with certain imported alcoholic beverages.

72 According to the panel, trends are particularly important in the context of experience-based consumer items and it would be unrealistic and, indeed, analytically unhelpful to attempt to separate every piece of evidence and disregard that which discusses implications for market structure in the near future.

73 The panel found that the total tax on diluted soju was 38.5 percent; on distilled soju and liquors 55 percent; on vodka, gin, rum, tequila, and admixtures 104 percent; and on whisky, brandy, and cognac 130 percent.

74 According to the panel, it was based on a very broad generic definition which was defined as soju and then there were specific exceptions corresponding very closely to one or more characteristics of imported beverages that were used to identify products which received higher tax rates. There was virtually no imported soju so the beneficiaries of the tax structure were almost exclusively domestic producers, and the only domestic product which fell into a category with higher tax rates was distilled soju which represented less than one percent of Korean production.

75 *Chile – Taxes on Alcoholic Beverages*, Report of the Panel, WT/DS87/R and WT/DS110/R (June 15, 1999) (99-2313); and *Chile – Taxes on Alcoholic Beverages*, Report of the Appellate Body, WT/DS87/AB/R and WT/DS110/AB/R (December 13, 1999) (99-5414), AB-1999-6.

76 In the panel's view, studies or surveys that reveal the following all serve as evidence of substitutability in end-uses: (i) a tendency among consumers to regard products as substitutes

in satisfying a particular need; (ii) that the nature and content of marketing strategies of producers indicate that they are competing for the expenditure of potential consumers in a particular market segment; and (iii) that distribution channels are shared with other goods.

77 According to the panel, although whisky and pisco were distilled from different substances, namely barley and grapes respectively, they share the characteristics of being potable liquids with high alcohol content, which was the product of distillation, as well as being receptive to mixing with non alcoholic beverages. In any event, even the differences in ingredients between whisky and pisco were not sufficient to render these two distilled alcoholic spirits, both of which have a high alcohol content and more or less satisfy a similar need, incapable of being substituted for each other. As for brandy, cognac and some other spirits, the differences in physical characteristics were only post-distillation differences such as color and smell which were not sufficiently significant to change the basic character of spirits essentially made from grapes or other fruits.

78 According to the panel, between 70 and 80 percent of Chilean production consisted of products with less than 35 degrees alcohol content and, therefore, enjoyed the lowest tax rate of 27 percent. Over 90 percent of pisco was in this category.

79 Under Chilean regulations, most of the imported beverages, such as whisky, had generic names that required them to contain at least 40 degrees of alcohol. Thus, almost 95 percent of imports would be taxed at the highest rate of 47 percent or would lose their ability to retain their generic name or would be required to change an important physical characteristic, namely their water/alcohol ratio.

80 Chile argued that its objectives of the tax measure were maintaining revenue collection; eliminating tax distinctions based on the types of alcoholic beverages; discouraging alcohol consumption; and minimizing the potentially regressive aspects of the reform of the tax system. Examining the relationship between the stated objective and the measure in question, the panel claimed that there was no rational reason why such a structure as devised by Chile was necessary for the purpose of maintaining revenue neutrality, as Chile had acknowledged that the same revenue result could be achieved with a single *ad valorem* rate at some point between 27 and 47 percent. Similarly, the panel claimed that the New System did not achieve the purpose of eliminating type distinctions because the favorable tax treatment accorded to products called "pisco" was removed, but the system was replaced with one providing unfavorable tax treatment for any products called "whisky," "gin," "vodka," or "rum," which happened to be primarily imports. Likewise, the panel claimed that there was no direct correlation between the objective of discouraging alcohol consumption and the measure because the tax differential between products with 35 degrees of alcohol and 39 degrees of alcohol was not the same as the differential between products with, for instance, 40 degrees and 44 degrees of alcohol as the tax rate almost doubled between 35 and 39 degrees but was the same between 40 and 44 degrees. Since the system was based not just on alcohol content, but on *ad valorem* rates qualified by the additional criterion of alcohol content, there appeared to be no correlation between value and alcohol consumption. Finally, minimizing the regressive aspects of the tax reform would be true only if the factual situation were to remain static. In many markets there were quite low priced whiskies sold at the same alcohol content as high priced whisky.

81 For a brief commentary on the Appellate Body decision in this case, see Raj Bhala and David Gantz, "WTO Case Review 2000" in *Arizona Journal of International and Comparative Law*, vol. 18, No. 1, pp. 1–101.

82 *Indonesia – Certain Measures Affecting the Automobile Industry*, Report of the Panel, WT/DS54/R, WT/DS55/R, WT/DS59/R and WT/DS64/R (July 2, 1998) (98-2505).

83 *United States – Standards for Reformulated and Conventional Gasoline*, Report of the Panel, WT/DS2/R (January 29, 1996) (96-0326).

84 See Hudec, *supra* note 19, at 363. For comments on the panel and Appellate Body decisions in this case, see Jennifer Schultz, "The Demise of 'Green' Protectionism: The WTO Decision on the US Gasoline Rule" in 25 *Denver Journal of International Law and Policy* 1 (Fall 1996) (arguing that the case was correctly decided).

85 This case was appealed but Appellate Body did not make any ruling on National Treatment because the issue was not raised in the appeal, see *United States – Standards for Reformulated and Conventional Gasoline*, Report of the Appellate Body, WT/DS2/AB/R (April 26, 1996) (96-1597), AB-1996-1.

86 *European Communities – Regimes for the Importation, Sale and Distribution of Bananas*, Report of the Panel, WT/DS27/R (May 22, 1997) (97-2069) (97-2070) (97-2077) (97-2078); and *European Communities – Regimes for the Importation, Sale and Distribution of Bananas*, Report of the Appellate Body, WT/DS27/AB/R (September 9, 1997) (97-3593), AB-1997-3. The issues at dispute related to Article III:4 of GATT were the EEC procedures and requirements for the distribution of licenses for importing bananas among eligible operators within the EEC, which provided for the allocation of import licenses in regard to 30 percent tariff quota for third country/nontraditional ACP imports to the operators that had marketed EC and/or traditional ACP bananas, on the basis of the average quantities of such bananas marketed in the three most recent years for which data were available, and the issuance of hurricane licenses exclusively to EEC producers or operators including or directly representing a producer adversely affected by a tropical storm who was unable to supply the EEC market. These rules were explicitly discriminatory but the main question was whether or not the provisions of Article III:4 applied to these rules. Once it was concluded that Article III:4 did apply in respect of these rules, the discrimination based on the origin of products was evident. The said EEC licensing procedures and requirements were contested as being inconsistent with the National Treatment obligations of both GATT Article III and GATS (General Agreement on Trade in Services) Article XVII. Both the panel and Appellate Body found these licensing procedures as being inconsistent with both the GATT and GATS National Treatment obligations. For a brief commentary on this case, see Terence P. Stewart and Mara M. Burr, "The WTO's First Two and a Half Years of Dispute Resolution," in 23 *North Carolina Journal of International Law and Commercial Regulation* 481 (1997/1998).

87 *Japan – Measures Affecting Consumer Photographic Film and Paper*, Report of the Panel, WT/DS44/R (March 31, 1998) (98-0886).

88 *Canada – Certain Measures Affecting the Automotive Industry*, Report of the Panel, WT/DS139/R and WT/DS142/R (February 11, 2000) (00-0455); and *Canada – Certain Measures Affecting the Automotive Industry*, Report of the Appellate Body, WT/DS139/AB/R and WT/DS142/AB/R (May 31, 2000) (00-2170), AB-2000-2. For a brief history and objectives behind the Canadian measures as well as the analysis and commentary on the Appellate Body decision in this case, see Raj Bhala and David Gantz, *supra* note 81.

89 This conclusion was based on the facts that, in making the commitments, the companies acted at the request of the Government of Canada ("the Government"); the anticipated Auto Pact between the United States and Canada was a key factor in the decision of the companies to submit these undertakings; the companies accepted responsibility *vis-à-vis* the Government with respect to the implementation of the undertakings contained in the letters, which they described as "obligations" and in respect of which they undertook to provide information to the Government and indicated their understanding that the Government would conduct yearly audits; and until recently the Government gathered information on an annual basis concerning the implementation of the conditions provided for in the letters. The panel rejected the Canadian argument that the commitments expressed in the letters

of undertaking were not "requirements" within the meaning of Article III:4 because the Government of Canada did not negotiate for them, and compliance with the letters was neither legally enforceable nor a condition to obtain an advantage.

90 *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, Report of the Panel, WT/DS161/R and WT/DS169/R (July 31, 2000) (00-3025); and *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef,* Report of the Appellate Body, WT/DS161/AB/R and WT/DS169/AB/R (December 11, 2000) (00-5347), AB-2000-8.

91 The measure in dispute was the Government of Korea's Management Guidelines for Imported Beef which specified that imported beef (except for pre packed imported beef) might only be sold in specialized imported-beef shops and that large-scale distributors (department stores, supermarkets, etc.) must provide a separate sales area for imported beef. Stores selling imported beef were also mandatorily required to display a "Specialized Imported Beef Store" sign to distinguish them from domestic meat sellers. Australia and the United States complained that Korea's requirement was inconsistent with Article III:4. Korea defended the dual retail system for beef on the grounds that it did not impose less favorable treatment on imported beef as domestic and imported beef both were sold in separate shops and there were no limitations on the number of imported-beef shops that could be opened.

92 The Appellate Body noted that the reduction of commercial opportunities was reflected in the much smaller number of specialized imported beef shops (around 5,000 shops) as compared with the number of retailers (around 45,000 shops) selling domestic beef.

93 The Appellate Body also stated that it was *not* holding that a dual distribution system that was *not* imposed directly or indirectly by governmental regulation, but was rather solely the result of private entrepreneurs acting on their own calculation of comparative costs and benefits of differentiated distribution systems, was unlawful under Article III:4.

94 *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, Report of the Panel, WT/DS135/R (September 18, 2000) (00-3353); and *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, Report of the Appellate Body, WT/DS135/AB/R (March 12, 2001) (01-1157), AB-2001-11. For analysis of various aspect of this case, see Laura Yavitz, "The World Trade Organization Appellate Body Report, European Communities – Measures Affecting Asbestos and Asbestos-Containing Products, March 12, 2001, WT/DS135/AB/R" in 11 *Minnesota Journal of Global Trade*, Winter 2002, p. 43; and Robert Howse and Elisabeth Tuerk, "The WTO Impact on Internal Regulations – A Case Study of the Canada–EC Asbestos Dispute," in G. de Burca and J. Scott (eds.), *The EU and the WTO: Legal and Constitutional Issues* (Oxford: Hart Publishing, 2001).

95 The panel simply ignored the arguments of the EEC that the measure itself was origin-neutral and did not seek to protect domestic products because France imports most substitute products from various third countries.

96 Appellate Body in *EEC – Asbestos*, see *supra* note 94.

97 Despite the existence of the same word and similar context, the Appellate Body's efforts to avoid for the purpose of Article III:4 the narrow definition of the word "like" given in the first sentence of Article III:2 seems to be influenced by the possible implication of such interpretation for the objective of the National Treatment principle. It stated that there is no sharp distinction between fiscal regulation covered by Article III:2 and nonfiscal regulation covered by Article III:4 because both forms of regulation can often be used to achieve the same ends. According to it, it would be incongruous if, due to a significant difference in the product scope of these two provisions, Members (of WTO) were prevented from using one form of regulation (for instance, fiscal) to protect domestic production of certain products, but were able to use another form of regulation (for instance, non fiscal) to achieve those results.

98 The Appellate Body also said that where the physical properties are very different, an examination of the evidence relating to consumers' tastes and habits is an *indispensable* – although not, on its own, sufficient – aspect of any determination that products are "like" under Article III:4.

99 However, the Appellate Body in this case did not examine further the interpretation of the term "treatment no less favorable" in the context of Article III:4.

100 He argued that it was difficult for him to imagine what evidence relating to competitive relationships as reflected in end-uses and consumers' tastes and habits could outweigh and set at naught the undisputed deadly nature of chrysotile asbestos fibers, compared with PCG fibers, when inhaled by humans, and thereby compel a characterization of the "likeness" of chrysotile asbestos and PCG fibers. However, he also clarified that he was not suggesting that any kind or degree of health risk, associated with a particular product, would *a priori* negate a finding of the "likeness" of that product with another product, under Article III:4. His suggestion was limited only to the circumstances of *EEC – Asbestos* case, and confined to chrysotile asbestos fibers as compared with PCG fibers.

101 *United States – Tax Treatment for "Foreign Sales Corporations" – Recourse to Article 21.5 of the DSU by the European Communities*, Report of the Panel, WT/DS108/RW (August 20, 2001).

102 *India – Measures Affecting the Automotive Sector*, Report of the Panel, WT/DS146/R, WT/DS175/R (December 21, 2001) (01-6327).

103 In this case, the issue relating to Article III:4 was certain provisions of the 2000 *FSC Repeal and Extraterritorial Exclusion Act* of the United States which was enacted to comply with the DSB recommendations and rulings in *United States – Tax Treatment for "Foreign Sales Corporations*." The EEC claimed, *inter alia*, that the provisions of the said Act which excluded certain extraterritorial income derived from the sale or lease of "qualifying foreign trade property" from taxation were contrary to Article III:4 of the GATT. "Qualifying foreign trade property" was the property made within or outside the United States, and sold for ultimate use outside the United States, no more than 50 percent of the fair market value of which was attributable to "articles manufactured, produced, grown, or extracted outside the United States" and "direct costs for labor … performed outside the United States," which meant that the exclusion from taxation provided by the Act was not available in respect of income derived from the sale or lease of property more than 50 percent of the fair market value of which was attributable to articles made, or costs of direct labor performed, outside the United States. The EEC argued that this foreign articles/labor limitation was inconsistent with Article III:4 as it was a requirement contained in a law which provided less favorable treatment to imported parts and materials than to like domestic goods with respect to their internal use in the production of goods within the United States.

104 See *United States – Tax Treatment for "Foreign Sales Corporations" – Recourse to Article 21.5 of the DSU by the European Communities*, Report of the Appellate Body, WT/DS108/AB/RW (January 14, 2002) (02-0152), AB-2001-8

105 The issue in this case relating to Article III:4 was the indigenization condition contained in Public Notice No. 60 issued by the Government of India under Foreign Trade (Regulation and Development) Act of 1992 and the MOUs required to be signed by manufacturers in order to gain the right to apply for an import license to import the restricted kits and components. The measure in question required the MOU signatories to commit to achieving a level of indigenization of components up to a minimum level of 50 percent in the third year or earlier and 70 percent in the fifth year or earlier, in order to obtain import licenses. The indigenization requirement was, thus, an obligation to use a certain proportion of local parts and components in the manufacture of cars and automotive vehicles. The United States and the EEC argued, *inter alia,* that this requirement accorded

less favorable treatment to imported parts and components and therefore was contrary to Article III:4.

106 See Robert Howse and Donald Regan, *supra* note 37.

107 See Michael Trebilcock, "International Trade and Labour Standards," in Stefan Giller, *International Economic Governance and Non-Economic Concerns* (Spunger-Wien, New York, 2003), at 289.

108 See Frieder Roessler, "Diverging Domestic Policies and Multilateral Trade Integration," in Jagdish Bhagwati and Robert E. Hudec (eds.), *Fair Trade and Harmonization: Prerequisites for Free Trade?*, Volume Two, Legal Analysis (MIT Press, 1996).

109 Donald H. Regan, "Regulatory Purpose and "Like Products" in Article III:4 of the GATT (with Additional Remarks on Article III:2)," 36(3) *Journal of World Trade* 443–78 (2002); and Donald H. Regan "Further Thoughts on the Role of Regulatory Purpose Under Article III of the General Agreement on Tariffs and Trade: A Tribute to Bob Hudec", 37(4) *Journal of World Trade* 737 (2003).

110 See Gaetan Verhoosel, *National Treatment and WTO Dispute Settlement: Adjudicating the Boundaries of Regulatory Autonomy* (Oxford: Hart Publishing, 2002).

111 For a critique, from another perspective, of the type of means-ends test under Article III as advanced by Verhoosel, see Donald Regan; *supra* note 109, at 745–48.

112 For similar economic perspectives on Article III, see Damien Neven, "How Should Protection be Evaluated in Article III GATT Disputes?" (2001) 17 *European Journal of Political Economy* 421; Henrik Horn and Petros Mavroidis, *supra* note 44; and Won-Mog Choi, *Like Products in International Trade Law* (Oxford University Press, 2003).

113 See Michael Trebilcock, Ralph Winter, Paul Collins, and Edward Iacobucci, *The Law and Economics of Canadian Competition Policy* (University of Toronto Press, 2002), chapter 4.

114 See Steven C. Salop and David T. Scheffman, "Raising Rivals' Costs," 73(2) *The American Economic Review*, *Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association* (May, 1983), pp. 267–71; and Thomas G. Krattenmaker and Steven Salop, "Anticompetitive Exclusion: Raising Rivals' Costs to Achieve Power Over Price," 96 *Yale Law Journal* 209 (1986).

115 See Robert Howse and Elisabeth Tuerk, *supra* note 94.

116 Steve Charnovitz also questions the optimism that future WTO panels will tolerate origin-neutral PPMs in the context of Article III, claiming that such optimism "would be unfounded." See Steve Charnovitz, "The Law of Environmental "PPMs" in the WTO: Debunking the Myth of Illegality," 27 *The Yale Journal of International Law* 59, at 92.

117 See *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, Report of the Appellate Body, WT/DS58/AB/R (October 12, 1998) (98-3899), AB-1998-4; and *United States – Import Prohibition of Certain Shrimp and Shrimp Products – Recourse to Article 21.5 of the DSU by Malaysia*, Report of the Appellate Body, WT/DS58/AB/RW (October 22, 2001), AB-2001-4.

118 *Lobster from Canada*, Final Report of the FTA panel (May 25, 1990), T.C.T. 8182.

119 Regan, *supra* note 109.

120 See Hudec, *supra* note 19.

121 *Canada – Measure Affecting Exports of Unprocessed Herring and Salmon*, BISD.35S (1988) 98.

122 Similarly, Steve Charnovitz's suggestion for also taking into account the degree of multilateral approval or disapproval of the regulatory measure in question, especially in the PPM measures, as a factor in evaluating the appropriateness of such measures may also be considered in this regard. See Steve Charnovitz; *supra* note 116, at 105–16.

123 See Joseph H. Weiler, "The Rule of Lawyers and the Ethos of Diplomats: Reflections on the Internal and External Legitimacy of WTO Dispute Settlement," 35 *Journal of World Trade* 191 (2001).

7

Do not Ask Too Many Questions: The Institutional Arrangements for Accommodating Regional Integration within the WTO

Petros Constantinos Mavroidis

Chapter Outline

When it comes to regionalism, economic theory asks the question as to the welfare implications of a preferential trade agreement (PTA). The legal test of consistency of a PTA with the relevant WTO rules only tangentially, if at all, addresses this issue. Instead, the legal test is primarily preoccupied with one question: how to render deviations from the nondiscrimination principle more onerous? The consistency of PTAs with the multilateral rules will be established before the Committee on Regional Trade Agreements ((CRTA), where all PTAs have to be notified) and, eventually, before WTO panels. In practice, for reasons having essentially to do with the voting mechanism in the CRTA, the question of consistency of PTAs with the WTO has been largely left unanswered. On the other hand, WTO Members often lack the incentives to submit to WTO panels cases relating to the consistency of PTAs. As a result, regional groupings of questionable WTO-consistency are being “tolerated” within the multilateral trading system.

1 By Way of Introduction

1.1 The Two Paradoxes of Regionalism

Sapir (1998, p. 718) mentions that when the WTO was established all but three of its original 120 Members were parties to at least one of the 62 regional agreements still in force (the exceptions being Japan, Hong Kong, China, and Korea). This seems quite paradoxical in two ways: on the one hand, one would expect that, following the substantial reduction of tariffs at the multilateral level, there would be no much argument for going regional:[1] in a world of low protection, there is not much to gain for countries that want to further reduce existing protection. On the other hand, there is an obvious tension between "globalization" – notwithstanding the abusive manner that the term is currently used – and regionalism: it cannot be that countries simultaneously pursue the "global" and the regional perspective.

Both paradoxes can be explained though: "regionalism" is not pursued solely for commercial (trade) purposes. It could very well be the case that regionalism is the expression of the political will to "lock in" policies in a predefined perspective. This is what in all likelihood pushed Mexico to join NAFTA. This observation amounts to an automatic rejection of the thesis that the GATT/WTO can provide as safe a "lock in" as NAFTA. And this is not paradoxical. With the exception of TRIPS, the WTO remains essentially a negative integration contract which does not impose common policies on its membership. Nothing, in principle, stops NAFTA members from pursuing common policies beyond what is covered by the current WTO mandate.

Regionalism can be guided by other than – or indirectly related to – trade interests: South East Asian nations are in the process of mounting a common prudential regulation defense in an effort to avoid repetition of the recent financial crisis. On the other hand, it is not economically impossible that regional integration can be welfare maximizing.[2]

1.2 What Economists Ask and What Lawyers Ask About Regional Integration: Two Ships Passing by in the Night

The WTO contract (Article XXIV GATT and Article V GATS) does not put into question the *rationale* for going regional: the rationale for going regional is a sovereign, political decision. What is put into question by the WTO contract is the conditions under which a PTA can legitimately be formed. Essentially since PTAs are inconsistent with the legal obligation imposed on all WTO Members to respect nondiscrimination when conducting their trade policies (as expressed in Article I GATT and Article II GATS – the so-called MFN principle), the WTO contract aims at regulating the conditions under which departures from MFN will be multilaterally acceptable.

Let us make at this point a humble detour to economics. Economic science from early on cast doubt on the welfare maximizing effects of PTAs. Viner's work on trade creation/trade diversion is invariably quoted in this respect.[3] Modern economics

analysis is more subtle. Economic science examines the welfare implications of PTAs: in essence, economic science deals with the issue whether welfare is maximized as a result of the establishment of a PTA. Although a number of empirical papers still support the idea that regionalism diverts trade more than it creates, one can hardly make the argument anymore that regionalism by definition is an obstacle to freer trade since there is evidence that this is not necessarily the case: as Sapir (1998) notes, economists disagree as to whether regionalism represents a threat for the multilateral system.[4] A survey of the economics literature (including the notorious Kemp–Wan theorem) suggests that a formation, for example, of a customs union can enhance welfare of the participants without harming nonparticipating countries.

The legal test only tangentially, if at all, reflects the concerns of modern economic analysis. It essentially cares about one thing: how to make deviations from the MFN principle truly exceptional? In doing that, it imposes difficult, on their face, conditions to meet to any WTO Member aspiring to go regional. These conditions however, do not have directly anything to do with welfare effects as a result of the creation of a PTA.

The role that Article XXIV GATT and its corresponding Article V GATS are called to play is perfectly captured in the language of Article XXIV.4 GATT:

> The contracting parties recognize the desirability of increasing freedom of trade by the development, through voluntary agreements, of closer integration between the economies of the countries parties to such agreements. They also recognize that the purpose of a customs union or of a free-trade area should be to facilitate trade between the constituent territories and not to raise barriers to the trade of other contracting parties with such territories.

One could very well construct examples where trade between members of a PTA is facilitated, no additional trade barriers for third parties are created, and yet trade is diverted toward a member of the PTA, while the most efficient producer worldwide lies outside the PTA. This is especially the case when before the creation of the PTA, countries aspiring to form the PTA knew of high border protection. Article XXIV GATT (and Article V GATS) fail to capture the initial (pre-PTA) position in this respect: the formation of PTAs will be scrutinized in the same manner independently whether at the pre-PTA stage its constituents protected their home markets with high or with low border protection.[5]

1.3 Some Extra Problems Inherent in the Legal Test

The legal test is not a self-interpreting one. The Article XXIV GATT test, for example, conditions consistency of PTAs with the GATT if the *general incidence* of protection post-PTA is not higher than pre-PTA (external requirement) and if constituents of the PTA liberalize *substantially all trade* between them (internal requirement). What is *general incidence* and what *substantially all trade* actually means is open for interpretation.

The WTO and its predecessor, the GATT, have not been successful in clarifying all aspects of the legal test. During the GATT years, typically a Working Party

(participation to which was open to all GATT contracting parties) would be formed to review the consistency of a notified PTA with Article XXIV GATT (Track I). The multilateral control however leaves a lot to be desired. The outcome is disappointing for reasons having to do with the mode of voting in the Working Party (consensus) but also probably because, as Roessler (1993) points out, WTO Members entrusted with enforcing the legal test do not have the incentive to do so and maybe sometimes might have the opposite incentive: that is, not to enforce the test and especially the internal requirement of consistency. The GATT Working Parties were replaced with the entry into force of the WTO by the Committee on Regional Trade Agreements (CRTA). The voting mode however remains the same and the incentive structure for those participating has not changed either.

The multilateral control to review consistency of PTAs can also take place through submission of PTA-related disputes to the WTO adjudicating bodies (Track II). Such submissions however have been rare in practice, essentially because of the frequent absence of incentive to do so. As a result, PTAs have been left largely unchallenged, tolerated in GATT/WTO practice. Unchallenged regionalism essentially means that the WTO, like the GATT before it, has not been instrumental in taming regionalism.

1.4 Organization of the Chapter

This chapter deals with the two questions mentioned so far: what the legal test for consistency of a PTA with the WTO contract is and who should interpret it?

We will first examine whether, with the advent of the WTO and the extension of its mandate to cover trade in services as well, for PTAs to be consistent with the multilateral rules must cover both trade in goods and trade in services (2). Then, we will revert to an examination of the WTO test for consistency of PTAs (3). At the time of writing, the WTO CRTA has yet to conclude on a report in the context of a services PTA. This is the essential reason why we will be focusing on the Article XXIV GATT practice. For reasons that will be explained though, our conclusions under Article XXIV GATT are largely relevant for the purposes of Article V GATS as well. Then I will deal with the interplay between the two tracks available to WTO Members for reviewing consistency of notified PTAs with the WTO rules and ask the additional question whether preference should be given to one of them (4). In other words, in the following section of the chapter, I will ask whether, for reasons having to do with the institutional balance in the WTO, the question of consistency of any given PTA should be properly placed before a panel (judiciary) or whether it should be left to the discretion of the CRTA (where all WTO Members are represented). Brief concluding remarks will follow.

2 Should PTAs Cover Both Goods and Services in the WTO Era?

Before the entry into force of the WTO, in the GATT years, only trade in goods was regulated at the world level. Naturally, the rules on PTAs were confined to trade in goods. With the entry into force of the WTO Agreement, the picture has changed: trade

in services is part of the WTO contract. The question therefore is to what extent PTAs, in order to be compatible with the multilateral rules, can be confined to only trade in goods (or trade in services, as the case may be), or, conversely, whether deviations from MFN should be permitted only when both goods and services are covered.

Practice sometimes develops before rules are clarified (especially when rules are not clear) and often preempts this process. In the WTO so far, practice has developed in the sense that WTO Members have entered into PTAs confined to either goods or services. Such PTAs have been notified to the CRTA and are being customarily reviewed without any WTO Member objecting to their goods-only or services-only character. It is probably premature to talk of practice in the institutional sense of the term since the WTO contract entered into force only on January 1, 1995.

With this caveat, we can still legitimately ask the question whether WTO adjudicating bodies would adhere to this emerging practice? WTO panels have so far interpreted the Agreement establishing the WTO as one agreement to which GATT, GATS, and TRIPS are annexes.[6] The WTO Appellate Body endorsed this point of view stating that:

> Unlike the previous GATT system, the *WTO Agreement* is a single treaty instrument which was accepted by the WTO Members as a "single undertaking." Article II:2 of the *WTO Agreement* provides that the Multilateral Trade Agreements in Annexes 1, 2 and 3 are "integral parts" of the *WTO Agreement*, binding on all Members.[7]

This should mean that the GATT, the GATS, and the TRIPS are not autonomous agreements: on top of the identical membership to all three, the three agreements are technically annexes to the same agreement.

The next step in this analysis is to examine to what extent nondiscriminatory trade is a WTO- rather than a GATT- or GATS-principle. Were one to accept that the principle of nondiscrimination is a WTO-principle, then the legal consequence would be to seek a WTO exception for PTAs, that is an exception covering both trade in goods as well as trade in services. The opposite would be the case were one to answer the question in the negative.

The Agreement establishing the WTO does not mention in its text the principle of nondiscrimination, the Agreement itself being essentially of technical nature and containing institutional provisions. The preamble however states that WTO Members aim, *inter alia*, at "the elimination of discriminatory treatment in international trade relations." The preamble, of course, does not contain any legally binding language, although, technically speaking, it is part of the context of the relevant provisions. Consequently, contextual arguments could be advanced in favor of a "unitary" approach whereby for WTO Members to enter into a WTO-compatible PTA, they would have to cover both goods and services.

Such arguments[8] are of limited legal value. Adjudicating bodies will privilege binding over hortatory language and will pay little, if any, attention to the preamble. Hence, the unavoidable legal conclusion is that, under the existing WTO Agreement, WTO Members can legitimately enter into PTAs that cover either goods or services. If they so wish however, they can go ahead and sign PTAs that cover both goods and services. If they do so, they will have to notify (see *infra*) the WTO and be subjected to a control of consistency under both Article XXIV GATT and Article V GATS.

3 When is a PTA Compatible with the WTO Contract?

In what follows, I address the issue of consistency of a PTA (either a goods- or a services-PTA) with the multilateral procedural and substantive rules. I start the analysis from the GATT perspective and deal first with Track I, that is, multilateral review of a PTA by the CRTA.

3.1 Track I: The Article XXIV GATT Test in a Nutshell

Article XXIV GATT distinguishes between two forms of PTAs: customs unions (CUs) and free trade areas (FTAs). There is an overlap between a CU and an FTA: in both forms of regional integration, parties to the agreement will have to liberalize substantially all trade between them. The difference between these two forms of regional integration is that in the case of a CU, contrary to what happens in an FTA context, there is provision for a common external policy as well.

Essentially Article XXIV GATT imposes three obligations on WTO Members wishing to enter into an PTA:

(i) an obligation to notify the PTA (procedural requirement). Since regional integration essentially amounts to an exception from the basic obligation to treat international trade in a nondiscriminatory manner, the resulting legal consequence is that WTO Members wishing to enter into a PTA, and consequently deviate from the obligation to treat trade from all other WTO Members in a nondiscriminatory manner, will have the burden of proof to state that they have complied with the relevant multilateral rules (*Quicunque exceptio invokat ejusdem probare debet*);
(ii) an obligation to liberalize among constituents of the PTA substantially all trade (substantive requirement); and
(iii) an obligation not to raise the overall level of protection and make access of products of third parties not participating in the PTA more onerous (substantive obligation).

The first obligation is of procedural nature and paves the way for a forum to be provided where the two substantive obligations will be tested. The first obligation results from the exceptional character of Article XXIV GATT. The second is the so-called *internal requirement* whereas the third is the so-called *external requirement*.

As mentioned above, Article XXIV GATT deals with FTAs and CUs only. This does not mean, however, that WTO Members cannot go further than the CU level. The EC experience is a perfect example. Article XXIV GATT consequently, has a function similar to a minimum threshold: WTO Members wishing to integrate have to *at least* comply with what it dictates.

Until recently, maybe naively, it was assumed that, for example, an economic and monetary union by definition encompasses a CU or at the very least an FTA. Hence by examining the CU or FTA component of an economic and monetary union, the GATT test for consistency of a PTA would have been satisfied. As the recent Asian experience shows however, this is not necessarily the case. WTO Members might be willing to integrate their prudential regulations without entering into an FTA or a CU. It is still too early in the day to pronounce on the consistency of such agreements with the WTO. It seems that such examples have more to do with the political will to provide a strong common front against speculators and there is no assessment of their discriminatory – in case they exist – repercussions.

3.2 The Obligation to Notify

As stated above, WTO Members deciding to enter into a PTA have to notify the WTO of their intentions (Article XXIV.7 GATT). The CRTA is the WTO *ratione materiae* competent organ which will receive notifications and examine the compatibility of the notified PTA with the multilateral rules. I briefly go through the mechanics of the CRTA before analyzing the obligation to notify.

3.2.1 To notify where?

A brief presentation of the CRTA. The CRTA is the successor to GATT Article XXIV Working Parties. It was established through a decision by the WTO General Council on February 7, 1996.[9] All WTO Members can participate in the CRTA. The mandate of the CRTA is described in the mentioned decision of February 7, 1996:

(a) to carry out the examination of agreements in accordance with the procedures and terms of reference adopted by . . . and thereafter present its report to the relevant body for appropriate action;

. . .

(d) to consider the systemic implications of such agreements and regional initiatives for the multilateral trading system . . .

The CRTA decides by consensus. Rule 33 of the Rules of Procedure for Meetings of the CRTA stipulates that:

Where a decision cannot be arrived at by consensus, the matter at issue shall be referred, as appropriate, to the General Council, the Council for Trade in Goods, the Council for Trade in Services or the Committee on Trade and Development.[10]

CRTA reports are adopted independently whether CRTA Members have unanimously concluded that a given PTA is consistent or inconsistent with the multilateral rules. It could very well be the case that an adopted report reflects divergent views on this issue.[11] Hence, reports that are adopted by consensus do not necessarily conclude on the consistency or the inconsistency of a PTA with the multilateral rules: final reports can be inconclusive and yet adopted (more on this point, *infra* under 4).

From Working Parties to the CRTA: just a cosmetic change? As stated above, the CRTA is now playing the role of GATT Working Parties. In a sense, Article XXIV GATT Working Parties have been consolidated into one Committee, the CRTA. The rules of the game have not changed though: the CRTA decides by consensus.

The fact that the CRTA continues, like the Article XXIV GATT Working Parties before it, to decide by consensus does not mean that the passage to CRTA is without any effect. In a sense the change is not a purely cosmetic one: there is an argument that moving from *ad hoc* groups to one consolidated body will contribute to increased expertise and coherence when it comes to dealing with regional integration at the WTO-level. National delegates serving at the CRTA will be focusing on regional integration issues (since they will, in all likelihood, be appointed to serve at the CRTA for the duration of their mandate) and will not anymore deal on a pure *ad hoc* basis with such issues. Moreover, they will in all likelihood be confronted by their own "jurisprudence": more likely than not, they will have before them recurring themes. They cannot easily hide behind the argument "it was someone else's Working Party."

The powers of the CRTA. In principle the CRTA has unlimited powers. Article XXIV.7 GATT relevantly reads in this respect:

> ... the CONTRACTING PARTIES and shall make available ... such information ... as will enable them to make *such reports and recommendations to contracting parties as they may deem appropriate* (emphasis added).

The language suggests that the CRTA can make any recommendation it deems appropriate. In principle, one cannot exclude that the CRTA concludes that a notified PTA is GATT-inconsistent.

When it comes to notified *interim agreements* leading to the establishment of a CU or an FTA, there is no ambiguity that the CRTA enjoys such wide powers. Article XXIV.7(b) GATT reads:

> If ... the CONTRACTING PARTIES find that such agreement is not likely to result in the formation of a customs union or of a free-trade area ... the CONTRACTING PARTIES shall make recommendations to the parties to the agreement. *The parties shall not maintain or put into force, as the case may be, such agreement if they are not prepared to modify it in accordance with these recommendations* (emphasis added).

In such cases, the CRTA can effectively "strike down" a proposed CU or FTA. GATT CONTRACTING PARTIES however, never in the history of GATT have exercised their authority in such a drastic way. Nothing changed in the WTO era.

The point here however is that in principle, the CRTA has the power to recommend that a notified interim PTA should not enter into force unless modified according to the opinion of the CRTA.[12] It is true that the same "sweeping" powers are not explicitly mentioned in the body of Article XXIV.7(a) GATT with respect to notified CUs and FTAs that do not know of an interim phase. Should the CRTA's powers be any different in such cases?

There are good arguments in favor of an affirmative response: there are strong similarities in the language used in Article XXIV.7(a) and in Article XXIV.7(b) GATT.

In both cases, that is independently whether an interim phase is foreseen, if prospective action (as should be the case, see *infra* under 3.2.2) is notified, the PTA *will* happen. The fact that in the case of XXIV.7(b) an interim phase is forecasted is not enough by itself to justify a stricter standard for interim agreements?[13]

In practice, however, never has a Working Party gone so far as to unanimously pronounce on the incompatibility of a notified PTA with Article XXIV GATT.[14]

3.2.2 TO NOTIFY WHEN?

The law . . .. Article XXIV:7(a) GATT reads:

> Any contracting party *deciding to enter* into a customs union or free-trade area, or an interim agreement leading to the formation of such a union or area, shall *promptly* notify the CONTRACTING PARTIES and shall make available . . . such information . . . as will enable them to make such reports and recommendations to contracting parties as they may deem appropriate (emphasis added).

The language of Article XXIV GATT suggests that what is notified is *prospective action*. The language stops short of mentioning that Article XXIV GATT operates as a "green light," necessary to comply with, for a PTA to be GATT-compatible. It could, of course, be legitimately argued that since what is notified is prospective action and since the consistency of prospective action is the subject-matter of multilateral review, WTO Members should refrain from *practicing* their PTAs before they are given the "green light" to do so. As will be shown in the discussion on Track II, WTO Members that decide to go ahead and behave as if their PTA was legitimate while the final report of the CRTA is still pending, do so at their own risk: they might face a legal challenge against their discriminatory practices in favor of their PTA-partners before a WTO panel.

. . . And the practice. Political reality however has moved to the opposite direction. Most of the time, especially recently, PTAs have been notified, contrary to the wording of Article XXIV:7(a), *ex post facto* (after their establishment).

For example, the NAFTA was signed on December 17, 1992, entered into force on January 1, 1994 and a Working Party to examine its consistency with the GATT rules was established only on March 23, 1994. The respective dates for the EC–Visegrad Agreements (free-trade areas with Hungary, Poland, and the Czech and Slovak Federal Republic Interim Agreement) are December 16, 1991, March 1, 1992 and April 30, 1992.[15]

Working Parties (and now the CRTA) consequently, have been presented with a *fait accompli*: what was originally supposed to an *ex ante* review became slowly an *ex post* review with all the problems that such a shift in time might entail. Since the text of Article XXIV.7 GATT has not been modified with the advent of the WTO, the CRTA has the same powers with Article XXIV Working Parties before it to enforce the contract. The question becomes what is the appropriate remedy in case a notified PTA, which meanwhile has entered into force, is found not to comply with the WTO rules by the CRTA? We will revert to this question under Section 4.

3.2.3 NOTIFYING PTAS UNDER THE GATS

What has been described above is, with few deviations, relevant for PTAs notified under the GATS. As under the GATT, Article V.7(a) GATS, obliges WTO Members to notify the CRTA of their PTA (procedural obligation). Article V GATS does not distinguish between FTAs and CUs. When it comes to the substantive obligations, Article V GATS, like Article XXIV GATT, distinguishes between an internal and external requirement as well. The language however, is not the same. Let us first start with the internal requirement:

1. This Agreement shall not prevent any of its Members from being a party to or entering into an agreement liberalizing trade in services between or among the parties to such an agreement, provided that such an agreement:

(a) has substantial sectoral coverage; and
(b) provides for the absence or elimination of substantially all discrimination, in the sense of Article XVII, between or among the parties, in the sectors covered under subparagraph (a), through:
   (i) elimination of existing discriminatory measures, and/or
   (ii) prohibition of new or more discriminatory measures,

either at the entry into force of that agreement or on the basis of a reasonable time-frame, except for measures permitted under Articles XI, XII, XIV, and XIV bis.

The term substantial coverage, which is not identical to the term substantially all trade appearing in Article XXIV GATT is further explained in a footnote in the following manner:

This condition is understood in terms of number of sectors, volume of trade affected and modes of supply. In order to meet this condition, agreements should not provide for the *a priori* exclusion of any mode of supply (italics in the original).

We now move to the external requirement (Article V.4 GATS):

Any agreement referred to in paragraph 1 shall be designed to facilitate trade between the parties to the agreement and shall not in respect of any Member outside the agreement raise the overall level of barriers to trade in services within the respective sectors or subsectors compared to the level applicable prior to such an agreement.

We can conclude from the above that the structure of Article XXIV GATT is followed in Article V GATS, the absence of a distinction between FTAs and CUs in the latter notwithstanding. The language is somewhat different but the logic remains the same: as is the case with Article XXIV GATT, Article V GATS is not directly concerned with the welfare implications of a services-PTA. Similarities do not end here: as is the case under the GATT, Article V GATS also distinguishes between a procedural and two substantive obligations. My analysis in the rest of the chapter will focus on goods-PTAs where there is a critical mass of cases. At the moment of writing, the CRTA has issued no report on a services-PTA. *Mutatis mutandis* however, my analysis in the field of goods is relevant, for the reasons mentioned above, for services-PTAs as well. To the extent necessary, I will, in what follows, explicitly revert to a discussion of regional integration under the GATS.

3.3 Substantive Requirements

3.3.1 To enforce or not to enforce? That is the question

As already stated, WTO Members deciding to enter into a PTA will so notify the CRTA where participating WTO Members will review whether the two substantive requirements mentioned above are met. The CRTA however does not have a voice independent of that of its Members, that is the WTO Members participating in the process. Hence, at the end of the day the quantity and quality of the multilateral review will depend on the will of individual WTO Members. As a general observation (and this is a point to which I will come back later), it is clear why WTO Members nonparties to the PTA would be willing to enforce the external requirement: if external barriers are raised their export trade toward the PTA will suffer. It is maybe counterintuitive why they would be willing to enforce the internal requirement: it seems arguable that the less trade liberalization exists among constituents of a PTA, the less trade diversion will take place. Of course, the opposite could be the case as well. An example with complementary products would be most appropriate to illustrate this: the more members of a PTA liberalize trade in cars, the likelier it would be for WTO Members nonparties to the PTA to sell wheels in the PTA-market. At best, the argument why nonparties would be willing to enforce the internal requirement is less clear than with respect to the external requirement.

3.4 The External Requirement

3.4.1 The external requirement for free trade areas

As already stated above, the difference between a CU and an FTA is the presence in the former and the absence in the latter of a common external policy. This difference has direct implications in the regulation of the external requirement laid down in Article XXIV.5.

With respect to FTAs first, Article XXIV.5(b) reads:

> . . . duties and other regulations of commerce . . . shall not be higher or more restrictive than the corresponding duties and other regulations of commerce existing in the same constituent territories prior to the formation of the free-trade area . . .

Article XXIV.5(b) states that WTO Members participating in a PTA cannot modify their external protection when joining an FTA. This approach is dictated by the very nature of FTAs, since an FTA aims only at liberalizing trade within its constituents without addressing at all the question of external to the FTA protection.[16] Hence, since by joining an FTA, WTO Members retain their sovereignty with respect to external protection, there is no reason to see changes in this respect. This does not mean however that internal regulations cannot affect external protection: by adopting rules of origin members of an FTA can drastically change external protection. Rules of origin are particularly

"appropriate" in an FTA context: unless goods circulating through an FTA are accompanied by a "certificate of origin," exporters will have an incentive to the cheapest port of entry (since external protection remains an issue of national sovereignty and it could very well be the case that there are asymmetries as to the level of customs duties among members of an FTA). The first best of course would be that the national systems conferring origin at the pre-FTA stage remain in force unchanged post-FTA. This is almost never the case though. Members of the FTA renegotiate "regional" rules of origin that, from an empirical perspective, more often than not are stricter after the creation of a FTA.[17] In the absence of a WTO Agreement on Rules of Origin, "regionalisation" of rules of origin can only be challenged by the tenant of Article XXIV.5(b) GATT.

3.4.2 THE EXTERNAL REQUIREMENT FOR CUSTOMS UNIONS

With respect to CUs, the external requirement is distinguished into two specific obligations: an obligation not to raise the overall level of protection above a certain threshold and a specific obligation to compensate in cases where customs duties in some constituents of a CU had to be raised to match the CU-level.

CUs aim on top of liberalizing trade among members of the PTA to provide for a common external protection. In case of asymmetrical external protection between the prospective members of the CU (some being more and some being less open to foreign products) adjustments will have to be made. Only in exceptional circumstances (cases of perfect symmetry) no adjustments are needed. Such cases however, in practice are hard to find.

Consequently, the external requirement can pose problems in the context of a CU. Article XXIV.5(a) reads:

> . . . duties and other regulations of commerce . . . shall not *on the whole* be higher or more restrictive than *the general incidence* of the duties and regulations of commerce applicable in the constituent territories prior to the formation of such union . . . (emphasis added).

The italicized words mark the difference between the text of Article XXIV.5(b) GATT and that of Article XXIV.5(a) GATT. At first, the words seem to suggest the, in practice, inevitable need for adjustment. "On the whole" and "general incidence" invite for a comparison of the general (and not item by item) picture before with the picture after the formation of the CU.[18]

Subsequent practice seems to accept that in principle an item-by-item approach is unwarranted in the context of Article XXIV.5(a) GATT, but there is disagreement as to the precise level on which comparisons will take place. The 1983 Working Party report on "Accession of Greece to the European Communities" contains the view expressed by the EC that "Article XXIV.5 required only generalized, overall judgment on this point"[19] but it failed to reach a consensus among its members on this point. By the same token, the 1988 Working Party report on "Accession of Portugal and Spain to the European Communities" contains the view of the EC that "Article XXIV.5 only required an examination on the broadest possible basis."[20] The same report though contains the view of a member of the Working Party which "could not accept the Communities' contention

that the extension of the tariff of the EC/10 to the EC/12 was compatible with their obligations under Article XXIV.5(a) regardless of the effect on the tariffs of Spain and Portugal. Article XXIV.5(a) required a comparison with the pre-accession tariffs of the constituent territories and the relative size of those territories was not a relevant factor."[21]

Moreover, disagreements appeared often among Working Party members as to whether bound or applied rates should be used in the context of Article XXIV.5(a) GATT.[22]

The picture seems much clearer now with the entry into force of the "WTO Understanding on the Interpretation of Article XXIV of the General Agreement on Tariffs and Trade 1994." Panels have yet to pronounce on the legal nature of the WTO Understandings. It seems reasonable to conclude though that they constitute international agreements concluded by WTO Members and that they should be interpreted in accordance with the Vienna Convention on the Law of Treaties. The language used in the Understanding is self-explanatory:

> The evaluation under paragraph 5(a) of Article XXIV of the general incidence of the duties and other regulations of commerce applicable before and after the formation of a customs union shall in respect of duties and charges be based upon an overall assessment of weighted average tariff rates and of customs duties collected. This assessment shall be based on import statistics for a previous representative period to be supplied by the customs union, on a *tariff-line basis* and in values and quantities, broken down by WTO country of origin. The Secretariat shall compute the weighted average tariff rates and customs duties collected in accordance with the methodology used in the assessment of tariff offers in the Uruguay Round of Multilateral Trade negotiations. For this purpose, the duties and charges to be taken into consideration shall be the applied rates of duty. It is recognized that for the purpose of the overall assessment of the incidence of other regulations of commerce for which quantification and aggregation are difficult, the examination of individual measures, regulations, products covered and trade flows affected may be required. (emphasis added).

Moving to tariff-lines in itself provides the necessary precision to the terms "on the whole" and "general incidence." Hence, with respect to Article XXIV.5(a) GATT it seems appropriate to conclude that the test for consistency post-Understanding is precise enough for interpreters.

There is an additional provision relevant only to the formation of CUs. Article XXIV.6 GATT reads:

> If, in fulfilling the requirements of subparagraph 5(a), a contracting party proposes to increase any rate of duty inconsistently with the provision of Article II, the procedure set forth in Article XXVIII shall apply. In providing for compensatory adjustment, due account shall be taken of the compensation already afforded by the reduction brought about in the corresponding duty of the other constituents of the union.

Article XXIV.6 GATT deals only with customs duties. In that it is *lex specialis* to Article XXIV.5(a). An example can help best illustrate the function of XXIV.6 GATT. Countries A, B, and C decide to enter into a CU. Assume that before the formation of the CU, the tariff protection (bound rates)[23] of the automotive sector in the three

countries was the following:

A	20%
B	30%
C	40%

When the CU is formed, A, B, and C decide to bind customs duties at 30 percent for the automotive sector for the CU as a whole. Arguably, they have met their obligations under Article XXIV.5(a) GATT. They have not necessarily met their obligations under Article XXIV.6 GATT though. As will be shown in what follows, if Article XXIV.5(a) GATT is violated, Article XXIV.6 GATT will be *ipso facto* violated; However, compliance with Article XXIV.5(a) GATT does not automatically lead to compliance with Article XXIV.6 GATT.

Article XXIV.6 GATT comes into play because A had to raise its pre-CU duty from 20 to 30 percent. In such cases, according to Article XXIV GATT, Article XXVIII GATT negotiations will start. This means that WTO Members which have "initial negotiating rights," "principal supplying interest," or "substantial interest" will participate in the negotiations with the members of the CU; such negotiations aim at compensating those WTO Members which will have more difficult access to A's market as a result of the formation of the CU.

The second sentence of Article XXIV.6 GATT makes it clear that "in-built compensation" will be taken into account, that is new market opportunities in those constituents of the CU that were forced to lower their pre-CU duties in order to comply with Article XXIV.5(a) GATT.

First scenario: A is a low per capita income small country, whereas C is a high per capita income large country. Neither A nor C produce cars, or their domestic production allows for substantial amount of imports. The fact that C lowers its duties from 40 to 30 percent, in all likelihood, overcompensates the fact that A raised its own duties from 20 to 30 percent. This is the notion of "in-built" compensation: C will import more cars and hence exporters will be compensated for their losses due to the fact that they will export less to A.

Second scenario: A is the high per capita income large country, whereas C is the low per capita income small country. In this case, the amount of trade lost because A had to raise its duties is, most likely, not compensated by the fact that C lowered its own duties. In such cases, there is nothing like "in built" compensation in the formation of the CU for exporters. Hence, something has to be done. Article XXIV GATT calls for compensation which will be offered to the WTO Members in accordance with an Article XXVIII-type of procedure.

3.4.3 THE EXTERNAL REQUIREMENT FOR PTAS IN THE GATS-CONTEXT

As stated above, the GATS regime does not deal with external protection. However, Article V.5 GATS makes it plain that compensation analogous to that offered in the

GATT context under Article XXIV.6 GATT will be offered in the GATS-context as well, in case a WTO Member entering into an RIA has to modify its schedule.

3.5 The Internal Requirement

3.5.1 NO CONSENSUS ON WHAT "SUBSTANTIALLY ALL TRADE" MEANS

According to Article XXIV.8 GATT, WTO Members wishing to enter into a CU or an FTA will have to eliminate duties and other restrictive regulations of commerce with respect to substantially all trade in products originating in the constituents of the RIA. Contrary to what happened in the context of Article XXIV.5(a) GATT, no clarification of the terms included in Article XXIV.8 GATT took place during the Uruguay Round negotiations. And, as it will be shown, this is probably the priority area where clarification was needed. Article XXIV.8 GATT throughout the years has led to major controversies: Working Party members simply could never agree on the actual meaning of the terms "substantially all trade," or "other regulations of commerce" included in the body of Article XXIV.8 GATT.

Substantially all trade: a matter for interpretation? The attitude of GATT contracting parties reveals a reluctance to submit the interpretation of the term to panels. No such cases are reported throughout the GATT years although dozens of PTAs came into being in the same period. This is not hard to understand though and it should not quickly lead to the conclusion that GATT contracting parties did not trust panels with such "loaded" issues: nonsubmission to a panel could simply be due to the lack of incentive that any GATT contracting party had to play the "policeman" of the GATT. As already stated above, although in some cases more internal liberalization could provide nonparticipants in an FTA with more export opportunities (that cars and wheels example stated above), a good argument can be made that such cases are not typical.

True, the conclusion above does not rest on empirical evidence. In fact, there is no comprehensive empirical evidence on this issue. However, it would seem odd that GATT contracting parties not participating in PTAs could win from increased internal liberalization and yet never attempted to enforce Article XXIV.8 GATT. This is particularly odd since, as will be shown later in this chapter, most PTAs fail the Article XXIV.8 GATT requirement on rather obvious grounds.

The Vienna Convention on the Law of Treaties constraint. Fifty years of GATT practice later, the term "substantially all trade" remains unclear. To interpret this term, we will have to observe the constraint imposed by Article 3.2 DSU, that is, any interpretation advanced has to be in conformity with the customary rule of interpretation. It is reminded here, that constant case law in the WTO (both at the panel and at the Appellate Body level) holds for the proposition that when Article 3.2 DSU refers to customary rule of interpretation, it refers to the rule of interpretation embodied in the Vienna Convention on the Law of Treaties (VCLT).

Textual and contextual arguments support a restrictive reading of Article XXIV.8 GATT which would entail as a consequence a "tough" control in this respect. The term

is not self-interpreting; to the contrary, it invites interpretation. In essence, the drafters of GATT felt that further precision was unwarranted or inappropriate (or maybe could not agree on a more precise term) and left it to the discretion of members of the various Article XXIV GATT Working Party reports to pronounce on a case-by-case basis. Practice in this field however, developed in a way that it offers minimal guidance, if any, as to the meaning of the term.

It is reminded that the interpretative rule contained in Article 31 VCLT obliges the interpreter to examine:

the ordinary meaning of the terms;

in their context;

in light of their object and purpose;

taking into account any subsequent decision;

taking into account subsequent practice.

In case the interpreter finds that, at the end of the exercise described above, the meaning of the terms is manifestly unreasonable, recourse should be made to supplementary means of interpretation (Article 32 VCLT). By "supplementary means of interpretation," we understand essentially the preparatory work (*travaux préparatoires*) of an international treaty (in our case, the GATT). Other means however (such as, e.g., the circumstances surrounding the conclusion of the treaty), cannot, in principle, be excluded.

It is true, that recourse to Article 32 VCLT can also be made in cases where the interpreter has reached a satisfying solution, by exhausting all elements mentioned in Article 31 VCLT, and still wants to confirm the interpretation reached. The International Court of Justice (ICJ) though, in a series of decisions on this issue, has repeatedly stated that recourse to Article 32 VCLT simply in order to confirm a meaning is not warranted.

One can safely make in addition the argument nowadays that the ICJ is quite skeptical when presented with evidence based on the preparatory work of an international agreement: it has not accepted arguments based on the preparatory work if one of the parties to the dispute, which is a member of the international agreement at hand, did not participate in the preparatory work. Although the ICJ's attitude does not have to be emulated by WTO panels, it is definitely relevant when it comes to judging whether recourse to preparatory work should be made at all. Moreover, one should note that in this respect the case law of the ICJ is particularly relevant for the WTO-context, where traditionally only a few Members participate in the preparatory work (compare, for example, the recent experience in the Uruguay Round). It should be noted that GATT/WTO panels, in general, show the same reticence as the ICJ to have recourse to the means laid down in Article 32 VCLT.[24] Consequently, when interpreting the two terms, we will have recourse to Article 32 VCLT only in case we conclude that the interpretation reached is manifestly unreasonable.

The ordinary meaning of the terms. The term "substantially all trade" is qualified by the use of the word "substantially." The term "substantially" unambiguously indicates that not all trade between members of a CU or of an FTA has to be covered for a CU or an FTA to be judged compatible with Article XXIV:8 GATT. At the same time though,

the question remains *how much* trade should be covered for a CU or an FTA to pass the Article XXIV:8 GATT-test?

It has been suggested in the past that the term has both a *quantitative* as well as a *qualitative* component, in the sense that it covers a percentage of trade and at the same time no major sector of a national economy can be excluded.[25] This distinction is misleading: the quantitative aspect of the term largely suffices to address the issue here, since, depending on what percentage of trade one is talking about, one could end up with no major sectors uncovered.

This interpretation seems to be further supported by one word in the body of Article XXIV.8 GATT that often escapes the attention in discussion on this issue: the word "all." The term is not "substantially most of the trade" or "a substantial portion of the overall trade"; it is "substantially *all* trade." This suggests that the tendency should be to measure everything by reference to 100 percent of the trade involved and check how close one is to this figure. The term used supports if anything a quantitative test to measure what substantially all trade can mean.

The context. The context of an article of an international treaty is at least the remaining articles of the same agreement and, sometimes, other agreements dealing with the same (or, even comparable) subject matter. We have already presented the relevant subparagraphs of Article XXIV (4, 5, 6, and 7) GATT. None of them however sheds any light on the interpretation of the term.

Consequently, we turn to the rest of the GATT. Article XXIV GATT must be read in the context of Article I GATT which lays down the basic obligation imposed on WTO Members not to discriminate between products of other WTO Members (the MFN-clause).[26] This assertion is mandated, *inter alia*, by the second sentence of Article XXIV:4 GATT which relevantly reads:

> . . . They also recognize that the purpose of a customs union or a free-trade area should be to facilitate trade between the constituent territories and not to raise barriers to the trade of other contracting parties with such territories.

Hence, there is institutional recognition that a PTA can, in principle, upset the fundamental GATT-obligation not to discriminate. The discrimination element is included in Article XXIV.8 GATT which serves two purposes:

> on the one hand, it allows members of a PTA to treat other members of a PTA in a way more favourable than the way they treat other WTO Members, since obstacles to substantially all trade between members of the PTA are eliminated (and this is not the case for the other WTO Members which do not participate in the PTA);
>
> on the other, it sets a threshold which must be complied with, that is for a PTA to be a GATT-compatible PTA, it has to, *inter alia*, guarantee the elimination of duties and other restrictive regulations of commerce with respect to substantially all trade between the members of the PTA.

Article XXIV.8 GATT is an exception to Article I GATT only to the extent that Article XXIV.8 GATT has been complied with. Legally, to acknowledge that a treaty provision is of exceptional character has one consequence: the burden of proof to demonstrate compatibility rests with the party invoking the exception. This is a

well-established principle in public international law (*Quicunque exceptio invokat, ejusdem probare debet*).

The panel report on *Turkey – Restrictions on Imports of Textile and Clothing Products*[27] in a way "ducked" the issue. It definitely acknowledges that Turkey had the burden of proof to show that Article XXIV GATT had been complied with. It did so however largely because Turkey only advanced the argument that its textile policies were anyway justified through recourse to Article XXIV GATT (para. 9.58). Hong Kong, China, a third party to the dispute, defended the point of view that, in accordance with public international law, Turkey was obliged to carry the burden of proof anyway since Article XXIV GATT is an exception to Article I GATT. The panel noted Hong Kong, China's argument (para. 9.58 *in fine*) but did not take a clear position either in favor or against it.

On appeal, the Appellate Body went even further and in an *obiter dictum* made it clear that it expects parties invoking an Article XXIV GATT defense to carry the burden of proof that they have complied with the requirements of Article XXIV GATT.[28]

The object and purpose of the agreement. The purpose and the object of the WTO are, of course, to liberalize international trade. Such liberalization should occur on a nondiscriminatory basis. Any deviation, consequently, from the nondiscrimination principle must be well justified and should be kept to the minimum possible so as not to undo the basic purpose and object of the GATT.

The complementarity between nondiscrimination and regional trade is, as already stated above, institutionally acknowledged in Article XXIV GATT: pursuance of one goal does not necessarily make the other redundant; if this were the case, then the interpreter would have a hard time to achieve an effective treaty interpretation in accordance with the VCLT.

The object and purpose of the Agreement consequently, because of the complementarity of regional and multilateral trade, do not offer any further guidance as to the exact meaning of the term.

Subsequent decisions. By this term, we understand any formal amendment or interpretative note concluded between the WTO Members regarding the interpretation of the term. No such decisions exist. Throughout GATT history, the GATT CONTRACTING PARTIES never adopted an interpretation of the term in this direction. Unadopted proposals that have been at times advanced do not have any legal significance for the purposes of this work.

Subsequent practice. Subsequent practice in this context is rich. By subsequent practice, we understand Working Parties' reports that have examined throughout the years the compatibility of notified PTAs with Article XXIV GATT.

There is no doubt (at least not for the time being) that the CONTRACTING PARTIES when they adopted a report in the context of an examination of compatibility of a notified PTA with Article XXIV GATT did so under the authority vested to them in accordance with Article XXV GATT. Consequently, Working Party reports adopted in the context of Article XXIV GATT constitute subsequent practice in accordance with Article 31 VCLT.

Maybe the most appropriate way to sum up practice in this field is offered by the Working Party report on "EC – Agreements with Portugal"[29] where the EC noted "there is no exact definition of the expression referring to the term 'substantially all trade.' " It is true. There is no such exact definition. The opinion has been expressed that it is "inappropriate to fix a general figure of the percentage of trade which would be subjected to internal barriers;"[30] it should be noted that in the same Working Party, EC Members expressed the view that "a free-trade area should be considered as having been achieved for substantially all trade when the volume of liberalized trade reached 80 percent of total trade."[31]

The Working Party report on "EFTA" on the other hand, records the view that "the percentage of trade covered, even if it were established to be 90 percent, was not considered to be the only factor to be taken into account."[32] Other Working Party reports reflect the view that the exclusion of a whole sector, no matter what percentage of current trade is contrary to the spirit of both Article XXIV and the GATT.[33]

Nothing changed in this respect in more recent years.[34] The GATT Analytical Index (vol. 2) on p. 824 in footnote 162 provides an exhaustive list of Working Party reports dealing with this issue: the outcome is still the same; the term under examination has not been clearly defined in relevant GATT practice.

Very recently, Australia tabled a proposal on the clarification of the term which is at least worth discussing.[35] Australia parts company with the usually mentioned but hardly defined concept that "substantially all trade" has both a quantitative and a qualitative element. In Australia's view, there is only a quantitative element that can come under "substantially all trade" and future negotiations should concentrate on putting a number next to the concept.

Australia specifically proposed that "substantially all trade" should be defined as coverage by an FTA or an agreement establishing a customs union of 95 percent of all the six-digit tariff lines listed in the Harmonized System. In its responses to questions by other WTO Members,[36] Australia accepted that the 95 percent figure is an arbitrary figure but intended to move negotiations out of a deadlock and provide a rule of thumb. Australia was also mindful of the fact that in case trade is concentrated in only a few products, the 95 percent figure could exempt sizeable trade flows. This is why it also proposed an assessment of prospective trade flows under an arrangement at various stages.

If accepted, this proposal would replace the never-ending discussion on the meaning of "substantially all trade" through a "rule of thumb." So far however, it has not provoked any meaningful discussion.

Preparatory work. Having exhausted recourse to the interpretative means laid down in the body of Article 31 VCLT one will have to concede that a general tendency is recognizable, but that at the end of the day the meaning of the term is "obscure" and that therefore recourse to supplementary means of interpretation is warranted.

Unfortunately however, preparatory work is not much of a help. The WTO CRTA proceeded to discuss the meaning of the term. In a series of papers that the WTO Secretariat prepared for the Committee,[37] the conclusion was inescapable: the drafters of the treaty used the term loosely. This however, does not mean that the issue is not justiciable. Working Parties have in a way already contributed through their practice: we

know that many of the Working Parties have flirted with the idea of a quantitative threshold. We also know that at least some of them went even further and requested that no sector of the economy be excluded as such, independently of its impact on the intra-PTA trade.

The problem is that Working Parties have not provided a precise meaning to the term that can be used as a "rule of thumb" in future cases. As already noted above, it is questionable whether outsiders, except in particular circumstances, ever had the incentive to enforce the internal requirement. Since no one had the incentive to play "international watchdog" – and in the absence of a meaningful WTO *ex officio* review power in this context – unsurprisingly, the precise meaning of the term remains elusive.

Is a clarification necessary? The answer depends on how much WTO Membership trusts its adjudicatory bodies (more on this issue, in Section 5). A clarification would limit the discretion of both the CRTA and subsequent panels dealing with this issue and, from a legal security perspective, is very welcome. It would presumably at the same time though outlaw prospective PTAs very comparable to those that have already passed the Article XXIV-consistency test. Confusion in this respect is welcome by those aspiring to participate in a PTA in the future and at the same time does not put into question past behavior: both incumbents[38] and new entrants can profit. And this probably explains why not much progress has been noted in the context of the CRTA in this respect.

3.5.2 Other Restrictive Regulations of Commerce

The ordinary meaning of the term. Article XXIV.8 GATT outlaws restrictive regulations of commerce (other than those embodied in Articles XI, XII, XIII, XIV, XV, and XX GATT) between constituencies of a PTA (either FTA or CU as XXIV.8(a) and XXIV.8(b) make clear). The term "restrictive regulations of commerce" is not further precised. There are however, interpretative tools that can help define it in a more clear way:

both paras (a) and (b) refer to *duties* and other restrictive regulations of commerce. This should consequently, be understood as any regulation which has a restrictive effect (like duties do). In principle, *any* regulation restrictive of commerce is captured;[39]

both paras (a) and (b) refer to duties and *other* restrictive regulations of commerce. This language suggests that *any* other restrictive regulation of commerce, without any additional qualification, is captured. An interpreter, consequently, would have to undertake a two-steps procedure: first see whether a regulation restricts commerce and second, outlaw it unless it can justify it through invocation of one of the exceptions mentioned in the first paragraph of this subsection;

the exceptions mentioned in the parenthesis (XI, XII, XIII, XIV, XV, and XX) underscore this point. Protection can occur essentially in two ways: either through duties or through quantitative restrictions (QRs). Duties are anyway outlawed. QRs as well, since XI contains the general prohibition to have recourse to such form of protection.

The ordinary meaning of the term as a consequence, seems to outlaw any regulation which restricts commerce (other than those mentioned in the parenthesis).

The context. The items mentioned in the parenthesis pose interpretative problems. In fact, what poses problems is what is not mentioned in the parenthesis. Does this mean that other restrictive regulations of commerce (like VI, XVIII, XIX, or XXI) cannot be invoked, or is it simply a regrettable omission?

On its face, it seems that anything not mentioned in the parenthesis is not tolerated after the formation of a CU or FTA. This would essentially mean that no antidumping action can take place between members of an FTA, or that for an FTA to be judged compatible with Article XXIV.8 it must abolish imposition of antidumping duties between its members. To our knowledge however, only two of the numerous PTAs (the EC and the ANZCERTA) have eliminated antidumping duties between their members.[40] Notorious PTAs (like NAFTA) continue to provide for the possibility that one member can impose antidumping duties against another.

Moreover, whereas good arguments could be advanced with respect to the elimination of antidumping duties or safeguard actions among members of a PTA, this is hardly the case with respect to measures based on security exceptions. In fact, it is not sustainable to argue that just by agreeing to form an FTA, members of an FTA acknowledge that their security would never again be threatened by actions of their partners.

The context hence rather perplexes the issue. It would seem to support the view that the items mentioned in the parenthesis cannot be part of an exhaustive list. Mavroidis (1997) has argued that even if the list is deemed to be indicative, restrictive regulations of commerce for economic motives cannot be added to what already exists. As will be shown later, this view has been overturned by the Appellate Body, at least with respect to safeguards.

With respect to restrictive regulations of commerce adopted for noneconomic motives, the list should not be viewed as exhaustive: except for Article XXI GATT considerations, one could well imagine that a WTO Member of a PTA might be willing to block trade from another member of its PTA if the latter violates *jus cogens*. Whether or not *jus cogens* is reflected in Article XX(a) GATT is an open issue.[41] What cannot be put into question though is the validity of Articles 53 and 63 VCLT which make it plain that a treaty in contravention of *jus cogens* is null and void.

Subsequent agreements. The WTO Members following the pragmatic attitude of the GATT CONTRACTING PARTIES did not proceed to a clarification of this point. As mentioned above, the WTO Understanding on Article XXIV GATT does not contain any interpretations of the terms included in the body of Article XXIV.8.

Subsequent practice. Subsequent practice is an illustration of divergent opinions. Starting with the 1970 Working Party report on "EEC – Association with African and Malgasy States" one notes, on the one hand, the opinion of members of the Working Party, which were not members to the notified PTA, to the effect that "free trade within the meaning of Article XXIV.8(b) did not exist" in view of the continued imposition by certain parties to the Convention of fiscal charges on imports from other members; at the

same time the members of the PTA were declaring that "the provisions of Article XXIV, concerning the concept of a free-trade area concerned only protective measures. The taxes referred to were of a fiscal character, not protective . . . ."[42]

There were also discussions on the issue whether Article XXI GATT should figure in the body of Article XXIV GATT in the Working Party report on the EEC. The view of the EEC Member States was reflected that "it would be difficult, however, to dispute the right of contracting parties to avail themselves of that provision which related, *inter alia*, to traffic in arms, fissionable materials, etc., and it must therefore be concluded that the list was not exhaustive."[43] Other Working Party reports reflect a series of discussions on this issue.[44]

During the Uruguay Round negotiations, a draft decision was tabled to clarify the relationship of Article XIX GATT with Article XXIV GATT. It read:

> When an Article XIX action is taken by a member of a customs union or free-trade area, or by the customs union on behalf of a member, it [need not] [shall not] be applied to other members of the customs union or free-trade area. However, when taking such action it should be demonstrated that the serious injury giving rise to the invocation of Article XIX is caused by imports from non-members; any injury deriving from imports from other members of the customs union or free-trade area shall not be taken into account in justifying the Article XIX action.[45]

Had this proposal been accepted it would have, on the one hand, provided a much needed clarification in this field but most importantly would have led to a more general discussion on what is and what is not permitted between members of a PTA. The proposal was rejected and the possibility to engage in a meaningful discussion in this field became a remote one.

In fact, subsequent practice has developed in the contrary way. On two occasions, first the Appellate Body and then a panel subsequently dealing with the same issue were faced with the question whether a member of a PTA (a CU in the case of Argentina[46] and an FTA in the case of the United States[47]) could impose safeguards against other members of the PTA. It should be kept in mind that on both occasions, the WTO adjudicating bodies were requested to interpret the WTO Safeguards Agreement. However, through their interpretation they were clearly "influencing" the question whether the items figuring in the parenthesis of Article XXIV.8 GATT form an exhaustive list or not.

On both occasions, the WTO adjudicating bodies held for the proposition that members of a PTA can impose safeguards against other members of a PTA provided that they respect a parallelism: they can do so, if they have counted PTA-imports when assessing injury; they cannot do so, if they have not counted PTA-imports when assessing injury.

Following these events, it is clear now that the list in the parenthesis of Article XXIV.8 GATT is not an exhaustive one. Safeguards for sure are now included.[48] What else however is included? Should one by analogy be brave enough and apply the stated reasoning in all forms of "contingent protection?" Only future experience will tell. The odds, however, are that WTO adjudicating bodies are willing to adopt a "pragmatic" attitude and redesign Article XXIV GATT according to modern reality. A textualist approach which does not keep up with modernity[49] is certainly not the one envisaged by the VCLT system. Subsequent practice is probably the single most important factor

mentioned in Article 31 VCLT. Subsequent practice in the present discussion suggests that one could well imagine other forms of contingent protection to be soon added to the list figuring in the parenthesis of Article XXIV.8 GATT.

Preparatory work. Recourse to preparatory work is meaningful here only to see whether the precise scope of the term is anywhere reflected in the records. This is not the case though. Hence, the conclusion that we can draw with respect to the term "restrictive regulations of commerce" is twofold:

that in principle, any regulation of commerce which can restrict trade is caught by the standard reflected in Article XXIV.8 GATT;

that, on the other hand however, the current tendency of WTO adjudicating bodies is to interpret Article XXIV.8 GATT in light of practice which arguably evolved against the unambiguous letter of Article XXIV.8 GATT. Hence, by extending the list of items figuring in the parenthesis of Article XXIV.8 GATT, WTO adjudicating bodies show a "conciliatory attitude" *vis-à-vis* reality and their willingness to keep up with modern times. The impact of course will be an easier to meet standard for aspiring PTAs.[50]

4 The GATS Regime

Article V GATS seems to set a looser standard than Article XXIV GATT. It essentially imposes two conditions for a PTA to be judged WTO-consistent:

(a) that it has substantial sectoral coverage; and
(b) that it abolishes discrimination.

According to a footnote to the text of Article V GATS, the term "substantial sectoral coverage" is to be understood "in terms of number of sectors, volume of trade affected and modes of supply. In order to meet this condition, agreements should not provide for the *a priori* exclusion of any mode of supply."

This approach is not as odd as it might seem at first sight. Liberalization in the GATS context is not of the same width and breadth as in GATT. It probably was considered politically untenable to request, from WTO Members aspiring to enter into PTAs, to include sectors that they have not previously liberalized on an MFN-basis. On the other hand, had such a requirement been introduced, regionalism in GATS could prove to be a vehicle for liberalization.[51]

In the absence of relevant practice,[52] it seems reasonable to conclude therefore that while some sectors can be excluded without putting into question the consistency of the PTA at hand with Article V GATS, no exclusion of mode of supply is permitted for the sectors covered.[53]

5 The Legal Nature of a CRTA Report

One would expect that, taking into account the ample powers attributed to Working Parties and now the CRTA by virtue of Article XXIV.7 GATT, a Working Party (CRTA)

report would operate like a "green light" for all notified PTAs. Practice however, shows a different trend. Never has a notified PTA been rejected by the GATT CONTRACTING PARTIES even in some rather extreme cases (like the very recently notified CU between EC and Turkey which fails to meet a series of requirements established by Article XXIV GATT). Lack of unanimity however, did not prevent adoption. And adopted panel reports usually, that is with the exception of the five reports mentioned *supra* where it was by consensus concluded that the notified PTAs met the substantive and procedural requirements of Article XXIV GATT, represent a variety of opinions (very often conflicting) on the same issue.

In essence, Working Party (and now CRTA) reports traditionally look more like inventories for future disputes. Some Members will complain about issue x, others about issue y, thus signaling to the members of the notified PTA what kind of disputes they should be expecting in the future.

There are even reported cases of Working Parties withholding judgment.[54] The 1985 panel report on "EC – Tariff Treatment on Imports of Citrus Products" (which remains unadopted) noted, *inter alia*, that:

> . . . the CONTRACTING PARTIES had withheld judgment . . . the question of the conformity of the agreements with the requirements of Article XXIV and their legal status remained open.[55]

This did not prevent however, the EC enlargement to the North in 1973 and the South in 1981 and 1986. Practice shows therefore that Working Parties have not functioned as the necessary "green light" for a PTA to go through.

What is then the legal nature of a Working Party report? In the WTO era, it is clear that WTO Members will adopt CRTA reports in conformity with Article IX of the Agreement Establishing the WTO the presented reports. As far as the already adopted GATT reports are concerned, it seems plausible to argue that they form part of GATT 1994, since they should come under "other decisions of the CONTRACTING PARTIES to GATT 1947" in accordance with Article 1(b)(iv) GATT 1994. If this is the case, then they should be construed as subsequent practice (if not subsequent agreement) of the parties to the original agreement relevant for the interpretation of Article XXIV GATT. But as already stated above, they are not helpful at all since they usually reflect an ongoing disagreement of the parties to the GATT as to the interpretation of the terms included in the body of Article XXIV GATT.

What the GATT CONTRACTING PARTIES have adopted in this context (and what the WTO Members will most likely be adopting in the future) is a disagreement. Unfortunately, there is no consistency as far as the subjects of the disagreement are concerned either, since, depending on whether they participate or not in a PTA, WTO Members adopt a relaxed or aggressive attitude as far as the interpretation of Article XXIV GATT is concerned.

At the end of the day, one could state that Working Party reports formally are decisions with limited (if not totally absent) practical relevance: they are typically inventories of potential future disputes.

6 The Interplay between the Two Tracks to Establish Consistency of PTAs: GATT Working Parties; They almost always came with nothing at the End

Upon notification, a Working Party would be established, the Terms of Reference of which were:

> to examine the Agreement on the Establishment of a . . . between . . . and . . . signed on . . ., in the light of the relevant provisions of the General Agreement, and to report to the Council.

Sometimes there would be an explicit subtitle in the report on the compatibility of the notified scheme with the GATT[56] and sometimes questions relating to the compatibility of the notified scheme would be raised under headings reflecting the essence of Articles XXIV.5 and XXIV.8 GATT.[57] There was no doubt in anyone's mind however that the task of such Working Parties was to examine the consistency of the notified scheme with the GATT rules.

And this in itself is an important observation. Independently of whether the Working Party's review should be viewed as an exclusive forum, it is important to note already that the GATT architecture and practice called for multilateral scrutiny of PTAs.

PTAs amount of course to a permanent waiver from the most-favored-nation clause (MFN), the cornerstone of the GATT edifice. Indeed the carrot for the "new kids in the block" was essentially MFN. One should not forget that the GATT started with 23 members only and that its major attraction for outsiders was the guarantee that accession equaled nondiscriminatory access to a sizeable number of markets.

The GATT Working Parties adopted a very pragmatic attitude toward PTAs. Pragmatic is a term that needs further definition. In essence, it means that the attitude adopted is probably below the benchmark intended by the drafters of the Treaty. Pragmatic in this sense means tolerant. And the GATT showed a lot of tolerance *vis-à-vis* PTAs, up to the point that, before the renewed confidence in the multilateral trading system with the coming-into-being of the WTO, intense skepticism reigned in the mind of scholars as to the direction of the GATT. Hudec (1972, p. 1362) notes:

> The seeming collapse of the MFN rules is probably the single most important cause of the present day pessimism about the GATT substantive rules.

Hudec twenty years later (1993, p. 154) remarks:

> the GATT's somewhat benign attitude toward RAs is merely one part of this larger tolerance toward departures from MFN in general.

Roessler (1993, p. 321) seems in agreement:

> The record under the current procedures is not encouraging. During the past three decades about 50 working parties have been established to examine RIAs. None of them was able to reach a unanimous conclusion on the GATT-consistency of the agreement examined . . ..

RAs and RIAs (regional agreements and regional integration agreements) are shorthand for PTAs. It is submitted that both scholars provide an accurate description of the GATT-era: Working Parties would be established, sometimes faced with a *fait accompli*, to routinely prepare a report which would reflect a number of disagreements which would eventually lead to disputes, albeit not so often.

And at the end, their adopted report reflected a consensus to disagree on the consistency of the reviewed PTA (an inventory of future disputes, as we noted *supra*) or, as Roessler (1993, p. 321) mentions, "the legal status of all these agreements has therefore remained undetermined." In fact, the dead end to which Working Parties routinely arrived probably explains the frustration of the Chairman of the Working Party examining the consistency of the Canada/United States FTA who asks:

> . . . what point was there in establishing a working party if no-one expected it to reach consensus findings in respect of specific provisions of such agreements, or to recommend to the participants how to meet certain benchmarks. . . . As further agreements came along, there might be a risk that they would be treated increasingly superficially and that contracting parties would lose – if they had not already done so – the ability to distinguish between agreements of greater or lesser GATT-consistency.[58]

Precisely because challenging the conformity of a PTA with the GATT rules is not always an incentive-compatible structure for other GATT parties,[59] the GATT dispute settlement system was not called often enough to finish the "unfinished business." We already alluded that the incentive for nonparticipants to enforce at least the internal requirement was more often than not missing. Schott (1989, p. 25) offers another reason why the incentive was very often not there:

> Besides the ambiguity of its provisions, political considerations have often outweighed other factors in decisions to accede to the terms of the agreements. In addition, affected third countries have been reticent to criticize preferential deals because the majority of GATT members participate in such agreements.

As a result, at the end of the GATT years we were left with dozens of PTAs the status of which under the WTO contract is uncertain. Moreover – although good arguments can be advanced in the opposite direction – it is legally questionable whether a WTO Member can now challenge a PTA that the GATT/WTO system "tolerated" for a good number of years.

As Roessler (1993, p. 321) points out, the need to reach consensus for each and every decision in the GATT years is probably in itself the basic reason explaining why the dead end was created in the first place. And the GATT-version of consensus was a perverse one: contrary to the maxim known in almost each and every legal system *nemo judex in sua causa esse potest*, the agreement of notifying GATT parties to a PTA was essential for consensus to be achieved.

The need for consensus probably explains why no massive requests for panels were tabled. After all without the agreement of the other party, no panel would be established, no report would be adopted, no countermeasures would be authorized.

7 GATT Panels: Not Much Practice

As to the justiciability of Article XXIV GATT-related claims, the following seems pertinent. Three panels were established to examine such claims.[60] Two reports were issued and they both remain unadopted. The first report from this experience is the *EC – Citrus* panel report which holds for the proposition that GATT panels can examine individual measures but not the overall consistency of a PTA with the multilateral rules. The relevant passage of the report reads:

> The Panel noted that at the time of the examination of the agreements entered into by the European Community with certain Mediterranean countries, there was no consensus among contracting parties as to the conformity of the agreement with Article XXIV.5 . . . The agreements had not been disapproved, nor had they been approved. The Panel found therefore that the question of conformity of the agreements with the requirements of Article XXIV and their legal status remained open.[61]

This report, as noted, remains unadopted, and hence, of limited legal relevance.[62]

The second report is the mentioned *EEC – Bananas* report, which essentially made the point that one-way preferential arrangements are *per se* inconsistent with Article XXIV GATT; obligations to liberalize must be assumed by all participants. Para. 159 of the report relevantly reads in this respect:

> This lack of *any* obligation of the sixty-nine ACP countries to dismantle their trade barriers, and the acceptance of an obligation to remove trade barriers only on imports into the customs territory of the EEC, made the trade arrangements set out in the Convention substantially different from those of a free trade area, as defined in Article XXIV:8(b).

Unsurprisingly, the panel went on to conclude in para. 164 that the Lome Convention did not correspond to the type of agreements which Article XXIV GATT covers. This report as well remains unadopted and, although the view expressed in the cited passage is sound, the legal value of the report is minimal.

8 The GATT-Era: PTAs Tolerated

GATT Working Parties (with the four exceptions mentioned in Schott 1989) did not end up with a final yes or no on the consistency of notified PTAs. Panels, although in principle competent to adjudicate PTA-related disputes were requested only twice to do so. With respect to the competence of panels to adjudicate such disputes there can be no opposite opinion. Not only the practice mentioned suggests that this was the case but the view of the delegate of the EC, the absolute champion in numerical terms of PTAs, is recorded stating in the 1978 Working Party report on the Agreement between the EEC and Egypt that

> . . . as regards the possibility of consultations with the contracting parties concerning the incidence of the Agreement on their trade interests . . . nothing prevented these countries from invoking the relevant provisions of the General Agreement, such as Articles XXII and XXIII.[63]

We can hence conclude that with respect to the GATT era:

there is an agreement among institutional players (contracting parties and GATT panels) that PTAs can be reviewed by GATT panels;

there is no agreement as to the extent of judicial review in such cases;

and that the two-track system (notification to Working Parties and eventually submission of disputes to GATT panels – with the noted disagreement concerning the ambit of judicial review) can be proud of a very unsatisfactory record.

Now, what was the WTO response to the situation described above?

9 Tracks I and II in the WTO Phase: Possibly Less Tolerance for PTAs?

When the WTO contract was negotiated, negotiators no doubt were well aware of the deficiencies of the two-track system. In the GATT-era consensus was needed for a Working Party to pronounce against the consistency of a PTA with the GATT rules; consensus, however, was needed for a panel to be established to adjudicate a PTA-related claim (and, of course, for its report to be adopted, countermeasures to be authorized in case of nonimplementation).

Dispute settlement, on the other hand, knows of one additional hurdle: incentives. In a decentralized system of enforcement, it is quite hard to imagine WTO Members substituting their own individual claims with public interest-based claims. The *ex officio* part of the review would unavoidably be left with the multilateral bodies.

So what was the response of the drafters of the WTO contract? A drastic renegotiation of dispute settlement where practically all hurdles were removed; an explicit directive to all concerned that Article XXIV GATT-related claims are justiciable in a new hurdles-free system; and a mundane embellishment of the other track without however modifying at all the picture in this respect. We explain.

9.1 Dispute Settlement in the WTO

As explained elsewhere in this volume, the new dispute settlement system operates in a negative consensus-mode: panels will be established if a potential plaintiff so wishes, reports will be adopted, if a party to the dispute is happy with the outcome and countermeasures will be authorized if the winning party is not happy with the implementing actions of the party found to have violated its WTO obligations. All of that, automatically.

Hence, the fear that the establishment of the panel or of its report being adopted or of an eventual request for countermeasures be blocked by the other party (indeed by the party which has the incentive to behave in a noncooperative manner) is eliminated. WTO Members now know that by submitting to the WTO dispute settlement system, they will end up with an outcome; favorable, or unfavorable, an outcome anyway.

9.2 The Understanding on Article XXIV GATT

Along with the other Uruguay Round agreements, negotiators concluded and adopted an Understanding on the Interpretation of Article XXIV of the GATT 1994. The § 12 of the Understanding pertinently reads:

> The provisions of Articles XXII and XXIII of GATT 1994 as elaborated and applied by the Dispute Settlement Understanding may be invoked with respect to *any* matters arising from the application of those provisions of Article XXIV relating to customs unions, free-trade areas or interim agreements leading to the formation of a customs union or a free-trade area (emphasis added).

Hence, on its face, the Understanding seems to grant a large review power to WTO adjudicating bodies: their power extends to any matter arising from the application of Article XXIV GATT. Roessler (2000) argues that the terms of the Understanding lean toward a restrictive understanding of its scope: the reference made is to the application of Article XXIV and not to Article XXIV GATT as such. This argument does not sufficiently address though that the reference is to Article XXIV in a horizontal manner and not to a particular CU or FTA. Application of Article XXIV GATT from a horizontal perspective, covers anything that can come under the ambit of the mentioned Article and is not case-specific constrained.

"Any matter relating to the application of Article XXIV GATT" thus, signals the willingness of negotiators to affirm their intention to see WTO panels deal with PTA-related issues. Note that, although GATT practice as noticed *supra* suggests that panels are competent to adjudicate PTA-related claims, this is the first time in GATT history that such an acknowledgment is made explicitly. And note further, that this acknowledgment takes place in an era where there is no consensus-hurdle to establish panels.

9.3 The Understanding, the CRTA and the Art of Cosmetic Changes

The Understanding adds nothing to what we already knew about the functioning of Working Parties in the GATT-era. The §§ 7–11 of the Understanding reflect some trivial transparency requirements but do not touch at all upon the major source of dysfunction of the Working Parties: the consensus rule.

In this respect there is an imbalance, a voluntary one, in the Understanding: first, this is the explicit institutional acknowledgment in the history of the GATT/WTO that adjudicating bodies can review Article XXIV GATT-related issues. In a sense, in this respect, the Understanding codifies the limited GATT practice but also the underlying consensus. Second, it makes it clear that one track of the review (the Working Party track) will remain as it is, that is, to a large extent, an inefficient means to review PTAs; whereas the other track of the review (dispute settlement) has undergone substantial modification and its efficiency (the move to negative consensus) has been greatly enhanced.

Soon after the entry into force of the WTO contract, the first track (Working Party) underwent, as noted above, another cosmetic type change: through a decision of February 7, 1996, the WTO General Council established CRTA. As a result, instead of establishing a Working Party each time a PTA is notified, a permanent forum was created which would entertain all such notifications.

The rules of the game did not change though. The passage from positive to negative consensus was limited to dispute settlement. The CRTA would continue to operate under the consensus-rule, the source of all its misfortunes in the past. So although consolidating all Working Parties to one body, the CRTA, maybe contributes to making the latter more accountable, consensus continues to be a major hurdle to reach concrete outcomes when reviewing PTAs in Track I.

It took little time for WTO practice in the context of CRTA to prove that absent a drastic rethinking of the consensus rule, not much change should be expected from the passage of Working Parties to the CRTA. At the time of writing, there is no decision on any of the notified PTAs so far. The mentioned report on the CU between the Czech and the Slovak Republics was issued on July 15, 1994 and thus predates the CRTA.

So what is the WTO response in a nutshell? The mandate is quite clear: the track that was tried and failed in the past is not revamped. It remains as it is for the time being at least. The track that has not been tried becomes an efficient tool in the hands of potential users and it has explicit jurisdiction over PTA-related issues. It remains to be seen whether the message will reach interested ears and how practice will evolve.

9.4 Did We Get the Message? Relevant Practice in WTO

And the message was heard. Maybe not with the intensity that one could have anticipated. Still whereas in the 40 GATT years only one PTA-related dispute was submitted for adjudication, it took only three years to see an equivalent case find its way before the WTO.

The *Turkey – Textiles* panel report[64] records the view that WTO adjudicating bodies are competent to examine PTA-related issues but should stop short from providing an overall assessment of consistency of a PTA with the WTO contract. The view of the panel on this issue is reflected in paras 9.52 and 9.53 of the report. We quote:

> As to the second question of how far-reaching a panel's examination should be of the regional trade agreement underlying the challenged measure, we note that the Committee on Regional Trade Agreements (CRTA) has been established, *inter alia*, to assess the GATT/WTO compatibility of regional trade agreements entered into by Members, a very complex undertaking which involves consideration by the CRTA, from the economic, legal and political perspectives of different Members, of the numerous facets of a regional trade agreement in relation to the provisions of the WTO. It appears to us that the issue regarding the GATT/WTO compatibility of a customs union, as such, is generally a matter for the CRTA since, as noted above, it involves a broad multilateral assessment of any such custom union, i.e. a matter which concerns the WTO membership as a whole.

As to whether panels also have the jurisdiction to assess the overall WTO compatibility of a customs union, we recall that the Appellate Body stated that the terms of reference of panels must refer explicitly to the "measures" to be examined by panels. We consider that regional trade agreements may contain numerous measures, all of which could potentially be examined by panels, before, during or after the CRTA examination, if the requirements laid down in the DSU are met. However, it is arguable that a customs union (or a free-trade area) as a whole would logically not be a "measure" as such, subject to challenge under the DSU (footnotes omitted).

In the panel's view, for reasons having to do more with the administrative burden, the CRTA is the more appropriate forum to review consistency of notified PTAs. The panel report was appealed. The Appellate Body report holds for a different proposition. In its view, the Article XXIV GATT defense holds only if two conditions are met:

First, the party claiming the benefit of this defence must demonstrate that the measure at issue is introduced upon the formation of a customs union that fully meets the requirements of sub-paragraph 8(a) and 5(a) of Article XXIV. And second, that party must demonstrate that the formation of that customs union would be prevented if it were not allowed to introduce the measure at issue. . . . We would expect a panel, when examining such a measure, to require a party to establish that both of these conditions have been fulfilled. It may not always be possible to determine whether the second of the two conditions has been fulfilled without initially determining whether the first condition has been fulfilled.[65]

In the Appellate Body's view WTO adjudicating bodies must request from parties raising the PTA-defense to first establish that they have fulfilled the conditions to raise such defense. Such an approach would be in perfect symmetry with the maxim *quicunque exceptio invokat, ejusdem probare debet.* To what extent the cited *obiter dictum* of the Appellate Body will be followed in future experience, remains of course to be seen.

More recently, the panel report on *United States – Definitive Safeguard Measures On Imports Of Circular Welded Carbon Quality Line Pipe From Korea* (WTO Doc. WT/DS202/R of October 29, 2001) faced an argument by the United States that, since it is a member of NAFTA, it was entitled to treat imports from NAFTA different than imports from non-NAFTA sources when imposing a tariff quota. The panel first addressed the issue of burden of proof:

As the party seeking to rely on an Article XXIV defense . . . the onus is on the United States to demonstrate compliance with these conditions (§ 7.142 of the report).

The same report addressed the issue of the **quantum of proof** that the party carrying the burden of proof has to provide in order to establish a *prima facie* case of consistency of a PTA with the multilateral rules. § 7.144 of the report reads in this respect:

In our view, the information provided by the United States in these proceedings, the information submitted by the NAFTA parties to the Committee on Regional Trade Agreements ("CRTA") (which the United States has incorporated into its submissions to the Panel by reference), and the absence of effective refutation by Korea, establishes a *prima facie* case that NAFTA is in conformity with Article XXIV:5(b) and (c), and with Article XXIV:8(b).

The information provided by the United States in the proceedings is reflected in § 7.142 of the report and essentially is exhausted in a statement that duties on 97 percent of the NAFTA-parties' tariff lines will be eliminated within 10 years, whereas with respect to "other regulations of commerce" a reference to "the principles of national treatment, transparency, and a variety of other market access rules" is made.

As to the legal significance of the fact that the CRTA had not issued its report at the time the dispute was submitted to the panel, § 7.144 reads:

> Concerning Article XXVIII:8(b), we do not consider the fact that the CRTA has not yet issued a final decision that NAFTA is in compliance with Article XXIV:8 is sufficient to rebut the *prima facie* case established by the United States. Korea's argument is based on the premise that a regional trade arrangement is presumed inconsistent with Article XXIV until the CRTA makes a determination to the contrary. We see no basis for such a premise in the relevant provisions of the Agreements Establishing the WTO.

9.5 An Inventory of the Modern: Institutional Balance in the WTO

So where do we stand now? There seems to be a disagreement between the approach advocated by the Panel and the Appellate Body in *Turkey – Textiles* as to the extent of judicial review. As noted however, the panel report in *Turkey – Textiles* echoes the findings of the *EC – Citrus* panel report and argues in favor of an, in principle, limited judicial review of PTAs by panels for reasons having to do with the administrative burden rather than with anything else.

Roessler (2000) has recently argued that, for reasons having to do with the institutional balance of the WTO, a limited judicial review by WTO adjudicating bodies is the most appropriate one. The institutional balance of course, has been established by the founders of the WTO.

Negotiators, even if they wanted to reinforce the CRTA-track, would undoubtedly have found it difficult to do so during the Uruguay Round negotiations. For the single most formidable obstacle to reinforcement is consensus and it would be quite difficult to contain its removal within the four corners of CRTA without any spillovers in the rest of the WTO contract.

On the other hand, the argument could be advanced that negotiators were quite happy with an institutional balance which calls for a very benign control of PTAs by the intergovernmental body (currently, the CRTA) and a limited (but not precisely specified) control by adjudicating bodies. Such an argument must be rejected.

As the cited passage from Roessler (1993) shows, there was dissatisfaction with the benign attitude. Moreover, if such dissatisfaction suffers from selection bias (not every Chairman of an Article XXIV GATT Working Party made similar comments), it is still difficult to explain why negotiators added § 12 to the WTO Understanding on Article XXIV GATT especially after having agreed on the passage to negative consensus in the context of dispute settlement.

The one argument in favor of strengthening the (now) CRTA-track is of course, the comprehensiveness of the endeavor in the context of CRTA. Time constraints alone

might prove an obstacle for panels wishing to follow the Appellate Body's *obiter dictum* in the *Turkey – Textiles* case and opt for a comprehensive review of PTAs. And of course, it would be ideal for WTO adjudicating bodies to be in a position to rely on acts of other WTO institutions and use them in their findings. Their review would be reduced and they would probably be in a position to avoid controversial issues. The problem is though that in practice such acts simply do not come by. The CRTA does not decide like the GATT Article XXIV Working Parties never did. Moreover, nothing in the WTO contract obliges WTO adjudicating bodies from refraining in PTA-related issues. If at all, they must refrain on their own initiative. But refrain from what? History teaches that, in PTA-related issues, they are confronted with a vacuum. They do not overturn a Working Party decision. They simply rule for the first time.

The reasonable interpretation of the Uruguay Round Agreements is that negotiators, even if they believed that a CRTA-type control is optimal, knew that it could hardly prove effective in practice because removal of consensus was a costly option. If they wanted to keep the *status quo*, then they did not need to enact the Understanding. Its enactment proves that they shared the belief that review of PTAs so far had been less than satisfactory and that they believed that the only feasible way to correct this distortion was through inciting use of dispute settlement procedures.

So at the end one can reasonably conclude that WTO Members agreed to tilt the institutional balance in favor of Track II. Are WTO panels well equipped to live up to the expectations of the WTO Membership? After the Uruguay Round Understanding on Article XXIV GATT, it seems that the only contentious, as far as interpretation is concerned, issue is the "internal requirement" and, more precisely, what is covered by the term "substantially all trade." With respect to Article XXIV.5 GATT, there is now enough normative guidance for panels to always reach coherent outcomes.

As mentioned above, WTO Members have failed to reach an agreement on the "internal requirement" issue. Are panels dealing with this issue – assuming that they from now religiously follow the mentioned *obiter dictum* in the Appellate Body report on *Turkey – Textiles* and request WTO Members to first show consistency of their PTA with WTO rules before they can avail themselves of the possibility offered in Article XXIV GATT – prejudging future action by the CRTA?

Not at all. The WTO Membership can at any time adopt an interpretation of the term which will restrict the discretion of future panels dealing with the issue. Such interpretation cannot be applied retroactively and hence what panels dealing previously with the issue ended up with is simply not a concern for the WTO Membership. Of course, panels' practice could provide useful input for any future WTO authoritative interpretation.

This issue is maybe more problematic if viewed on an *ad hoc* basis: for the sake of the argument imagine that the same PTA is simultaneously before a panel and the CRTA. If the CRTA by consensus concludes on the consistency/inconsistency of the notified PTA with the multilateral rules (a rather unrealistic scenario in view of the time limits of the CRTA review) there is good reason to believe that the panel subsequently dealing with the issue will follow the opinion reflected in the CRTA. The panel report on *India – Quantitative Restrictions on Imports of Agricultural, Textile and Industrial Products*[66] which dealt with a similar issue (i.e., to what extent a panel dealing with

an issue where the Balance of Payments Committee has already decided) relevantly provides in para. 5.94:

> . . . we see no reason to assume that the panel would not appropriately take those conclusions into account. If the nature of the conclusions were binding . . . a panel should respect them.

The same should be true for panels dealing with PTA-related issues. This of course raises the question whether panels should stop short of deciding PTA-related disputes in case the CRTA has not pronounced on the overall consistency of the given PTA. Such a "peace clause" of course does not exist. Panels of course, could through their attitude effectively "insert" such a "peace clause" by exercising judicial restraint. Such an approach is hardly recommendable. For if panels were to indeed exercise restraint and avoid ruling on PTA-related issues (in the name of an ill-defined institutional balance) they would effectively be depriving WTO Members of their MFN-rights: the CRTA will invariably take a long time to reach consensus and the consensus will invariably reflect a disagreement.[67]

Should one view the issue of reviewing a PTA as tantamount to a political question which escapes the judicial review by WTO panels? Are in other ways PTA-related questions not justiciable? Again, the answer should be in the negative.

First, the wording of the WTO Understanding on Article XXIV GATT goes to the opposite direction when it entrusts adjudicating bodies with the possibility to review PTA-related disputes. Second, the only way one could conceivably make the argument that in this context we are dealing with a political question is by highlighting the rationale for going regional. But the rationale for going regional is the one element that panels cannot put into question. The wording of Article XXIV GATT in itself makes the rationale for going regional a nonconcern for subsequent interpretations. In other words, independently of the reasons why a WTO Member decides to opt for regionalism, it must always abide by the standard laid down in Article XXIV GATT. Finally, the question of political doctrine was raised by India in the balance of payments-related dispute mentioned above and rejected by the Appellate Body essentially on textual grounds.[68] Textual grounds argue in favor of a similar outcome in the PTA context as well: whenever drafters of the GATT wanted to reflect in a legal text their intention to see a relaxed standard of review, they did so in unambiguous terms. It is not accidental that the body of Article XXI GATT which deals with the security exception – the only genuine political question – reads in a manner that makes it obvious that the margin of discretion rests primarily with the state invoking the exception. This is not the case of Article XXIV GATT which acknowledges the discretion of multilateral organs to decide whether a given PTA is in conformity with WTO law.

The last question to address is whether panels are likely to be called to evaluate the overall consistency of PTAs in the years to come. It seems rather unlikely for reasons having to do with the incentives of nonparticipants described above. However, submission of a PTA-related dispute to a panel is the only threat for aspiring regional groupings. To what extent the threat will prove to be a credible one will largely depend on the willingness of Members to submit and the use panels will make of such submissions.

10 What is the Law when Dealing with PTAs?

At the end of this chapter the question can legitimately be asked what is the law with respect to PTAs? And the fact that the question in itself deserves to be raised is proof enough that the existing system has failed so far to provide the necessary clarity.

GATT practice reveals widespread tolerance of PTAs: Track I has been very disappointing and Track II has been hardly used. In the WTO-era, there is a willingness to "tighten the screws." This tendency is essentially evidenced at both the legislative and the judiciary level: the interpretation of the "general incidence" requirement through the adoption of the Understanding on Article XXIV GATT; the explicit acknowledgment in the same Understanding that WTO adjudicating bodies are competent to review PTA-related disputes; the attitude of the Appellate Body to effectively monitor compliance as reflected in the cited *obiter dictum* from the *Turkey – Textiles* report.

One should not see too much though in the changes in the attitude that the Uruguay Round Agreements brought about. Track I in the WTO-era continues where GATT-era left off and panels are not flooded with PTA-related cases. And since panels are not flooded with such cases, tolerance of PTAs (irrespective of their consistency or the extent of their inconsistency with WTO rules) continues to be the ruling "paradigm" in the multilateral system.

On the other hand, one should not lose perspective: even the mentioned legal efforts to tame regionalism are not informed by an economics-friendly analysis on the welfare implications of regionalism. In fact, the legal test for consistency of PTAs essentially deals with one question: how to ensure that deviations from MFN will come by harder rather than easier? Whether such deviations maximize or minimize welfare is simply not a concern.

Hence, at the end of the day, the Article XXIV GATT-test is an uninformed by economic science test which in practice is not observed.

Acknowledgments

I would like to thank Kyle Bagwell, Bill Davey, Michael Finger, Gene Grossman, Jim Hartigan, Rob Howse, Axel Kallmayer, Edwini Kessie, and especially Andre Sapir for very helpful comments. Serra Ayral helped me with numerous WTO documents. Henrik Horn has generously reviewed the manuscript and provided me with detailed comments. He is of course only partially responsible for any remaining errors.

Notes

1 To be sure, "going regional" should not be understood as an option to integrate at geographically regional level. Whereas this was indeed originally the case (regional integration between geographic neighbors), recent preferential trade agreements (PTAs) have defeated the geographical dimension (see, e.g., the FTA between the United States and Israel).

2 Baghwati (1996) however points out that countries opting for regional preferences could always include an "opt in" clause, whereby any third country which conforms to the requirements of the incumbents should be allowed to participate in the PTA established.

3 For a survey of the economics literature in this respect, see Horn and Mavroidis (2000).

4 See, e.g., the assessment of Bhagwati and Paningariya (1996) and that of Winters (1996).
5 Indeed, if trade diversion was the only motive for going regional, it would seem that in a world where customs duties are quite low, there would be less incentive for governments to exercise this option. Irwin (1996) argues that duties calculated on dutiable imports stood at 19.34 percent in 1947 and 13.87 percent in 1948. However, were one to include duty free trade, the ration of duties collected to total imports amounted to 7.55 percent in 1947 and 5.71 percent in 1948. The average tariffs on dutiable imports (weighted with the 1939 trade values) in 1947 was 32.2 and 25.4 percent in 1948. As 60 percent of US imports was duty free in 1947, the average tariff for total imports were much lower. In addition, were tariff averages to be weighted with the 1947 trade weights (and prices and not with 1939 values) then the post-Geneva average tariff on dutiable imports drops from 25.4 percent to 15 percent. Including duty free imports, the pre-Annecy (GATT 1948) tariff average of all US imports was estimated by the US Tariff Commission to amount to 5.9 percent. This number fits almost unbelievably well with the ration of duties calculated to total imports (free and dutiable) of 5.97 percent in 1950.
6 This interpretation is in line with § 4 of the Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations where it is stated that "the representatives *agree* that the WTO Agreement shall be open for acceptance as a whole" (emphasis in the original).
7 See the Appellate Body report on *Brazil – Dessicated Coconuts* WT/DS22/AB/R at 12.
8 Messerlin seems to be sympathetic to this idea, see Messerlin (1995).
9 See WTO Doc. WT/L/127.
10 See WTO Doc. WT/REG/1 of August 14, 1996.
11 In fact, if CRTA continues the practice of Article XXIV Working Parties in this respect, the final reports will invariably reflect disagreements.
12 One should keep in mind that, whereas the Agreement Establishing The WTO states in unambiguous terms that WTO Agreements (primary law) are binding on the WTO Membership, it does not do the same with respect to decisions by WTO organs (secondary law). This does not mean that decisions by WTO organs are *per se* nonbinding (because they are secondary law). Sometimes the Agreement itself lays out the consequences in case of nonrespect of a decision by a WTO organ (like, e.g., the right to request countermeasures as a result of the nonrespect of a panel/Appellate Body finding). Besides, WTO institutional practice shows that WTO Members overwhelmingly abide by what they vote. The fact that decisions are taken by consensus (in the vast majority of cases) obviously operates as a "legitimizing" factor in the eyes of WTO Members.
13 The argument of course could be made that most notified PTAs are interim agreements since de facto most (if not all) of them do not proceed to trade liberalization from one day to the other. In such a scenario, the CRTA would have quite extensive powers that have not been used so far precisely because it is almost impossible to reach consensus on such issues and further because in the absence of consensus WTO Members are unlikely to proceed to voting in this context.
14 Schott (1989, 25) mentions only four unanimous decisions. In all four cases, Working Parties actually admitted that the notified PTA was GATT-consistent. The same was true in the only other case decided by consensus ever since: the CU between the Czech and the Slovak Republics.
15 For information on this issue, on all notified PTAs, see the table in the Analytical Index 858ff.
16 This is why goods originating in a non-FTA State are in practice accompanied by a certificate of origin when moving from one FTA member to another.

17 See Serra et al. (1997) on this point.

18 Incidentally, this seems to have been the intention of the drafters. Quoting from the preparatory materials: "The phrase 'on the whole' . . . did not mean that an average tariff should be laid down in respect of each individual product, but merely that the whole level of tariffs of a customs union should not be higher than the average overall level of the former constituent territories," see EPCT/C.II/38 at 9 reproduced in the GATT Analytical Index: Guide to GATT Law and Practice, Updated 6th Edition (1995) at 803 (hereinafter the Analytical Index); "The Sub-Committee recommended that the words 'average level of duties' be replaced by 'general incidence of duties' in paragraph 2(a) of the new Article. It was the intention of the Sub-Committee that this phrase should not require a mathematical average of customs duties but should permit greater flexibility so that the volume of trade may be taken into account," Havana Reports reproduced in Analytical Index at 803.

19 See BISD 30S/168 at 184.

20 See BISD 35S/293 at 295–6.

21 *Idem* at 311.

22 See, e.g., the discussions in the Working Party examining the compatibility of the EEC with Art. XXIV, GATT Doc. SR.18/4 at 46–54 and also in C/M/8, SR.19/6-7; see the Working Party report on "Accession of Greece to the European Communities," op. cit., at 175; see also the 1991 Working Party report on "Free Trade Agreement Between Canada and the United States," BISD 38S/47 at 66.

23 Art. XXIV.6 GATT refers to bound and not to applied duties.

24 With the notable exception of the Tuna–Dolphin panel reports, see Chapter 4 on TRADE AND ENVIRONMENT in this book.

25 The Analytical Index, e.g., reflects reports discussing this distinction at 824–5.

26 As Andre Sapir in rather colloquial terms often remarks in conferences on Regionalism, there must be some reason why the founding fathers of the GATT placed the nondiscrimination principle in Art. I GATT and the possibility to form PTAs 23 articles later.

27 See WTO Doc. WT/DS34/R of May 31, 1999 (hereinafter referred to as *Turkey – Textiles*).

28 See WTO Doc. WT/DS34/AB/R of October 22, 1999 at paras 58 and 59. The fact that the Appellate Body made this point in an *obiter dictum* does not reduce the value of the statement with respect to precedent-setting considerations: it is true that a statement that is not reflected in the *ratio decidendi* of a report is of reduced legal value; at the same time however, the Appellate Body does not have to add *obiter dicta*. When it does, it is usually because it wants to signal its future reaction to a particular set of cases. A reasonable interpretation of this signal by the Appellate Body would be that, in the future, WTO Members which want to benefit from an Art. XXIV GATT defense will have to establish first that their PTA is in full conformity with the WTO rules. In Section 5 *infra* we deal with the question whether this is a good policy or not.

29 See B.I.S.D. 20S/171 at para.16.

30 See the Working Party report on EEC, B.I.S.D. 6S/100 at para. 34.

31 See B.I.S.D. 6S/70 at para. 30.

32 See B.I.S.D. 96/83 at para. 48.

33 See the Working Party report on "EEC – Agreements with Finland," B.I.S.D. 29S/79 at para. 12.

34 See the Working Party report on "Free Trade Area between Canada and the US," B.I.S.D. 38S/73 at para. 83.

35 See WTO Doc. WT/REG/W/18 (November 17, 1997).

36 See WTO Doc. WT/REG/W/22/Add. 1 of April 24, 1998.

37 See WTO Docs. WT/REG/W/17 (October 31, 1997); WT/REG/W/17/Add. 1 (November 5, 1997); WT/REG/W/17/Corr. 1 (December 15, 1997); WT/REG/W/17/Rev. 1 (February 15, 1998).
38 Incumbents (existing PTAs) legally will not be submitted to the same review, assuming of course that a new standard is found. They still are subject to the law in place where they formed a PTA. Political considerations rather than legal grounds argue that nonprecision of the term will avoid comparisons where harmful.
39 This point is particularly relevant for the discussion on rules of origin. As Serra et al. (1997) point out, rules of origin have proved to be one of the most forceful barriers to market access within an FTA. For the time being though, there is no WTO Agreement on Rules of Origin, the negotiators having repeatedly failed to settle the issue. In the absence of a WTO Agreement any WTRO Member would be obliged to apply its own laws in a nondiscriminatory manner.
40 On this issue, see Mavroidis (1997).
41 See the excellent analysis of Charnovitz (1998) in this respect.
42 See B.I.S.D. 18S/133 at 135–7.
43 See B.I.S.D. 6S/70 at 97.
44 See the Analytical Index at 820ff.
45 See WTO Doc. WT/REG/W/17/Rev. 1 at 4.
46 See *Argentina – Safeguard Measures On Imports Of Footwear*, WTO Doc. WT/DS121/AB/R of December 14, 1999.
47 See *United States – Definitive Safeguard Measures On Imports Of Wheat Gluten From The EC*, WTO Doc. WT/DS166 of July 31, 2000.
48 The Appellate Body had the opportunity to pronounce on the legal status of GATT adopted panel reports: according to its case law, such reports form part of the so-called GATT *acquis* and provide useful guidance to subsequent panels dealing with the same issue. It had never had the opportunity so far to pronounce on the legal status of WTO adopted reports. From a political perspective GATT adopted reports seem to carry enhanced legitimacy because of the *imprimatur* given to them through positive consensus of the losing party (which is not the case in the WTO era with the passage to negative consensus). From a legal perspective however, it would seem untenable to support the thesis that WTO adopted reports do not enjoy at least the same status as their GATT counterparts. So although there is nothing like *stare decisis* in WTO law (indeed, there is nothing like *stare decisis* in public international law) adopted WTO reports exercise, as modern practice shows, considerable influence on subsequent reports dealing with the same issue. Following prior decisions when appropriate, and motivating departures when needed is probably the surest way for WTO adjudicating bodies to acquire legitimacy.
49 On this issue see Dorf (1998).
50 I accept here that antidumping duties should be viewed as a "restrictive regulation of commerce," i.e., that they form an integral part of the *internal requirement*. They should automatically, by consequence, be excluded from the ambit of the *external requirement*. As a result, WTO Members that enter into PTAs can legitimately impose antidumping duties against nonparticipants to their PTA. This is a legally safe conclusion with strong underpinnings in GATT/WTO practice.
51 Bhagwati's "open regionalism" aspiration could gain full effect in this context.
52 At the time of writing this chapter, the only Working Party which had made substantial progress in examining the consistency of a notified PTA with Art. V GATS was the Working Party on NAFTA, see WTO Doc. S/C/N/4.
53 Compare the discussions in the Art. XXIV Working Party reports where it was often stated that no sector should be excluded.

54 For example, the Working Party examining the 1973 enlargement of the then EEC to the United Kingdom, Republic of Ireland, and Denmark did not issue a final report, see GATT Doc. L/5453, adopted on March 9, 1983, BISD 30S/168 at 174.
55 L/5776, unadopted, dated February 7, 1985, para. 4.10.
56 See, e.g., the FTA Agreement between Israel and the United States, in GATT Doc. B.I.S.D. 34S/58ff. at 63, title III.
57 See, e.g., the Accession of Portugal and Spain to the European Communities, GATT Doc. B.I.S.D. 35S/293ff.
58 Excerpt re-produced in Roessler (1993, 322).
59 Concurring Roessler (1993, 321–2).
60 The first, after a request by Canada in 1974 in connection with the accession to the European Community of Denmark, Ireland, and the United Kingdom (GATT Doc. C/W/250) was not activated because the parties to the dispute reached an agreement (GATT Doc. C/W/259). The second led to an unadopted panel report in *EC – Tariff Treatment on Imports of Citrus Products from Certain Countries in the Mediterranean Region*, GATT Doc. L/5776 (hereinafter the *EC – Citrus* panel report). The third report is on *EEC – Import Regime of Bananas*, GATT Doc. DS38/R of February 11, 1994 (hereinafter the *EEC – Bananas* panel report) which also remains unadopted.
61 See GATT Doc. L/5776, dated February 7, 1985 at para. 4.6 and at para. 4.10.
62 See on this issue the conclusions of the WTO Appellate Body in *Japan – Taxes on Alcoholic Beverages.*
63 See GATT Analytical Index at 781.
64 See the panel report on *Turkey – Restrictions on Imports of Textile and Clothing Products*, WTO Doc. WT/DS34/R of May 31, 1999.
65 See WTO Doc. WT/DS34/AB/R of October 22, 1999 at §§ 58–9.
66 See WTO Doc. WT/DS90/R of April 6, 1999.
67 For a relatively recent expression of a typical disagreement, see paras 31–8 of the Working Party report on the Free Trade Agreement between EFTA and Turkey, adopted on December 17, 1993 (L/7336).
68 See WTO Doc. WT/DS90/AB/R of August 23, 1999 at paras 98ff.

References

Baldwin, Robert E. 1997: The causes of regionalism. *World Economy*, 20, 865–88.

Bhagwati, Jagdish and Arvind Paningariya 1996: Preferential trading areas and multilateralism: strangers, friends or foes? In Bhagwati, Jagdish and Arvind Paningariya (eds.), *Free Trade Areas or Free Trade? The Economics of Preferential Trading Agreements*. Washington DC: AEI Press.

Charnovitz, Steve 1998: The moral exception in trade policy. *Virginia Journal of International Law*, 38, 689–732.

Dorf, Michael 1998: The supreme court 1997 term, foreword: the limits of socratic deliberation. *Harvard Law Review,* 4–83.

Horn, Henrik and Petros C. Mavroidis 2001: Economic and legal aspects of the most favoured nation clause. *European Journal of Political Economy* (forthcoming).

Hudec, Robert E. 1972: GATT or GABB? The future design of the general agreement on tariffs and trade. *The Yale Law Journal*, 80, 1299–386.

Hudec, Robert E. 1993: GATT's influence on regional agreements: a comment. In Jaime de Melo and Arvind Paningariya (eds.), *New Dimensions in Regional Integration*. Cambridge, Mass: CEPR, Cambridge University Press, 151–5

Irwin, Douglas 1996: Changes in US tariffs: prices or policies. NBER Working Paper 5665.

Mattli, Walter 1999: *The Logic Of Regional Integration: Europe and Beyond.* Cambridge, UK: Cambridge University Press

Mavroidis, Petros C. 1997: The treatment of dumping, subsidies and anti-competitive practices in regional trade agreements, In Paul Demaret, Jean-Francois Bellis, and Gonzalo Garcia Jimenez (eds.), *Regionalism and Multilateralism After The Uruguay Round.* Brussels: European Inter-university Press, 389–96.

Mavroidis, Petros C. 2000: Remedies in the WTO legal system: between a rock and a hard place.

Mavroidis, Petros C. 2002: Judicial supremacy, judicial restraint and the issue of consistency of preferential trade agreements with the WTO: the apple in the picture. In Dan Kennedy and James Southwick (eds.), *The Political Economy of the International Trade Law, Essays in Honor of Robert E. Hudec.* Cambridge, UK: Cambridge University Press, 583–601.

Messerlin, Patrick 1995: *La nouvelle OMC*. Paris: Dunod.

Roessler, Frieder 1993: The relationship between regional integration agreements and the multilateral trade order. In Kym Anderson and Richard Blackhurst (eds.), *Regional Integration and the Global Trading System.* Exeter, UK: Harvester Wheatsheaf, 311–25.

Roessler, Frieder 2000: Are the judicial organs overburdened? Paper presented in the *Conference in the Honour of Raymond Vernon.* Harvard, Kennedy School, 2–3 June 2000.

Sapir, Andre 1998: The political economy of EC regionalism. *European Economic Review*, 42, 717–32

Schott, Jeffrey 1989: More free trade areas? In Jeffrey Schott (ed.), *Free Trade Areas and US Trade Policy.* Washington D.C.: Institute of International Economics, 1–58

Serra Jaime, Guillermo Aguilar, Jose Cordoba, Gene Grossman, Carla Hills, John Jackson, Julius Katz, Pedro Noyola and Michael Wilson 1997: *Reflections on Regionalism; Report of the Study Group on International Trade. Carnegie Endowment for International Peace*. Washington D.C.: The Brookings Institution Press.

Trebilcock, Mike and Robert E. Howse 1999: *The Regulation of International Trade*, 2nd edn. New York and London: Routledge.

Winters, Alan L. 1996: Regionalism versus multilateralism. Discussion Paper No. 1525. London: CEPR.

8

Trade and Informal Institutions

James E. Anderson

Chapter Outline

This chapter surveys recent work on informal institutions and trade. Trade barriers other than tariffs, quotas, and transport costs are apparently very large even between developed countries such as the United States and Canada. Nations exhibit wide variation in their use of informal institutions in trade, suggesting complex relations between formal and informal institutions and the volume of trade. New institutional economic research on informal institutions and trade attempts to explain these phenomena. The institutions presumptively lower trade costs but impose costs of their own. Informal institutions presumptively substitute for but may complement formal institutions. Better institutions are not always in every trader's interest.

Markets are shaped by institutions, a fact so obvious that only an economist could ignore it. Recent progress in the new institutional economics has opened the door to building tractable models of markets in which institutions play a role. International borders are necessarily bridged with institutions, so progress with models of institutions promises to greatly enhance understanding of barriers to and benefits from trade. This chapter reviews the progress in the economic research program on trade and informal institutions. By institutions we mean rules and procedures for enforcing the rules; informal refers to non-State rules and/or enforcement.

Underneath formal trade policy such as tariffs, there are very substantial barriers to trade posed by corrupt officials, unreliable courts, and outright predation. Recent empirical work demonstrates that even modern economies such as the United States and Canada pose very substantial barriers to trade with each other (McCallum, 1995; Anderson and van Wincoop, 2003). Anderson and Marcouiller (2002) show that 1996 trade flows to Latin America are constrained as much by poor Latin institutions as by

high Latin American formal trade barriers. Institutions differ across environments and this impacts the pattern of bilateral trade. For a wonderfully interesting and rich description of institutional ingenuity in medieval long-distance trade, see Braudel (1992). For a good description of international trading intermediaries, see Cho (1987).

Economists have in the past mostly ignored institutions[1] because of their research program of constructing simple mathematical models of important market forces.[2] The standard market model of the economics principles textbooks, the Marshallian scissors diagram with demand and supply schedules plotted in price–quantity space, is therefore institution-free. The model can as well represent the market for slaves in fourth century BC Athens as the market for 30-day Treasuries in present day New York City. When the analysis requires recognition that some costs inhibit trade, a common device is to posit exogenous transactions costs typified by "iceberg melting" costs which reduce the amount of goods shipped by a fixed proportion of the volume.

The new research program on institutions and trade opens the black box of transactions costs to model the costs mathematically as the result of opportunistic behavior, on the one hand, interacting with institutions to constrain such behavior, on the other hand. The rules and procedures that define the institution are further understood to be selected by the forward-looking rational agents in some anticipation of the effect of the rules on those actions. Thus, the institutions have a rationale in efficiency. While much productive work on institutions can be done ignoring the international border, the context of the border between sovereign States is an especially rich area for analysis because (i) the interests of sovereign States do not always agree and (ii) the border itself provides opportunities for selfish behavior, which imposes costs on trade. The analysis reviewed in this chapter on cross-border trade can in many cases carry over to domestic trade, but by shutting off consideration of domestic exchange institutions the analysis does unrealistically shut down feedback mechanisms between cross-border and domestic exchange institutions.

Many important aspects of trade and institutions are not treated here to provide a compact focus. Formal institutions of international trade policy, for example, international institutions such as the World Trade Organization and national institutions such as the US International Trade Commission operate to both facilitate and hinder trade, modulating the forces of national interests. For a good survey of recent theory on this subject, see Staiger (1995). This chapter abstracts almost entirely from rich descriptions of the nation-state and its interactions with other nation-states. The division between formal and informal is not everywhere a bright line, but it is quite useful. The formal institutions of national and international policy, and the informal institutions interact with each other, a rich subject that this review avoids for simplicity. For a good recent survey of one form of interaction, which is important for international trade and investment, that of corruption, see the contributions in Elliott (1997). Formal trade policy institutions and informal institutions also interact with the institutions of government whereby individuals' interests clash and are aggregated; the subject of political economy models. See Grossman and Helpman (2001a) for a full research report. The review here abstracts almost entirely from political economy. Eventually, research must combine all these important channels of institutional interaction, but abstraction and simplification are essential to understanding the mathematical economic models required in the new institutional economics.

How do institutional barriers to trade operate? Can they explain much of the apparently large barriers that characterize even modern economies? Why are poor institutions chosen by national entities that apparently could choose better ones? Weak states leave a vacuum into which parastatal institutions move. What is the nature of state (formal) and parastatal (informal) institutional interaction? Recent work provides some answers to these questions.

The chapter summarizes how recent new institutional economic research has modeled informal institutions, which facilitate or hinder trade. Models are internally consistent logical structures which can be set out with mathematical description and thus are objects of potential measurement. Economic models contain decision-making agents who behave rationally, maximizing some self-interested objective. The economic environment in which the agents interact contains elements whereby the institution (its rules and other structure) is explained as modifying self-interested behavior in a way that benefits at least some agents. To succeed in this, each model typically abstracts from realistic features of the real world, which become the focus of another model. The format of the survey thus classifies the research by its model environment, by the type of economic problem, which the institution can be understood to address. As a result, informal institutions such as Mafias make several appearances as a solution to several economic problems.

It is useful to extract common themes that run through what is at first sight a quite heterogeneous collection of research. One theme of this work is that causation runs in both directions: not only do good institutions cause trade, trade opportunities can cause good institutions. A second theme is that better institutions are not always unanimously preferred. Conflict of interest can arise through several channels: (i) terms of trade effects (price changes) can be expected to accompany changes in trade costs, and the losers from this may not gain sufficiently from the other aspects of institutional improvement; and (ii) many of the models of transactions costs explored below imply externalities of one sort or another, and institutional improvements affect the size of the externality. International trade conflicts of interest must presumptively be resolved with unanimity rather than some political resolution process, a key difference from intranational trade.[3] A third theme of the work is path dependence, meaning that quite different equilibria can eventuate from initial conditions that may not differ very much. (Physical examples of path dependence are familiar, such as nearly identical golf strokes producing quite different "equilibrium" positions of golf balls on the course.) Path dependence in institutional models is primarily due to the conflict of interest theme and how it plays out in the models. Fourth, institutions that lower the barriers to trade also introduce new distortions of their own. For example, monopoly intermediaries can reduce barriers but will take advantage of their monopoly power. The final theme is that the obvious idea that state and informal institutions are substitutes and thus rival can be wrong: formal and informal institutions can be complements in some interesting and revealing ways.

Section 1 surveys recent work on contract enforcement and trade. Anarchic interaction presents agents with a holdup problem – negotiations between traders at the point of exchange will ignore sunk costs (such as the cost of transporting the goods to the point of exchange). When sunk costs are significant or the trader's return from the bargaining is low, trade which would otherwise be mutually profitable is deterred. Thus, some efficient investments in trade are deterred unless a contract can bind *ex ante* commitments

that the traders make to each other. In the absence of perfectly enforced and complete contracts, informal institutions arise to mitigate the holdup problem. These institutions introduce costs and distortions of their own, which can reduce trade and impede the progress toward better institutions.

Section 2 surveys some recent work on predation and trade. Goods in transit are exposed to theft and extortion by thieves and corrupt officials, a problem which is more acute at international borders. Institutions of specialized private protection arise to reap some of the gains from trade in this setting. But the institutions exploit their power in self-interested fashion, which can reduce trade.

Section 3 sketches an approach to the institutions of risk diversification supporting trade. In a good contracting environment, legitimate business risks can be fully diversified through insurance. Where contracts are weak or the potential losses are illegal (bribe or extortion payments) then insurance is unavailable and other institutions arise to provide partial diversification. The rise of formal institutions damages risk diversification through informal institutions and can result in harm to some traders.

Section 4 reviews work on information asymmetry and institutions which mitigate it. These institutions benefit insiders and can harm outsiders.

1 Contract Enforcement

Trade involves significant commitments of costs prior to meeting partners to exchange goods for other goods or for money. These costs are sunk when the partners meet. Thus they are irrelevant to the bargaining which occurs between partners upon their meeting. A forward-looking trader who is unable to recover his sunk costs from his payoff expected from the bargaining will not enter into trade at all. In situations where the arbitrage margin (the gain from trade) is sufficient to pay the trade costs of both sides, a failure to trade is inefficient. Yet, such a failure can easily arise when one side is weak in bargaining or when sunk costs are nontrivial. In fact, it can be shown that efficient trade in this situation would be an accident, occurring with probability equal to zero. This is the holdup problem. Holdup is less important as good outside options strengthen the weaker party's hand in bargaining. For example, finding another partner with the same willingness to pay or sell provides a good option. But in rather common circumstances (costly search in thin markets, relationship-specific design components) the outside options are poor and holdup is important.

Holdup can be resolved or at least ameliorated by three different institutions: formal contract enforcement, informal enforcement of agreement, and ownership. Each of these is being studied in recent research with promising results. By formal enforcement we mean a rule-driven process operated by a third party with no direct interest in trade. This can take several forms, one being the familiar law court which can use State sovereign power to enforce its decisions. Even here, much of the meaning of the contract arises from customary understanding and thus has informal aspects. Another as-if formal alternative is a Mafia, which uses illegitimate force to back its decisions. Informal enforcement occurs through a social network of traders which uses ostracism to enforce its decisions. Ownership means a combination of traders who divide the surplus according to ownership shares and other compensation schemes. This form avoids all

difficulties and costs of enforcement but incurs monitoring and other costs of its own, the subject of the classic Coase theory of the firm.

1.1 Formal Enforcement

A contract with formal enforcement is a commitment to trade goods of a given quality at an agreed price which is designed for enforcement by an outside authority.[4] It can be in both trading parties' interests to agree *ex ante* to set up or submit to a contract enforcement authority which can force them to execute their contracts even though one party would be better off *ex post* by reneging. This is obviously so if contracts are necessary to permit any trade at all. But most potential markets pose the problem less starkly – there will always be some traders with low enough fixed costs to be available from each side of the market. How might these traders improve their interaction through institutions? A key issue is that *international* transactions are between sovereign jurisdictions, so enforcement must be in the self-interest of both sides.

Marin and Schnitzer (1995) offer a model in which enforcement problems across sovereign jurisdictions are resolved by *counterpurchase*, the dominant form of countertrade. A high-technology export in the form of a plant is built in the poor country by the rich country's exporting firms. When completed, the exporter is to be paid in foreign exchange. The opportunistic rich country firm has an incentive to cheat on quality and the opportunistic poor country importer has an incentive to renege on the obligation to pay. The incentives can be aligned properly with a tied counterpurchase deal in which the rich country firm agrees to buy some goods (not those produced by the new plant – this form of countertrade is called buyback) at a mutually profitable price. The second deal, being worthwhile to both parties, acts like a performance bond ensuring the promised behavior on the first deal. They show that predictions of this model fit a sample of Austrian countertrade deals quite well.

Reputation can discipline opportunism, resulting in adherence to due process. In contract enforcement, the "rules of the game" as they affect foreigners are then the key to analysis. Anderson and Young (2000) show that it may not always be in both sides' interest to agree to commit to contract enforcement, but provide fairly broad conditions under which the rule of law dominates anarchy. In the absence of contracts, agents randomly match (once only) and bargain. In general, there will be unequal numbers of traders from the buyers' and sellers' sides of the market, so trade is inefficient. Contracts are always preferred by the scarce side of the market, but the excess side faces a tradeoff. On the one hand, without contracts they can exert their power in bargaining. On the other hand, they may not match. The desirability of contracts depends on the tradeoff. With a sufficiently high elasticity of response of traders from the scarce side of the market, contracts are preferred. Anderson and Young also show that other institutions such as monopoly middlemen can resolve holdup – the durability of the monopoly means that it has a valuable reputation to protect by sticking to its commitments on pricing policies, even when short-run gains could be had by reneging. But monopoly middlemen bring the distortion of monopoly pricing, and they will resist institutional reform moving toward the rule of law.

Anderson and Young (2002) show that when enforcement is imperfect, as is realistic, then under fairly general circumstances the optimal degree of enforceability (parameterized as the proportion of contracts in default which are forced to be executed) is less than perfect for one side of the market. Essentially similar elements explain the tradeoff faced by the excess side of the market in choosing the optimal enforcement regime. The model explains why developing countries that could choose otherwise offer such poor contract enforcement to outsiders: it is in their traders' interests to do so. Weak formal domestic institutions are associated with high bargaining power because commercial success selects for toughness in bargaining. When active international traders have high bargaining power, the model shows they prefer low enforcement.

McLaren (1999) analyzes the choice between detailed costly contracts and "handshakes" in markets. In thick markets (with trade again thickening the market) the handshake form faces fewer problems with holdup so it allows saving on the costly explicit contract form. The model predicts that informal arrangements will dominate formal contracting as the number of firms grows (market thickens) and as trade costs fall.

Grossman and Helpman (2001b) analyze the location of outsourcing in a model of imperfect contracting. There is a holdup problem between an input supplier and the purchaser who has specific requirements for its design. The supplier must make a relationship-specific investment prior to purchase, and only a parametric proportion of this investment is verifiable in court (hence contractable). The purchaser is located in the high-wage North and can search for suppliers in either the North or the low-wage South. The relative cost of customizing the input plays a role in the choice. But, also, thicker markets lower search costs. And finally, better contracting reduces the transactions cost. This setup is embedded in a general equilibrium model of two countries and two goods, generating a rich but still tractable set of interactions.

Descriptions of gangs and Mafias acting as contract enforcers in trade are available in the sociological literature (Gambetta, 1993). No new institutional economic model of this is available.

In all the preceding models, the contract enforcement authority is a black box. Given the formal complexities of the trading environment, this is an appropriate simplification. But progress requires opening the box. (1) Self-interested agents provide the enforcement, so their decisions must maximize their income (possibly also status or related nonmonetary reward). The process by which the compensation of enforcers and the scope of their power is decided must be modeled, embedding the self-interested behavior of the enforcer. (2) The environment out of which the enforcement process emerges must be modeled. Formal law courts historically have incorporated many of the informal rules and customs of the informal social networks which preceded them. But the two modes are not distinct even at a point in time. Even in the presence of impartial and inexpensive formal law courts, much enforcement continues to occur through social networks. Legitimate enforcement authority must in at least some areas compete with Mafia or gang enforcement. How this plays out cannot be analyzed until models of the behavior of each are first developed.

A seminal model which peeks into the black box is Greif, Milgrom, and Weingast (1994). A ruler establishes a market where he protects the trade from predators and

offers contract enforcement. Self-interest motivates the ruler to expropriate the traders. But, when traders are organized in a guild, they can credibly threaten to refrain from trading again with a ruler who expropriates. The social network (see the next section) aspects of the guild prevent the individual defections which otherwise would undo the power of this sanction. This mechanism can ensure the good behavior of the enforcer. A promising start on imperfect private enforcement is offered by Dixit (2004). He models enforcement intermediaries who resolve, at a price, the holdup problems arising between populations of randomly matching buyers and sellers. His model is fundamental enough to distinguish between an intermediary who merely assembles and disseminates information on reneging traders and an intermediary who has some punishing power.

1.2 Informal Enforcement

Holdup is alternatively resolved by the institution of social networks. The distinction between the formal contract enforcement and enforcement through social networks is that in a social network the participants in the market themselves act to enforce agreements. Because the traders have strong immediate interests to exploit temporary advantages, and because they must forgo immediate profits in order to punish reneging behavior, the main theoretical problem is to show that the long-run gains from sticking to an enforcement agreement to punish miscreants are sufficient to offset any temporary gains. The formal repeated game structure suggests that networks would be fragile, but the impressive evidence of history suggests their durability. It helps if the traders are outsiders in the environments where they operate, because noneconomic gains from solidarity help keep members in line. See Greif (1993) for an influential description of the Maghribi traders of the twelfth century, interpreted in light of this model. For other very interesting descriptions of networks see Braudel (1992). Social networks reap gains from trade but they also can serve to impede trade as insiders exploit their monopoly power relative to un-networked outsiders. Kranton and Minehart (2001) is notable for offering a rich description of multiply connected networks in which agents *choose* how many links to build. The multiple connections of an agent have a strategic value in reducing holdup. See Rauch (2001) for a good survey of networks.

Bernstein (2001) describes in fascinating detail how social networks enforce agreements in the US cotton trade. Interestingly, the growth of foreign sources of supply has not resulted in the internationalization of the social network, and presents enforcement problems for the trade.

Casella (1996) offers a model in which international arbitration supplements the work of courts. Traders can settle disputes through voluntary private arbitration with advance commitments to abide by arbitration decrees (which ultimately have to be enforced by national courts). Traders benefit from such a design. It takes away activity from national courts (so national formal institutions may oppose international arbitration). The point of the analysis is more general than the motivating example of arbitration: private arrangements endogenously shift to exploit market

opportunities. Casella's model treats the payoff to court proceedings as a black box, so it is difficult to integrate with other models of enforcement such as Anderson and Young's.

1.3 Ownership

Holdup can also be resolved by ownership structure. Where relationship-specific investment is necessary, recent literature has begun to analyze the vertically integrated form as an alternative to market interaction subject to holdup. Vertical integration is costly (or it would always be chosen), so there is a tradeoff of low direct costs through markets versus high monitoring costs in the firm which can explain institutional choice.

The international dimension of this story is that commodity flows across international boundaries can switch from arms' length to integrated (multinational) trade, or the substitution can be more complex as the change in organizational form could reduce or increase trade. The empirical literature on outsourcing points out the large and growing proportion of trade in intermediate goods, while the literature on multinationals points out the growth in multinationals and trade between subsidiaries of multinationals. Multinationals are in turn inextricably connected with international investment, particularly Foreign Direct Investment (FDI). This review avoids the large literature on FDI; for a good review see Markusen (1995).

Market thickness can help with holdup because it improves the outside options of the original parties to an agreement to trade and lessens the relative desirability of vertical integration. McLaren (2000) argues that market thickness, which increased due to trade liberalization and the reduction of other trade costs, can explain the global vertical disintegration of the auto industry over the last 20 years. The basic model has n USF (upstream firms) and DSF (downstream firms) which can either deal at arms length or integrate. The main idea is the market thickness principle: USF is more likely to find an alternative buyer for the customized input to use as a threat point the more unintegrated firms there are among the remaining buyers. There are many potential reasons for the market thickness principle but in this model it arises from the strategic decisions of the suppliers to specialize their inputs to the needs of DSFs. The DSF gains more with a customized input, creating a larger surplus to share between USF and DSF. The less specialized is the input, the more likely a good alternative buyer can be found to serve as a threat point in bargaining with the target DSF. The key ingredient is randomness in quality of the input. Uncertainty about quality is what allows each supplier to hope to have a good outside option while also having *ex post* some suppliers with no good option.

Grossman and Helpman (2000) use an essentially similar framework but build on a specification of incomplete contracting parameterized as the proportion of investment which is verifiable (in a court) and hence contractible. Thus arms' length contracting can cover a portion of USF's investment in specialization. They analyze the "make or buy" decision. Grossman and Helpman (2001b) use their model to analyze the location of outsourcing given a decision to buy rather than make, leading to an analysis of local versus external buying and a theory of trade related to institutional quality.

2 Predation and Enforcement

Predation either as outright theft or as extortion by unofficial or official opportunists is far more prevalent than standard discussion of international trade suggests. Businessman surveys show another picture. Anderson and Marcouiller (2002) and Marcouiller (2001) show that predation substantially reduces trade.

A natural environment in which to examine predation on trade is one of anonymous interaction where predators seek matches with traders and traders avoid matches with predators. In this, the predator–prey interaction resembles the anonymity of the market itself. An important property of the anonymous interaction of predator and prey is "safety in numbers," the externality which arises as more prey reduce the probability of meeting with a predator. Safety in numbers persists even after accounting for the response of increased predator supply, so long as the predator supply is less than infinitely elastic. In contrast to this setup, a literature on predation and production has focused on the productive agents' choice between production and defense, while predators can choose between offensive capability and production. A key feature of this literature is the bilateral interaction of predator and prey, resembling an arms race. This makes sense in considering production, which is naturally associated with a known location.

Anderson and Marcouiller (2003) explore a formal model of trade and predation under anarchy, and find that anarchy presumptively implies autarky. Anarchy means that all traders can do is evade their predators, no coordination of defense is possible, and the safety in numbers externality is not internalized. Anarchy also means that no insurance markets exist for diversifying the risk of predation. Agents have the choice of entering autarkic production, specializing production and trading, or becoming predators. Intuition suggests that trade would *not* be eliminated in equilibrium, because a predator–prey balance seems the likely outcome. As trade shrinks, the reward to predation would shrink as well. In contrast, the formal model shows that for most parameter values, autarky prevails. A key factor is the absence of risk diversification: with perfect insurance, most parameter values support trade. The model is motivated by the seventeenth-century collapse of Spain's American trade. The Crown clung to the convoy system and would not enforce insurance contracts for independent sailings, so the conditions approximate the model setup.

Traders subject to predation have an incentive to organize institutions to improve their lot (and predators may similarly gain from organizing). The state is one such institution, but the state can itself be a predator. This is most obviously true when the trade is illegal (e.g., cocaine or opium) but official predation significantly affects legal trade in much of the world. So informal institutions arise to enforce property rights in trade.

Anderson and Bandiera (2003) develop a model of informal enforcement by a Mafia. As compared to self-enforcement, the Mafia may either foster or hinder trade depending on the balance of three forces. On the one hand, the Mafia can internalize the safety in numbers externality and it can offer a better protection technology, modeled as a positive probability that stolen goods can be recovered. On the other hand, the Mafia uses its monopoly power to overprice security. The Mafia has to cover a fixed cost to enter a market, so low-value markets will be stuck with self-enforcement.

The Mafia and the State interact in two distinct ways. In illegal trade, the State and the Mafia are directly opposed. The State seeks to reduce the volume of trade either by preying on it directly or by damaging the Mafia's capacity to protect it. The model shows that most obvious policies can backfire. For example, trying to reduce the Mafia's effectiveness can result in a rise in trade volume. The model may provide an explanation for the spectacular failure of the US war on drugs.

In legal trade, the State and the Mafia are competitors in the enforcement business. In the formal sector the State provides protection from predators and charges taxes. In the informal sector (the grey market) the Mafia does the same (taxes are called fees here, inessentially). A key insight of the model is that the enforcement strategies of the State and the Mafia are strategic complements. This means that State attempts to reduce the volume of informal trade will hurt the State's own revenues. Strategic complementarity may offer an explanation why Mafias persist even in strong states (Italy, Japan, ...). The basic structure explaining strategic complementarity is safety in numbers, in this case a general equilibrium version of safety in numbers due to predators distributing themselves between markets to equalize their returns. Thus a rise in numbers in the formal sector raises the per predator payoff there and draws some more predators from the informal sector, the net effect being that the success rate for traders rises in both markets.

When goods must be customized, even if contracts are perfect, predation imposes the risk of nondelivery or at best of an inferior generic good being purchased as a substitute. Important aspects of this loss appear to be difficult to insure, since the "market value" of a stolen input is less than its value to the target customer. This trade environment subject to predation has not yet been modeled to my knowledge, but it appears useful and likely to be tractable.

3 Risk Diversification

Risk associated with trade arises in two important forms. First, if trade volumes are chosen after the realization of random variables (price or productivity) then the risk is to income. Second, if planned trade must be chosen prior to the realization of the shock to trade, then there is in addition to income risk a risk to consumption (or more precisely to the consumption mix). Each type of risk is significant and each requires different modeling. Upon the basic production and trade model it is necessary to build models for trade in risky assets. Finally, these must be set into institutional models.

When trade volumes are chosen after randomness is resolved, the mix of production can be chosen to minimize the risk to income. Suppose that perfect asset markets are to be found domestically, so that stock markets value the shares in output of each industry by appropriately valuing the risky returns based on anticipations of external prices and internal productivity shocks. Then the market will select the optimal production bundle; the invisible hand theorem holds. Income risk is still present, but is appropriately reduced given the asset market structure. In the limit, with extreme aversion to income risk, the optimal production choice will imply coinciding with the autarky choice (Anderson and Riley, 1976). But income risk can be diversified by international asset markets. Helpman and Razin (1978) show that with perfect international

asset markets of a special sort, the production choice is separated from risk diversification considerations. While this is an important benchmark result, its prediction is wildly at variance with empirical regularities. For example, diversified asset holders should have personal income which is independent of shocks to GDP. Instead, these variables are highly correlated (Asdrubali, Sorensen, and Yosha, 1996). Asset markets are neither completely closed nor completely integrated. While government controls account for some of the transactions cost which limits integration, empirical work suggests that institutional factors play a strong role, as is indeed intuitive. The theoretical literature has failed to develop a good model of partially integrated capital markets, so the new institutional economics of international risk diversification awaits a founder.

The alternative form of trade risk arises when the planned trade is chosen prior to the resolution of randomness. For example, theft and holdup affect the *ex post* trade which is realized based on planned trade levels. An important aspect of this type of risk is that the risk to the consumption mix which it induces is not diversifiable.[5] Anderson and Marcouiller (2003) analyze the consumption risk that an income-risk-neutral[6] producer/trader faces when there is a chance that the exports sent out will not yield imports in return due to predation. In a full general equilibrium model where agents can enter either predation or specialized producer/trader activity, they show that anarchy presumptively implies autarky. The presumption is based on simulation of a simple special case model, but the implication seems likely to be robust. The model is suggestively applied to explain the collapse of seventeenth-century Spanish–American trade under the weight of a shift in technology favoring predation combined with the Crown's adherence to the convoy system and a refusal to enforce insurance on unregistered shipments.

Risk diversification can be partially achieved by networks. Kranton and Minehart (2001) show that networks where buyers and sellers are multiply connected diversify the risk of difficulties facing any one trade partner. McLaren and Newman (2001) build on this idea to show how globalization, by offering better markets, will harm networks. The market offers better risk diversification, but in reducing the risk diversification available through networks, it is possible that a utilitarian welfare measure will fall. They show that small countries must gain on balance while large countries can lose.

Institutions to diversify consumption mix risk include trading companies. With a large number of independent shipments facing a given success rate, an effectively certain number of shipments are successful. In effect, the planned mix can always be realized. This will increase trade and real income. But since diversification implies scale economies, associated with large size and market power, institutions like trading companies acting as intermediaries between producer/traders can exploit their market power, reducing trade and real income. Finally, the emergence of domestic wholesale markets can diversify the risk of end-users who no longer trade internationally at all. If insurance is available, the size of international traders who arbitrage between domestic wholesale markets can be small and the market becomes competitive. In light of this discussion, it is a striking observation that contemporary US trade in homogeneous products like agriculture is carried on by very large firms. A rich descriptive source for trading companies is Cho (1987). Formal modeling of trading companies seems likely to yield useful insights.

Even the income risk associated with predation on trade may be important. Theft in well-established markets might be thought to represent an easily diversified risk. Even though moral hazard interferes with the efficiency of normal commercial theft insurance, it does not appear to limit its availability. In thin trade markets, however, insurance may not be so readily obtained. But extortion is a more important form of predation in modern times and here insurance is unavailable, since the payment of bribes is illegal. With undiversifiable income risk, the risk aversion of the traders comes into play and inhibits further their trade.

More theory but also more empirical observation of risk-sharing arrangements are needed.

4 Information

Imperfect and especially asymmetric information impedes the operation of markets. Portes and Rey (1999) show that greater density of telephone connectivity between bilateral trade pairs significantly raises trade and FDI. Gould (1994) shows that immigrant linkages significantly raise trade. Rauch and coauthors (e.g., Rauch and Trindade, 1999) have done most to document the importance of linkages, focusing especially on the overseas Chinese. These empirical investigations use the gravity model to control for other factors and make no attempt at a model of just how information impedes trade. The immigrant linkages results might alternatively stand for the role of social networks as informal agreement enforcers. Despite these critiques and others which can be made with respect to their use of the gravity model (see Anderson and van Wincoop, 2003), there is little doubt the qualitative finding is robust: information is imperfect and mechanisms which improve its flow do significantly raise trade.

Making progress in understanding how institutions reduce trade frictions due to information requires a model of the informational environment where agents do not coordinate information; atop which must sit a stylized model of an institution which yields gains to the participants. As with other institutions, gains to the insiders may come at the expense of outsiders, and membership in the club may be restricted to the benefit of members. Some aspects of imperfect enforcement and the institutional models of it reviewed above reflect information imperfection in the background; here the focus is on institutions to deal with information alone. Carrying out this research presents researchers with some daunting complexities, so it is not far advanced.

Casella and Rauch (1998) develop a first-generation model of information asymmetry with the features laid out above. Foreign and home agents randomly match with counterparties when there is no information. Social networks facilitate perfectly informed matching, and this yields more surplus. The operation of networks can reduce the surplus of outsiders. This package of properties is quite useful since the distributional implications do look like the world we see. Future progress is not easy because the model exogenously switches between no and perfect information as we exogenously switch on or off the social network. All the useful questions about investing in better information, extending the size of the network by a bit and the interaction of the network with a formal provider of information cannot be addressed in the model. This should be on the agenda of future research.

5 Conclusion

Recent progress in mathematical modeling of informal institutions promises to deepen understanding of hitherto hidden barriers to trade. Empirical work not tied closely to developing theory has already demonstrated that institutional variation impedes a lot of trade, so the topic is important. Better theory will make for more precise and informative empirical work. With better understanding will come an increased ability to change trade barriers with acts of policy, and also an increased ability to distinguish those trade-impeding barriers which are harmful from those which are useful.

An important goal of this chapter is communication across the boundary dividing economics from legal studies. Economists can appear bizarre even to themselves sometimes in the severe abstraction of their models. This chapter may give some assurance economists are incorporating in sequential fashion a host of important complications to the environment which give rise to the informal institutions which shape markets. Reading this chapter may stimulate lawyers to attempt to interest economists in institutional elements missing from the research thus far.

Notes

1 A narrow definition of "economists" is used here for simple classification purposes. The older school of institutional economists which flourished in the interwar period in the United States eschewed formal models, but it certainly used economic reasoning, regarded itself as scientific and would be astonished at exclusion from the set of economists.
2 This chapter is not the place for a philosophical discussion of the mathematical economics strategy. Pragmatically, among the set of economists who run the major PhD programs and the most prestigious journals, the folk wisdom is that "it doesn't count until you write down a model" and "it takes a model to beat a model."
3 In the absence of strong nation-states the point still applies: outsiders are treated differently and no group will treat them better than is in the group's perceived self-interest.
4 The requirement of design means that the terms must have common meaning to the outside authority as well as to the contracting parties. Personality should not matter, either of the contracting parties or the authority.
5 Fair insurance will pay off the value of the stolen goods *ex post* in return for *ex ante* payment of a premium which is equal to the expected loss. This smoothing of income does not compensate for the disutility of the shift in the consumption mix. Insuring the shift in consumption mix requires a type of insurance which is unlikely deal with problems of verifiability and moral hazard.
6 Neutrality to income risk could be due to the availability of normal theft insurance.

References

Anderson, James E. and Oriana Bandiera 2003: Traders, cops and robbers. NBER Working Paper No. 9572.
Anderson, James E. and Douglas Marcouiller 2002: Insecurity and the pattern of trade: an empirical investigation. *Review of Economics and Statistics*, 84(2), 345–52.

Anderson, James E. and Douglas Marcouiller 2003: Anarchy and autarky: endogenous predation as a barrier to trade. Boston College; revision of NBER Working Paper No. 6223; *International Economic Review* (forthcoming).

Anderson, James E. and John G. Riley 1976: International trade with fluctuating prices. *International Economic Review*, 17, 76–97.

Anderson, James E. and Eric van Wincoop 2003: Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93, 170–92.

Anderson, James E. and Leslie Young 2000: Trade implies law: the power of the weak. NBER Working Paper No. 7702.

Anderson, James E. and Leslie Young 2002: Imperfect contract enforcement. NBER Working Paper No. 8847.

Asdrubali, Pierfrederico, Bent E. Sorensen, and Oved Yosha 1996: Channels of interstate risk-sharing: United States, 1963–90. *Quarterly Journal of Economics*, 111, 1081–110.

Bernstein, Lisa 2001: Private commercial law in the cotton industry: creating cooperation through rules, norms, and institutions. 99 Mich. L. Rev. 1724.

Braudel, Fernand 1992: *The Wheels of Commerce*. Berkeley: University of California Press.

Casella, Alessandra 1996: On market integration and the development of institutions. The case of international commercial arbitration. *European Economic Review*, 40, 155–86.

Casella, Alessandra and James Rauch 1998: Overcoming informational barriers to international resource allocation: prices and group ties. NBER Working Paper 6628.

Cho, Dong-Sung 1987: *The General Trading Company*. Lexington, Massachusetts: Lexington Books.

Dixit, Avinash 2004: *Lawlessness and Economics: Alternative Modes of Economic Governance*. Princeton: Princeton University Press.

Elliott, Kimberly Ann 1997: *Corruption and the Global Economy*. Washington: Institute for International Economics.

Gambetta, D. 1993: *The Sicilian Mafia: The Business of Private Protection*. Cambridge, Massachusetts: Harvard University Press.

Greif, Avner 1993: Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition. *American Economic Review*, 83, 528–48.

Greif, Avner, Paul Milgrom and Barry R. Weingast 1994: Co-ordination, commitment and enforcement: the case of the merchant guild. *Journal of Political Economy*, 102, 745–76.

Gould, David M. 1994: Immigrant links to the home country: empirical implications for U.S. bilateral trade flows. *Review of Economics and Statistics*, 76(2), 302–16.

Grossman, Gene and Elhanan Helpman 2000: Integration versus Outsourcing in Industry Equilibrium. Harvard University.

Grossman, Gene and Elhanan Helpman 2001a: *Special Interest Politics*. Cambridge: MIT Press.

Grossman, Gene and Elhanan Helpman 2001b: Outsourcing in a Global Economy. Harvard University.

Helpman, Elhanan and Assaf Razin 1978: *A Theory of International Trade under Uncertainty*. New York: Academic Press.

Kranton, Rachel E. and Deborah F. Minehart 2001: A theory of buyer–seller networks. *American Economic Review*, 91, 485–508.

McCallum, John 1995: National borders matter: Canada–U.S. regional trade patterns. *American Economic Review*, 85(3), 615–23.

McLaren, John 1999: Supplier relations and the market context: a theory of handshakes. *Journal of International Economics*, 48, 121–38.

McLaren, John 2000: "Globalization" and vertical structure. *American Economic Review*, 90, 1239–54.

McLaren, John and Andrew Newman 2001: Globalization and insecurity. University of Virginia.

Marcouiller, Douglas 2001: Hijacking, hold-up, and international trade. Boston College Working Paper No. 477.

Marin, Dalia and Monika Schnitzer 1995: Tying trade flows: a theory of countertrade with evidence. *American Economic Review*, 85, 1047–64.

Markusen, James R. 1995: The boundaries of the multinational firm and the theory of international trade. *Journal of Economic Perspectives*, 9, 169–90.

Portes, Richard and Helene Rey 1999: The determinants of cross-border equity flows. NBER Working Paper No. 7336.

Rauch, James 2001: Business and social networks in international trade. *Journal of Economic Literature*, 39, 1177–203.

Rauch, James E. and Vitor Trindade 1999: Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84, 116–30.

Staiger, Robert 1995: International rules and institutions for trade policy. In G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, volume III, chapter 29. Amsterdam: Elsevier.

9

The Economics of Preferential Trade Agreements

Pravin Krishna

Chapter Outline

This survey is intended as a brief and accessible primer on the economics of preferential trade agreements (PTAs). It covers a number of analytical developments in the literature – from discussions on the questions relating to the "static" welfare impact of PTAs to the more recent contributions relating to the "dynamic" expansion of existing PTAs and the impact of PTAs on the world trade system. Particular emphasis is placed on those aspects of the economic analysis of PTAs that inform questions relating to institutional design.

1 Introduction

A cornerstone of the Geneva-based multilateral organization, the General Agreement on Tariffs and Trade (GATT), and its more recent incarnation, the World Trade Organization (WTO), is the principle of nondiscrimination: member countries may not discriminate against goods entering their borders based upon the country of origin. However, in a nearly singular exception to its own central prescript, the WTO does permit its members to enter into preferential trade agreements (PTAs), provided these preferences are complete, thereby sanctioning the formation of Free Trade Areas (FTAs), whose members are obligated to eliminate internal import barriers, and Customs Unions (CUs), whose members additionally agree on a common external tariff against imports from nonmembers.[1] Such PTAs are now in vogue. Even as multilateral approaches to trade liberalization – through negotiations organized by the GATT/WTO – have made substantial progress in reducing international barriers to trade, various countries have negotiated separate preferential trade treaties with each other in

the form of GATT/WTO-sanctioned PTAs. Among the more prominent PTAs currently in existence are the North American Free Trade Agreement (NAFTA), the European Economic Community (EEC) and the MERCOSUR (the CU between the Argentine Republic, Brazil, Paraguay, and Uruguay).[2]

This chapter is intended as a brief and accessible primer on the economics of PTAs.[3] It places particular emphasis on those aspects of the economic analysis of PTAs that inform questions relating to institution design. Are FTAs to be equated with free trade? Is preferential trade liberalization guaranteed to be welfare improving? If not, can necessarily welfare improving PTAs be designed? Can such PTAs accommodate noneconomic objectives? Do PTAs undermine the workings of the multilateral system? It is these questions that this chapter focuses on.

The rest of this chapter proceeds as follows. Section 2 is concerned with the basic welfare analytics of PTAs. It develops and builds on the well-known arguments of Viner (1950) concerning the generally ambiguous welfare effects of preferential trade liberalization. Given this ambiguity, the literature has focused considerable attention on ways in which necessarily welfare improving PTAs can be designed. Section 3 discusses these issues, taking care to distinguish between FTAs and CUs. Section 4 discusses the arguments made in the theoretical literature on the question of the impact of PTAs on the multilateral system. Section 5 concludes.

2 Welfare Analysis: Trade Creation and Trade Diversion

Does preferential trade liberalization in favor of particular trading partners have the same welfare consequences as nondiscriminatory trade liberalization in favor of all imports? Does a simple proportion of the welfare benefits of nondiscriminatory free trade accrue with preferential liberalization? That a country entering an FTA (where it eliminates tariffs against select partners) is doing something distinct from free trade as such (where it eliminates tariffs against all imports regardless of country of origin) should be easy to see. What this implies for the liberalizing country is a little more difficult to understand. Even a good half century after the economic implications of FTAs were first articulated by Viner (1950), the differences between FTAs and free trade remain a nuance that most policy analysts (and occasionally distinguished economists) appear to miss. So, how is the economics of free trade agreements and PTAs more generally different from that of free trade? And what does this distinction imply for the conduct of economic policy?

A thorough answer to these questions would require the reader to take a deep plunge into the abstruse world of the second best (whose existence and complexities were indeed first discovered and developed by analysts working on the economics of PTAs). But the idea may be introduced in a rudimentary fashion using the following "textbook" representation of Viner's analysis: consider the case of two countries, A and B, and the rest of the world W. A is our "home" country. A produces a good and trades it for the exports of its trading partners B and W. Both B and W are assumed to export the same good and offer it to A at a fixed (but different) price. Initially, imports from B and W are subject to nondiscriminatory trade restrictions: tariffs against B and W

are equal. Imagine now that A eliminates its tariffs against B while maintaining its tariffs against W. This is preferential tariff reduction as opposed to free trade, since the latter would require that tariffs against W be removed as well. It is very tempting to think that this reduction of tariffs against B is a step in the direction of free trade and therefore that this ought to deliver to Country A a proportionate fraction of the benefits of complete free trade. But Viner (1950) showed that this need not (and generally would not) be the case. Indeed, while a complete move toward free trade would be welfare improving for Country A, Viner demonstrated that the tariff preference granted to B through the FTA *could* in fact worsen A's welfare.

Figures 9.1 and 9.2 illustrate preferential tariff reform as respectively welfare-enhancing and welfare-worsening. The y-axes denote price and the x-axes denote quantities. M_A denotes the import demand curve of Country A. E_B and E_W denote the price at which countries B and W are willing to supply A's demand; they represent the export supply curves of B and W respectively. In Figure 9.1, B is assumed to be a more efficient supplier of A's import than is W: E_B is drawn below E_W, and its export price P_B is less than W's export price P_W. Let "T" denote the nondiscriminatory per-unit tariff that is applied against B and W. This renders the tariff-inclusive price to importers in A as $P_B + T$ and $P_W + T$, respectively. With this nondiscriminatory tariff in place, imports initially equal M_0 and the good is entirely imported from B. Tariff revenues in this initial situation equal the areas (1 + 2). When tariffs against B are eliminated preferentially, imports rise to M_{PT}. Imports continue to come entirely from B (since the import price from B now, P_B, is lower than the tariff-inclusive price of imports from W, $P_W + T$). The tariff preferences granted to B simply increase the volume of imports. This increase in the volume of trade with the country whose exports were initially being purchased by A anyway (i.e., with the more efficient producer) when tariffs against it

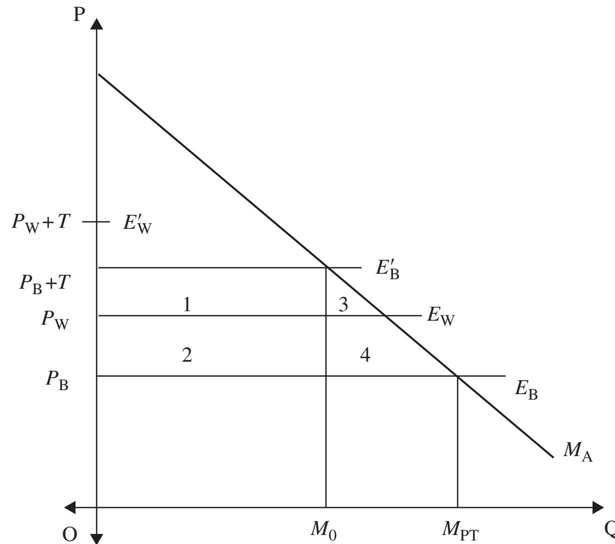

Figure 9.1 Trade creating tariff preferences: change in welfare = (3 + 4)

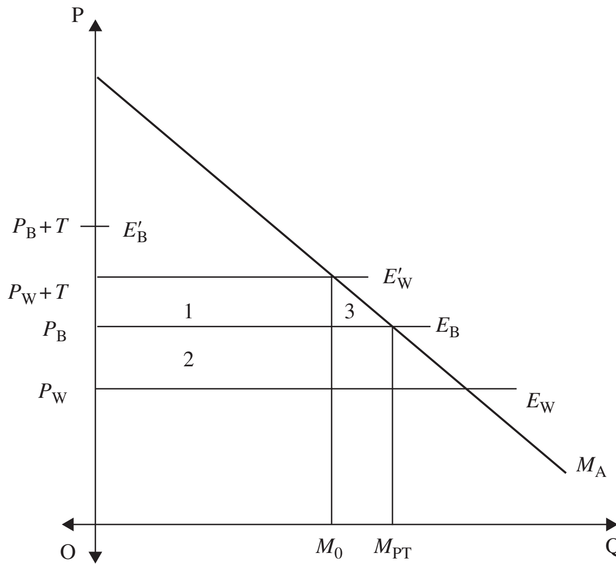

Figure 9.2 Trade diverting tariff preferences: change in welfare = (3−2)

are preferentially reduced is referred to as "trade creation." Trade creation here can be shown to be welfare improving. The increase in benefit to consumers (consumer surplus) in A following the reduction in consumption prices from $P_B + T$ to P_B equals the areas $(1+2+3+4)$. No tariff revenue is now earned and so the loss of tariff revenue equals areas $(1+2)$. The overall gain to A from this preferential tariff reduction equals areas $(1+2+3+4) - (1+2) =$ areas $(3+4)$, a positive number. The trade-creating tariff preference is thus welfare improving.

In demonstrating that the tariff preference we have considered is welfare improving for the home country, A, we have assumed that the partner which receives this tariff preference, B, is the more efficient supplier of the good. Figure 9.2 reverses this assumption, making W, the rest of the world, the more efficient supplier of the good. E_W is thus drawn below E_B. Initial imports are M_0. The tariff revenue collected is equal to the areas $(1 + 2)$. When tariffs are eliminated against B, the less efficient partner, the tariff-inclusive price of imports from W is higher than the tariff-exclusive price from B (this need not necessarily be the case, it is simply so as drawn). This implies that all trade is now "diverted" away from W to B. What is the welfare consequence of this trade diversion? The increase in consumer surplus is equal to the areas $(1 + 3)$ since consumers now pay a price equal to P_B for this good. The loss in tariff revenue is $(1 + 2)$. The overall gain to A equals the area $(3 - 2)$, which may or may not be positive. Thus, a trade-diverting tariff preference may lead to a welfare reduction.

The preceding examples illustrate well the central issues emphasized in the academic literature on the welfare consequences of preferential trade. Preferential trade liberalization toward the country from whom the good was imported in the initial

nondiscriminatory situation creates more trade and increases welfare; preferential liberalization that diverts trade instead may reduce welfare.

In more general contexts, with more than one commodity imported, this idea may be extended and modified: preferential tariff reduction on a basket of imports from a particular partner country will result in some trade creation and some trade diversion and, loosely speaking, the welfare of the liberalizing country is lowered if trade creation is dominated by trade diversion. Formally, this may be modeled as follows: consider again our trading world, which is composed of three countries, Country A, its prospective partner Country B, and a third Country W, representing the rest of the world. Each country only produces a single distinct good,[4] some of which it exports to pay for its consumption (imports) of the other two goods. Normalizing the border price of each good to be one, country A's budget constraint, representing the equality between expenditures and total revenues (revenues from sales of any goods produced plus import tax revenue), can be expressed as

$$E(1, 1 + t_{\mathrm{B}}, 1 + t_{\mathrm{W}}, U) = R(1, 1 + t_{\mathrm{B}}, 1 + t_{\mathrm{W}}, \bar{V}) + t_{\mathrm{B}}M_{\mathrm{B}} + t_{\mathrm{W}}M_{\mathrm{W}}, \qquad (1)$$

where E is the expenditure function associated with Country A, R is the revenue function (i.e., revenue derived from sales of any goods produced), U denotes country A's welfare, $\bar{V}$ denotes the (fixed) factor supplies used in production in A, and t_{B}, t_{W}, M_{B}, and M_{W} denote tariffs imposed against and imports from countries B and W, respectively. Equation (1) represents the budget constraint and the expenditure and revenue functions in the general form that is traditional in trade theory. We note first that our (Armington) assumption that each country only produces a single good and that factor supplies are fixed implies that the partials of the revenue function with respect to prices are zero. We are interested in analyzing the effect of a preferential reduction in tariffs imposed by Country A against Country B (holding tariffs against W fixed at $t_{\mathrm{W}} = \bar{t}_{\mathrm{W}}$. To get to this, we totally differentiate (1) and let E_i denote the partial derivatives of E with respect to the i'th domestic price to obtain

$$E_{\mathrm{B}}dt_{\mathrm{B}} + E_{\mathrm{U}}d_{\mathrm{U}} = t_{\mathrm{B}}dM_{\mathrm{B}} + M_{\mathrm{B}}dt_{\mathrm{B}} + \bar{t}_{\mathrm{W}}dM_{\mathrm{W}}. \qquad (2)$$

Since the partials of the expenditure function, E_i, denote consumption of the i'th good, it follows that $E_{\mathrm{B}} = M_{\mathrm{B}}$, and (2) therefore reduces to

$$E_{\mathrm{U}}dU = t_{\mathrm{B}}dM_{\mathrm{B}} + \bar{t}_{\mathrm{W}}dM_{\mathrm{W}}, \qquad (3)$$

where $E_{\mathrm{U}} > 0$, since it is simply the inverse of the marginal utility of income (which helps convert the real-income changes on the right-hand side into welfare units). This analysis extends the traditional framework of Viner (1950), in which the exportable supplied by B and W were perfect substitutes. Here too, however, we are able to see that concepts analogous to trade creation and trade diversion determine welfare outcomes. Expression (3) has the familiar intuitive interpretation: for welfare improvement to be guaranteed, both imports from the partner country and the rest of the world should increase. If, alternately, imports from the partner country increase (similar to trade creation), that is, $dM_{\mathrm{B}} > 0$, but the home country substitutes away from the rest of the world and imports from the rest of the world decrease (similar to trade diversion),

that is, $dM_C < 0$, welfare might drop instead.[5] Finally, welfare change following a reduction in tariffs against B is related to the tariff-weighted (and not the simple) sum of changes in trade volumes due to trade-creation and trade-diversion.

2.1 Revenue Transfer Effects

Thus far we have considered the case in which the home country is small relative to both the partner country and the rest of the world. In the Vinerian analysis illustrated in Figures 9.1 and 9.2, with the exportable from the partner and the rest of the world being perfect substitutes, when consumption is switched from the rest of the world to the partner country, the partner country is assumed to be able to satisfy all of the demand of the home country. What happens if B is so small that after receiving the tariff preference from A it is unable to satisfy all of A's demand for its importable. This implies that A continues to import some amount from the rest of the world W (which we assume for the moment is so large that it is able to handle all of the changes in A's demand without letting this affect its supply price) even after granting preferential access to B. Here, it can be shown that the home country loses unambiguously. The following example, provided by Panagariya (2000), illustrates. In Figure 9.3, the export supply curve of Country B is shown to be rising. The tariff inclusive supply curve faced by the home country is E_B^T. Total consumption of the importable initially is M_0 and imports from B are M_B^T. A tariff preference in favor of B simply shifts the effective export supply curve to E_B and the imports from B to M_B. Total imports stay at M_0. The domestic price of the importable in the home market in A is set by W (which continues to supply to A) and is

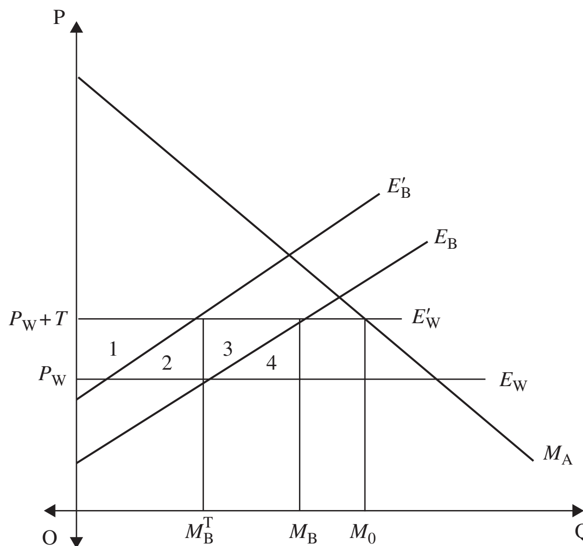

Figure 9.3 Change in welfare for home = −(1+2+3+4); change in welfare for partner = (1 + 2 + 3); and change in welfare for union = −(4)

the same as before (i.e., it stays at E'_{W}). The outcomes in this case are quite stark. Since consumption of the importable continues to be at M_0, there is no change in consumer surplus in the home country. There is however a direct tariff revenue loss since no tariff revenue is now earned on imports from the partner. The loss in tariff revenue (which is equal to the overall loss to A) equals the areas $(1 + 2 + 3 + 4)$. In what can effectively be seen as a tariff-revenue transfer to B, a gain of areas $(1 + 2 + 3)$ accrues to B in the form of an increase in producer surplus. Thus, preferential tariff liberalization leads to a loss in welfare for the liberalizing country, a (smaller) gain in welfare for its partner and a net loss of area (4) to the union as a whole.

In general, in the context of an *exchange* in tariff preferences negotiated under a PTA, we may expect that tariff-revenue losses to the home country in some sectors are made up for by gains in other sectors in which the home country gets preferential access to its partner's markets. Who gains more will depend upon the extent of tariff preferences exchanged and specific market circumstances (shapes of the supply and demand curves). The outcome is uncertain.[6]

We have so far studied the economic implications and income-redistribution effects that follow a preferential agreement between countries. The broadest message to take away from this discussion is that welfare gains with preferential trading are highly ambiguous. And we have seen that member countries may gain or lose with a PTA. Moreover, while we have not explicitly considered this so far, it should be easy to see that the rest of the world may lose as well if demand for their exports by the member countries in the PTA drops sufficiently so as to lower the price (i.e., worsen their terms of trade). It is worth noting, however, that the ambiguous welfare outcomes have resulted from trade preferences that have taken quite particular forms. In particular, we have analyzed tariff reductions in favor of the partner while external tariffs against all outside countries are maintained at their initial levels. An important question relating to institutional design can now be asked: by departing from this particular structure of trade preferences, is it possible to design PTAs where welfare improvement for all countries is guaranteed? It is to this topic we turn next.

3 Necessarily Welfare Improving Preferential Trade Areas

3.1 Customs Unions

As we have just argued, the literature on preferential trade areas rarely offers clear-cut answers with respect to the welfare impact of the formation of trading blocs between nations. A singular and important exception is the well-known result relating to customs unions, stated independently by Kemp (1964) and Vanek (1965) and proved subsequently by Ohyama (1972) and Kemp and Wan (1976): starting from a situation with an arbitrary structure of trade barriers, if two or more countries freeze their net external trade vector with the rest of the world through a set of common external tariffs and eliminate the barriers to internal trade (implying the formation of a CU), the welfare of the union as a whole necessarily improves (weakly) and that of the rest of the world does not fall. A Pareto-improving PTA may thus be achieved. The Kemp–Wan

construction offers a way to sidestep the complexities and ambiguities inherent in the analysis of PTAs. It has therefore played an important role in shaping the way that economists think about issues relating to the design and implementation of PTAs.

The logic behind the Kemp–Wan theorem is as follows: by fixing the combined, net extra-union trade vector of member countries at its pre-union level, we can guarantee nonmembers (i.e., the rest of the world) their original level of welfare. Moreover, taking the extra-union trade vector as an endowment, the joint welfare of the union is maximized by allowing free trade of goods internally (thus, in economics parlance, equating the marginal rate of substitution and marginal rate of transformation for each pair of commodities to each other and across all agents in the union). It should be easy to see that the PTA thus constructed has a common internal price vector implying further a common external tariff; it is therefore a necessarily welfare-improving CU.

A simple proof of the Kemp–Wan theorem demonstrating the possibility of Pareto-improving CUs runs as follows.[7] We continue to consider our trading world with the home and partner countries, A and B, and the rest of the world, W. Using lower and upper case letters to denote the home and the partner country respectively, we let $e(\cdot)$ and $E(\cdot)$ denote the standard expenditure functions, and $r(\cdot)$ and $R(\cdot)$ the standard revenue functions, in the home and the partner country. Consumer price vectors are denoted by p and P and welfare levels by u and U. An arbitrary structure of trade barriers may be assumed initially. The Kemp–Wan CU assumes that the total trade vector between the member countries and the rest of the world, W, is frozen – thus the terms of trade with the rest of the world are the same in the CU equilibrium as initially. We are interested in showing that the joint utility of the home and the partner country is (weakly) greater in the CU equilibrium we construct and that the welfare of the rest of the world is the same as it is initially.

The income–expenditure inequality for the union as a whole in the post-CU equilibrium is

$$e(p^{\mathrm{f}}, u^{\mathrm{f}}) + E(P^{\mathrm{f}}, U^{\mathrm{f}}) = r(p^{\mathrm{f}}) + R(P^{\mathrm{f}}) + (p^{\mathrm{f}} - P_{\mathrm{W}}^{\mathrm{o}})m^{\mathrm{o}} + (P^{\mathrm{f}} - P_{\mathrm{W}}^{\mathrm{o}})M^{\mathrm{o}}, \tag{4}$$

where the superscripts o and f denote the original equilibrium and the final (i.e., post-CU) equilibria respectively, $p^{\mathrm{f}} = P^{\mathrm{f}}$ is the common internal price vector with the CU, m^{o} and M^{o} are vectors of quantities imported by the home and the partner country, respectively, from the rest of the world (i.e., they do not include the initial imports from each other) initially and (by construction) in the post-CU equilibrium. We note that m^{o} is defined to include any goods from W that may have entered home through the partner in the pre-CU equilibrium. That is to say, in fixing the external import vector of a member, we include in it any goods that may have been imported or exported indirectly through the partner. A similar statement applies to M^{o}. Vector $P_{\mathrm{W}}^{\mathrm{o}}$ is the world price vector in the post-FTA equilibrium, which coincides with the pre-CU equilibrium since we freeze the joint external trade vector of the member countries with the rest of the world at their pre-CU levels. By the definition of the expenditure function, we have

$$e(p^{\mathrm{f}}, u^{\mathrm{o}}) + E(P^{\mathrm{f}}, U^{\mathrm{o}}) \leq p^{\mathrm{f}} d^{\mathrm{o}} + P^{\mathrm{f}} D^{\mathrm{o}}, \tag{5}$$

where d and D denote domestic consumption. For the home country, domestic demand (d^o) is satisfied by domestic production (x^o), imports from the rest of the world (m^o) and imports from the partner country (n^o). The same is true for the partner country. Using the fact that $n^o = -N^o$, and that $P_W^o(m^o + M^o) = 0$ given trade balance between the member countries and the rest of the world, we have

$$e(p^f, u^o) + E(P^f, U^o) \leq p^f x^o + P^f X^o + (P^f - P_W^o)(m^o + M^o). \tag{6}$$

For the joint welfare of the member countries to be greater in the CU equilibrium, it is sufficient that $e(p^f, u^f) + E(P^f, U^f) \geq e(p^f, u^o) + E(P^f, U^o)$. Note that (4) and (6) imply that for this to be true, we need that

$$r(p^f) + R(P^f) \geq p^f x^o + P^f X^o, \tag{7}$$

which is trivially satisfied given the definition of the revenue function. Thus the joint welfare of the home and partner country in the CU equilibrium is greater than (or equal to) their initial level of welfare. Note that the common external tariff of the CU is the difference between the internal price vector $P^f = p^f$ and the external price vector P_W^o. This completes the proof.[8]

This demonstration of a necessarily welfare improving PTA has (at least) two practical limitations. First, it is limited to CUs and does not consider the more common type of preferential agreement – FTAs. The second limitation relates to the common external tariff. While the theory is quite clear on what the nature of this external tariff ought to be (i.e., that it should eliminate trade diversion), there are a number of interesting questions regarding the suitable design or choice of the common external tariff that arise in practice – resulting, for instance, from the fact that the external tariff will have to be determined in advance of the CU agreement rather than as an equilibrium outcome (as the difference between domestic and world prices), as in the Kemp–Wan construction. We return to the discussion of external tariffs after first considering an analogous proof of necessarily welfare improving FTAs only recently provided by Panagariya and Krishna (2002).

3.2 Free Trade Areas

It should be straightforward to see that a demonstration regarding welfare-improving FTAs is substantially more complex than that for CUs: in the case of an FTA, member-specific tariff vectors imply that the domestic-price vectors differ across member countries and the FTA generally fails to equalize marginal rates of substitution across its members. This implies (in turn) that it is not possible to extend the *original* Kemp–Wan methodology to directly cover FTAs. Panagariya and Krishna (2002) have, however, recently provided a corresponding construction of necessarily welfare improving FTAs. The Panagariya–Krishna FTA, in complete analogy with the Kemp–Wan CU, freezes the external trade vector of the area, with the essential difference that the vector is frozen country by country. That is to say, the trade vector of *each* member country with the rest of the world is frozen at the pre-FTA level. Since in FTAs different member

countries impose different external tariffs, it is necessary to specify a set of rules-of-origin to prevent a subversion of FTA tariffs by import through the lower-tariff member country and direct transhipment of goods to the higher-tariff country (which, if allowed, would bring the FTA arbitrarily close to a CU). Interestingly Panagariya and Krishna find that the rules-of-origin necessary to support their welfare improving FTA are quite simple: all goods for which any value is added within the FTA are to be traded freely. The proportion of domestic value added in final goods does not enter as a criterion in the rules-of-origin.[9]

As we have already noted, the framework we have used above to prove the Kemp–Wan proposition was developed by Panagariya and Krishna (2002) and facilitates a relatively easy extension of the case FTAs. Before we get to this, we should point out again that the essential difference between CUs and FTAs relates to the nonequalization of prices within the FTA. To begin with, it is worth noting that this nonequalization of prices inside the FTA applies only to consumer prices. Producer prices are equalized since producers of goods within the FTA are free to sell goods in either country. Using the subscript c to denote consumers, we let p_c^f and P_c^f denote consumer prices in the home and the partner country and continue to let p^f and P^f denote producer prices in the home and the partner country. Goods can now be divided into three sets, to be denoted α, β, and γ. In set α, we place goods for which the consumer price in the home country exceeds that in the partner ($p_c^f > P_c^f$). For these goods, all within-FTA output is sold in the home country so that the FTA-wide producer price coincides with the consumer price at home (so that $p^f = P^f = p_c^f$). In set β, we include goods for which the consumer price in the partner exceeds that in home ($p_c^f < P_c^f$). In this case, the FTA-wide producer price coincides with the consumer price in the partner (so that $p^f = P^f = P_c^f$). Finally, in set γ, we have goods for which consumer and producer prices coincide FTA-wide (i.e., $p^f = P^f = P_c^f = P_c^f$).

As we have noted above, the Panagariya–Krishna FTA freezes the external trade vector of the FTA members country by country. Free internal trade is assumed. Focusing on a world with just final goods and following exactly the same logic as that used to arrive at equations (4)–(7), it is easy to show that the joint welfare of the FTA members is higher in the FTA equilibrium if

$$p^f x^o + P^f X^o \geq p_c^f x^o + P_c^f X^o + (p_c^f - P_c^f) n^o. \tag{8}$$

From the classification of goods into α, β, and γ goods and the discussion above, we know that $p^f = P^f = p_c^f$ for α goods, $p^f = P^f = P_c^f$ for all β goods and $p^f = P^f = P_c^f = p_c^f$ for all γ goods. This in turn implies that for (8) to be satisfied, we need that

$$[(p_c^f - P_c^f)(X^o + N^o)]_\alpha + [(P_c^f - p_c^f)(x^o + n^o)]_\beta \geq 0. \tag{9}$$

From our discussion above regarding prices and the fact that domestic output in each member country is at least as large as its exports to its partner, it should be clear that all the terms in the parentheses above are greater than (or equal to) zero. The member countries in the FTA have (weakly) greater welfare than before. The rest of the world is no worse off than before. The FTA is thus necessarily welfare improving, overall. This completes the proof.

It is worth noting that the member countries in the FTA just constructed each have their own individual tariff vectors with the tariff vector defined as the difference between the world price P_W^o and the consumer price vector within each country. The theory therefore provides us with specific guidance as to what form the external tariff vector in a (welfare-improving) FTA or a CU ought to take (i.e., that it should deliver the same level of trade with the rest of the world as initially – thereby eliminating trade diversion). However, there remain interesting practical problems with regard to the choice of this external tariff vector and in the elimination of internal barriers within the union. It is to these issues that we turn next.

3.3 Trade Barriers

The preceding discussion of necessarily welfare improving CUs and FTAs provided a precise description of the tariff vectors that ought to be implemented in these agreements. Specifically, internal barriers are to be completely eliminated and the external tariff vector in both cases (i.e., the CU or the FTA) should eliminate trade diversion – member countries should continue to import the same amounts from the rest of the world as they did initially.[10] Can these tariffs be implemented in practice? And where do existing GATT/WTO provisions stand in relation to the theoretical specification?[11]

Article XXIV of the GATT, which permits the formation of PTAs, also originally stipulated *broadly* that internal preferences needed to be complete (i.e., that internal barriers between the members were to be completely eliminated) and that external trade barriers were not to be more restrictive than initially. As we will discuss below, a number of questions arose in connection with GATT regulations regarding both internal and external tariffs – some having to do with their economic merit, others to do with implementation and possible abuse given the ambiguous and imprecise wording adopted in the original text of the GATT. As we will discuss further, while the more recent "Understanding on the Interpretation" of Article XXIV issued by the GATT in 1994 clarified some of these issues, other questions still remain.

On internal barriers to trade, two questions arise. The first relates to coverage – do GATT regulations require a removal of *all* internal barriers? The second relates to timing – how much time do countries have to comply with the rules? On the former issue, it should be clarified that while the putative intent of the GATT was to require that internal barriers be eliminated completely, the actual text of the GATT only required that restrictions be eliminated on "substantially all trade." The ambiguous phrasing through the use of the qualifier "substantially" opened up a number of possibilities for abuse. Whether "substantial" should have been taken to imply a full 100 percent or something smaller was not clear and has not yet been clarified. In this context, it is worth noting that for a given level of external tariffs, member country welfare is not necessarily maximized with zero internal barriers.[12] From a purely economic standpoint, given the level of external tariffs, welfare may well be maximized by maintaining some *particular* level of internal restrictions. It may therefore be *potentially* argued that the ambiguous phrasing permitting nonelimination of internal barriers allowed member countries to aim at welfare maximizing outcomes. This is however quite unlikely. Any retention of internal barriers within PTAs is probably better explained by selective protectionist

motivations on the part of country governments. Separately, it may be imagined that nonmember countries would have an incentive to monitor and ensure the full dismantling of internal trade barriers within PTAs. However, it is also quite likely that the welfare of countries outside the union is higher when the discrimination against them is lower (i.e., when internal preferences are less than complete). *Ex post*, the external monitoring incentive is therefore minimal. On the question of the timing and the phasing out of internal barriers to trade, GATT rules, rather than requiring an immediate removal of internal barriers in a PTA, allowed for this to take place within a "reasonable length of time," once again permitting substantial ambiguity in understanding and room for abuse.[13]

On external tariffs, the original GATT requirement was that external barriers not be more restrictive than initially. For FTAs, since countries retain individual tariff vectors, this could be taken to imply that no tariff was to rise. For CUs, since a *common* external tariff was to be chosen, and initial tariffs on the same good likely varied across countries, the tariff vector would necessarily change for each country. The expectation was then that the "general incidence" of trade barriers would not be higher or more restrictive than before. Given the imprecise phrasing, there was once more substantial ambiguity as to what is implied – should the common external tariff equal the unweighted mean of initial tariffs in the member countries? Should it be the trade-weighted mean? Or something else? As Dam (1970), Bhagwati (1993) and several others have noted, it is clear that Article XXIV's ambiguity in this regard left plenty of room for opportunistic (i.e., protectionist) behavior by member countries against nonmembers. The 1994 "Understanding on the Interpretation" of Article XXIV issued by the GATT provided substantial clarity on the issue of measurement and choice of the common external tariff – indicating that the GATT secretariat would compute weighted average tariff rates and duties collected in accordance with the methodology used in the assessment of tariff offers in the Uruguay Round of trade negotiations and examine trade flow and other data to arrive at suitable measures of non-tariff barriers. While this relieves, at least partially, the issue of measurement of external barriers and the comparison with barriers in place initially, the economic concern regarding trade diversion is not addressed. Clearly, leaving external barriers at their initial level and removing internal barriers does not eliminate trade diversion (as theoretically required in the Kemp–Wan and Panagariya–Krishna constructions of welfare improving PTAs). Indeed, with this configuration, trade diversion is practically guaranteed.

Having pointed to the deficiencies in existing GATT regulations in relation to the elimination of trade diversion, it may be noted that picking or designing tariff vectors *ex ante* that would ensure *zero* trade diversion, good by good, is a rather difficult task; the necessary measures of the exact sensitivity of external trade flows to external barriers of the CU or the FTA would be hard if not impossible to estimate accurately. So there is little prospect of identifying the exact trade-diversion-eliminating Kemp–Wan tariff vector and implementing it in practice.[14] Nevertheless, designing other disciplines to *minimize* diversion is less difficult; one can certainly say that lowering external barriers simultaneously with the formation of a CU or an FTA is likely to lower the degree of trade diversion (by minimizing the substitution away from the goods supplied by the rest of the world to within-union goods). McMillan (1993) has suggested as a test of admissibility of any PTA the measurement (estimation) of whether that PTA will

result in less trade with the rest of the world.[15] In a similar spirit, Bhagwati (1993), has suggested that the requirement of a simultaneous *pro rata* reduction of external trade barriers with the progressive elimination of internal barriers could replace the current requirements.

4 Preferential Trade Agreements and the Multilateral Trade System

The literature on PTAs has focused considerable attention on "static" welfare questions asking, as did Viner (1950), what would happen to welfare if arbitrarily specified FTAs or CUs were to occur. More recently, however the theoretical literature has taken a quite different approach by examining various "dynamic" aspects relating to the expansion of such agreements to include ever more members until (potentially) multilateral free trade is achieved. Space constraints preclude a detailed discussion and we will focus here only on a small number of contributions in this area. In particular, we will discuss here some recent papers that have variously examined political-economy issues – asking what the political-economy incentives are for arriving at such arrangements in the first place and, importantly, what the incentives are (if any) for the successive expansion of such agreements.[16]

We can best answer the "positive" question of why countries enter PTAs by carefully considering the domestic determinants of trade policy. The large and growing "interest group" literature on the political-economy of trade policy has argued that the governments that make it are less concerned with maximizing the nation's welfare than with advancing their own interests, interests that are best served by bending to the wishes of the interest groups and lobbies representing different sectors of the economy. Grossman and Helpman (1995) and Krishna (1998) both model the influence of powerful producers in decision making over entry into a PTA, and while the models and analytic frameworks differ in detail, they come to a similar and striking conclusion, that PTAs that divert trade are more likely to win internal political support. This is so because governments must respond to conflicting pressures from their exporting sectors, which gain from lower-trade barriers in the partner, and their import-competing sectors, which suffer from lower-trade barriers at home, when deciding on whether to form or enter a PTA. As Krishna (1998) argues, trade diversion effectively shifts the burden of the gain to member-country exporters off member-country import-competing sectors and onto nonmember producers, who have little political clout inside the member countries. Krishna (1998) also argues that such PTAs will lower the incentives for any subsequent multilateral liberalization – producers in trade diverting PTAs may oppose multilateral reform since this would take away the gains from benefits of preferential access that they enjoyed in the PTA that diverted trade to them. Under some circumstances, he shows, the incentive for further multilateral liberalization is completely eliminated.

Levy (1997) models trade policy as being determined by majority voting. Countries are assumed to differ (at least potentially) in their endowments of factors (labor and capital). Countries are also assumed to produce different varieties of goods – so that trade reform will result in gains to individuals due to the greater number of varieties

that are available for consumption. However, it should also be clear that any changes in trade policy result in changes in the distribution of income (by altering the relative rewards to the different factors of production). The arguments that emerge out of this framework are as follows. First, preferential trade integration with partners with similar relative factor endowments (i.e., with similar capital–labor ratios) is more likely to receive majority support – since this results in minimal income redistribution and still provides variety gains from trade. Second, bilateral agreements could render infeasible multilateral liberalization (which, even if they bring greater variety gains, would involve trade with countries with quite different relative endowments of capital and labor and could therefore result in much more drastic income redistribution).

Both sets of papers we have discussed above argue that bilateral agreements could impede progress toward multilateral free trade and thus undermine the multilateral trade system.[17] McLaren (2002) provides an ingenious analysis of the role of sunk costs and trade policy determination that reaches similar conclusions. He argues, roughly speaking, that the expectation of a PTA could induce agents in the economy to undertake costly and irreversible investments that makes the members within the bloc more specialized toward each other and less so toward the rest of the world (put differently, they increase dependence on each other and lower it toward the rest of the world). *Ex post* multilateral free trade is less desired.

Baldwin (1995), on the other hand, argues that PTA expansion could have "domino" effects – increasing the size of a bloc increases the incentive for others to join it (as they then gain preferential access to increasingly large markets). Assuming open-membership rules (i.e., insiders do not oppose the entry of new members who abide by the same rules as the members), the successive expansion of the PTA could then lead to multilateral free trade – a conclusion that is also reached by the work of Yi (1996) which uses a quite different economic model but addresses similar issues.

5 Conclusions

PTAs are a quintessential example of second-best economic contexts – their analysis is complex and subject to considerable ambiguity. It should be unsurprising then that economists and policy makers stand so divided on the question of the merits of PTAs. Nevertheless, the fundamental analytical insights of Viner (1950) and Kemp and Wan (1976) have helped to substantially clarify welfare-theoretic thinking in this area and to inform us on both the pitfalls of preferential trading and ways in which these can be side-stepped. While these insights do not appear to be fully reflected in existing WTO rules on PTAs, they remain central to ongoing discussions. It can be hoped that these ideas will shape future regulations on PTA admissibility and design.

An important concern we have expressed in this survey relates to the more general impact of PTAs on the world trading system. Many of the papers we have discussed have taken a sceptical view, arguing that there are plenty of reasons to worry that existing PTAs, rather than providing momentum toward multilateral free trade, will result in a fractionalized trading world. It has also been suggested that open-membership rules will likely guide countries toward more benign outcomes. While the literature has provided

some systematic ways to address these issues, much work remains to be done before answers can be provided with confidence.

Acknowledgments

This chapter was written when I was a visiting fellow at the International Economics Section (IES) in the Economics Department at Princeton University whose hospitality, intellectual, and financial support are gratefully acknowledged. I am grateful as well to Jagdish Bhagwati and Arvind Panagariya for many helpful discussions on this topic.

Appendix A.1

To relate the welfare expression (3) to country-pair characteristics that can be econometrically estimated (own- and cross-price effects resulting from preferential tariff liberalization), we make use of the fact that the compensated import demand functions M_B and M_W themselves are a function of prices and welfare and can be expressed as

$$M_B = M_B(1, 1 + t_B, 1 + \bar{t}_W, U) \tag{A.1}$$

and

$$M_W = M_W(1, 1 + t_B, 1 + \bar{t}_W, U). \tag{A.2}$$

Totally differentiating the above expressions gives us

$$dM_B = M_{BB}dt_B + M_{BU}dU \tag{A.3}$$

and

$$dM_W = M_{WB}dt_B + M_{WU}dU, \tag{A.4}$$

where M_{BB} and M_{WB} denote the own- and cross-price effects on imports from B and W respectively (i.e., the impact on imports from B and W resulting from preferential tariff liberalization by A on imports from B).

Substituting (A.3) and (A.4) into (3) gives us,

$$(E_U - t_B M_{BU} - \bar{t}_W M_{WU})dU = (t_B M_{BB} + \bar{t}_W M_{WB})dt_B. \tag{A.5}$$

Since E is homogenous of degree one in prices, E_W is also homogenous of degree one in prices. Using Euler's theorem, we then have

$$E_U = E_{AU} + (1 + t_B)E_{BU} + (1 + \bar{t}_W)E_{WU}. \tag{A.6}$$

Substituting (A.6) into (A.5) gives us a final expression for welfare

$$HdU = (t_B M_{BB} + \bar{t}_W M_{WB})dt_B \tag{A.7}$$

where H is given by

$$H = E_{AU} + E_{BU} + E_{WU} = \left(\frac{\partial(E_A + E_B + E_W)}{\partial I}\right) E_U, \tag{A.8}$$

and I denotes income (which is equal to expenditure E). Equations (A.7) and (A.8) may be understood as follows. In (A.8), the partials $\partial E_i/\partial I$ denote the extent to which consumption of good i increases with a unit increase in income (holding prices fixed). Given that E_U is simply the inverse of the marginal utility of income, H then denotes the total increase in the real value of the consumption basket (i.e., the value of the consumption basket at world prices) corresponding to a unit increase in welfare. As Dixit and Norman (1980) note, HdU then simply represents real-income change. We therefore interpret the expression on the right-hand side of (A.7) as the real-income change associated with the tariff reduction dt_B – with H serving to translate welfare changes into real-income changes. Note that if all goods are normal in consumption (as will be assumed here), then $\partial E_i/\partial I > 0, \forall i$ implying that H is positive as well.

Note that (A.8) implies that the term "H" is independent of the particular bilateral tariff reduction that is being considered. In other words, if we were to compare the welfare effect of a preferential reduction in tariffs by A against Country B, with the welfare effect when tariffs are preferentially reduced against some other Country C instead, a comparison of the right-hand side of (A.7) in the two cases would suffice to establish a welfare ranking. Of course, to estimate the right-hand side of (A.7) for preferential tariff reductions against each potential partner country, we need to estimate the own price and cross price effects on imports from the partner country and the rest of the world in each case.

Notes

1 More specifically, FTAs and CUs may be formed under Art. XXIV of the GATT, which permits PTAs provided that trade barriers between members are eliminated on substantially all trade – i.e., that the preferences are complete. Additional exceptions to the GATT's principle of nondiscrimination include the Enabling Clause, which permits developing countries to grant to each other whatever preferences they may agree to and the Generalized System of Preferences which allows developed countries to grant preferential access to developing countries.

2 It is perhaps worth pointing out that NAFTA and the EEC were formed under Art. XXIV of the GATT and that MERCOSUR, on the other hand, was formed under the Enabling Clause we have referred to earlier.

3 The magisterial survey by Panagariya (2000), which provides a more technical and comprehensive coverage than is given here, and Bhagwati's (1993) superb overview of economic and political issues in the debate over PTAs are both greatly recommended to interested readers. The survey by Baldwin and Venables (1995) is also recommended for its treatment of issues relating to imperfect competition. The edited volume by Bhagwati, Panagariya and Krishna (1999) collects a number of the central contributions in this literature including many that are discussed here in this chapter.

4 Thus, the partner Country B and the rest of the world, W, produce goods that are not perfect substitutes for each other – in contrast with the previous Vinerian analysis.

5 Appendix A.1 processes these expressions further to relate the trade creation and trade diversion terms in (3) to own- and cross-price elasticities of import demand in the home Country A. These elasticities may be estimated using standard econometric techniques in order to arrive at estimates of the trade creation and diversion that can be expected to result from preferential tariff liberalization by A. The "natural trading partners" hypothesis, which asserts that preferential liberalization toward geographically proximate partners is more likely to be welfare improving, has recently been evaluated by Krishna (2003), who conducts just such an estimation exercise for the United States and finds no support in the data for it.

6 Also, we have not considered the possibility that there may be extra-union terms of trade effects. As the analysis of Mundell (1964) has shown, allowing for extra-union terms of trade effects may complicate matters further: the terms-of-trade of the tariff-reducing country with respect to the rest of the world may rise or fall following a preferential reduction in its tariffs against a particular partner. On this point see also the recent analysis of Panagariya (1997).

7 The proof outlined here regarding the design of welfare improving CUs is borrowed from Panagariya and Krishna (2002) and is quite different from the one presented by Kemp and Wan. As will become clear from our discussion later, the Panagariya–Krishna framework allows a relatively easy extension to the case of FTAs whereas the Kemp–Wan methodology does not.

8 Since it is only the joint welfare of the union countries that is improved by the formation of the CU, we need to rely on a system of lumpsum internal payments here to guarantee that *each* member country actually benefits from it. See, however, Grinols (1981), who derives a suitable scheme of internal payments that does not have this lumpsum feature and is based on preunion observables. In a quite different extension of the Kemp–Wan results, Krishna and Bhagwati (1997) have shown that a variety of noneconomic objectives can be accommodated within the CU without diminishing its Pareto-improving virtues. They have shown that if, for instance, member governments were keen to maintain factor employment levels in particular industries at the same level as before, this can be achieved through the same Kemp–Wan external tariff *scheme* and the use of suitable production/factor employment subsidies internally.

9 For a more detailed discussion of the exact theoretical context and rules which suffice to prevent transhipment, see section 5 in Panagariya and Krishna (2002). For a general discussion on rules of origin in free trade areas, see also Krishna and Krueqer (1995).

10 It should, of course, be clear by now that with the Kemp–Wan CU the trade vector of the member countries is restricted to be the same as initially for the union members as a whole. With FTAs, this constraint needs to be met for each member country separately.

11 Our coverage of these issues will be quite brief. See Chapter 7 by Mavroidis in this volume for a detailed discussion on many of the relevant economic and legal issues surrounding the implementation of PTAs in the GATT/WTO.

12 It is important to keep in mind here that the elimination of internal tariffs maximizes the welfare of member countries for a given level of external trade (as in Kemp–Wan) and not for a given level of the external tariffs. With fixed tariffs, member country welfare may well be maximized with internal tariffs that are non-zero.

13 The more recent "Understanding on the Interpretation" of Art. XXIV issued in 1994 clarifies that the "reasonable length of time" should exceed 10 years in only "exceptional cases."

14 See, however, the paper by Srinivasan (1997) which attempts to identify and characterize the Kemp–Wan tariff vector in the context of a particular economic model.

15 Of course, the Kemp–Wan and Panagariya–Krishna schemes both require that the PTA trade *exactly* the same amount as before. A PTA that trades no less, as in the McMillan test, is not necessarily welfare improving, as Winters (1997) has argued.

16 In focusing on the positive aspects, we are neglecting a large normative literature that has recently examined the "dynamic" welfare impact of exogenously formed and expanded trade blocs, exemplified by the work of Krugman (1993) and Frankel, Stein, and Wei (1995). For discussions of the sensitivity of these frameworks to modeling assumptions, see Bond and Syropoulos (1996), Deardorff and Stern (1993) and Srinivasan (1993). Other recent "dynamic" analyses include Bagwell and Staiger (1997), which explicitly considers the impact of FTAs and CUs on multilateral tariff cooperation during the transition period when the agreed-upon lowering of tariffs within the FTA and CU are implemented. Finally,

Ethier (1998) and Freund (2000) study the interaction between multilateral liberalization and the incentives to form/sustain PTAs with the former paper considering additionally North–South differences in this regard.

17 A somewhat parallel literature has asked the question of what external tariffs will be chosen by PTA members – examining, in particular, whether external tariffs can be expected to rise or fall. No clear answers to this question emerge. While Panagariya and Findlay (1996), employing a lobbying model find that external tariffs rise after tariff preferences are granted, Cadot, de Melo, and Olarreaga (1996) find that external tariffs fall for an FTA but rise for a CU. See also Dinopoulos and Syropoulos (2002) who allow for tariff endogeneity in examining the question of whether trade blocs impede global free trade and arrive at mixed results.

References

Bagwell, K. and R. Staiger 1997: Multilateral cooperation during the formation of free trade areas. *International Economic Review*, 38(2), 291–319.

Baldwin, R. 1995: A domino theory of regionalism. In R. Baldwin, P. Haaparanta, and J. Kiander (eds.), *Expanding Membership of the European Union*. Cambridge University Press.

Baldwin, R. and A. Venables 1995: Regional economic integration. In Gene Grossman and Kenneth Rogoff (eds.), *Handbook of International Economics*, Volume III. Elsevier Science BV.

Bhagwati, J. 1993: Regionalism and multilateralism: an overview. In Jaime de Melo and Arvind Panagariya (eds.), *New Dimensions in Regional Integration*. Cambridge University Press.

Bhagwati, J. and A. Panagariya 1996: Free trade areas or free trade? The economics of preferential trade areas. *American Enterprise Institute*. Washington, DC: AEI Press.

Bhagwati, J., A. Panagariya, and P. Krishna 1999: *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements*. MIT Press.

Bond, E. and C. Syropoulos 1996: The size of trading blocs, market power and world welfare effects. *Journal of International Economics*, 40(3), 411–438.

Cadot, O., J. de Melo, and M. Olarreaga 1999: Regional integration and lobbying for tariffs against non-members. *International Economic Review*.

Dam, K. 1970: *The GATT: Law and International Economic Organization*. University of Chicago Press.

Deardorff, A. and R. Stern 1994: Multilateral trade negotiations and preferential trading arrangements. In Alan Deardorff, and Robert Stern (eds.), *Analytical and Negotiating Issues in the Global Trading System*. Ann Arbor, MA: University of Michigan Press.

Dinopoulos, E. and C. Syropoulos 2002: *Bilateralism: "Stepping Stone" or "Stumbling Block" to "Global Free Trade"?*, Mimeo.

Dixit, A. and V. Norman 1980: *International Trade*. Cambridge University Press.

Ethier, W. 1998: Regionalism in a multilateral world. *Journal of Political Economy*, 106(6), 1214–1245.

Frankel, J., E. Stein, and S. Wei 1995: Trading blocs: the natural, the unnatural and the super natural? *Journal of Development Economics*, 47(1), 61–95.

Freund, C. 2000: Multilateralism and the endogenous formation of PTAs. *Journal of International Economics*, 52(2), 359–376.

Grossman, G. and E. Helpman 1995: The politics of free trade agreements. *American Economic Review*, 85, 667–690.

Grinols, E. 1981: An extension of the Kemp–Wan theorem on the formation of customs unions. *Journal of International Economics*, 11, 259–266.

Kemp, M. 1964: *The Pure Theory of International Trade* N.J.: Prentice-Hall.

Kemp, M. and H. Wan 1976: An elementary proposition concerning the formation of customs unions. *Journal of International Economics*, 6(1), 95–8.

Krishna, K. and Krueqer, A. 1995: "Implementing Free Trade Areas: Rules of origin and hidden protection," in Deardorff, Levinsonn and Stern, eds., *New Directions in Trade Theory*, University of Michigan Press, Ann Arbor.

Krishna, P. 1998: Regionalism and multilateralism: a political economy approach. *Quarterly Journal of Economics*, 113, 227–251.

Krishna, P. 2003: Are regional trading partners "natural"? *Journal of Political Economy*, 111(1), 202–231.

Krishna, P. and J. Bhagwati 1997: Necessarily welfare improving customs unions: the Cooper–Massell–Johnson–Bhagwati conjecture. *Japan and the World Economy*, 9(4), 441–446.

Krugman, P. 1993: Regionalism versus multilateralism: analytical notes. In Jaime de Melo and Arvind Panagariya (eds.), *New Dimensions in Regional Integration*. England: Cambridge University Press, 84–9.

Levy, P. 1997: A political economic analysis of free trade arrangements. *American Economic Review*, 87(4), 506–519.

McMillan, J. 1993: Does regional integration foster open trade? Economic theory and GATT's Article XXIV. In Anderson and Blackhurst (eds.), *Regional Integration and the Global Trading System*, N.Y.: St. Martin's Press.

McLaren, J. 2002: A theory of insidious regionalism. *Quarterly Journal of Economics*, 2, 571–608.

Mundell, R. 1964: Tariff preferences and the terms of trade. *Manchester School of Economic Studies*, 1–13.

Ohyama, M. 1972: Trade and welfare in general equilibrium. *Keio Economic Studies*, 9, 37–73.

Panagariya, A. 1997: Preferential trading and the myth of natural trading partners. *Japan and the World Economy*, 9(4), 471–489.

Panagariya, A. 2000: Preferential trade liberalization: the traditional theory and new developments. *Journal of Economic Literature*, 38(2), 287–331.

Panagariya, A. and R. Findlay 1996: A political economy analysis of free trade areas and customs unions. In Feenstra, Irwin and Grossman (eds.), *The Political Economy of Trade Reform: Essays in Honor of Jagdish Bhagwati*.

Panagariya, A. and P. Krishna 2002: On the existence of necessarily welfare improving free trade areas. *Journal of International Economics*, 57, 353–367.

Srinivasan, T. N. 1993: Comment on regionalism versus multilateralism: analytical notes. In Jaime de Melo and Arvind Panagariya (eds.), *New Dimensions in Regional Integration*, England: Cambridge University Press, 84–9.

Srinivasan, T. N. 1997: Common external tariffs of a customs union: the case of identical Cobb Douglas tastes. *Japan and the World Economy*, 9(4), 447–465.

Vanek, J. 1965: *General Equilibrium of International Discrimination*. Harvard University Press.

Viner, J. 1950: *The Customs Unions Issue*. New York: Carnegie Endowment for International Peace.

Winters, A. 1997: Regionalism and the rest of the world: the irrelevance of the Kemp–Wan theorem. *Oxford Economic Papers*, 49(2), 228–234.

Yi, S. 1996: Endogenous formation of customs unions under imperfect competition: open regionalism is good. *Journal of International Economics*, 41, 153–177.

10

Conditionality, Separation, and Open Rules in Multilateral Institutions

Paola Conconi and Carlo Perroni

Chapter Outline

We examine the implications for the viability of multilateral cooperation of different legal principles governing how separate international agreements relate to each other. We contrast three alternative rules: conditionality – making cooperation in one area a condition for cooperation in another; separation – forbidding sanctions in one area to be used to enforce cooperation in others; and open rules, i.e., the absence of any restriction on the patterns of cross-issue cooperation arrangements and sanctions. As an example, we focus on trade and environmental agreements. Our analysis suggests that conditionality is more likely to facilitate multilateral, multi-issue cooperation in situations where the environmental policy stakes are small relative to the welfare effects of trade policies; when the costs of environmental compliance are high, a conditionality rule can hinder multilateral cooperation. Separation can undermine cooperation by limiting punishment, but can also promote broad cooperation by making partial cooperation more difficult to sustain.

1 Introduction

The increasing economic and social interdependence between nations has intensified the need for international coordination of policies, both in the economic sphere and

outside it. This is achieved through internationally agreed upon rules of international law, often embodied in international institutions, such as GATT/WTO, the International Monetary Fund (IMF), the International Bank for Reconstruction and Development (IBRD), various UN bodies such as UNCTAD and UNDP, as well as specialized organizations such as the Bank for International Settlements (BIS), WCO, ITU, WIPO, FAO, and the ILO.

Typically such rules and institutions each focus on a specific policy dimension, and are the result of separate negotiations. In some cases, however, multiple policy dimensions are dealt with by a single piece of international legislation or covered by a single agreement. For example, the World Trade Organization (WTO) provides the framework for the conduct of international trade in goods and services and for the protection of intellectual property rights. In other cases, agreements and rules in different areas remain formally distinct but contain provisions regulating their interaction. For example, the General Agreement on Tariffs and Trade (GATT) allows member countries to take measures that it considers necessary for the protection of its essential security interests as well as those taken in pursuit of its obligations under the United Nations Charter for the maintenance of international peace and security. Sometimes we observe distinct rules and agreements, which contain no formal mention of each other, coming into being as a result of joint negotiation. An example is the changes in patent and royalty regulations, energy pricing, and energy supply arrangements that were negotiated between Canada and the United States in the 1980s concurrently with (albeit not jointly with) the Canada–US Free Trade Agreement (Perroni and Whalley, 2000).

The examples just described represent instances of "linkage" between different dimensions of international cooperation. What does economic theory have to say about such linkages? One of the key insights that economics has brought to the understanding of laws and institutions is the idea of Coasian bargaining. This is the notion that, whatever the initial distribution of legal rights, the interested parties can arrive at a mutually advantageous (and economically efficient) exchange of concessions through bargaining – as long as the pattern of concessions is unconstrained. If multiple issues are on the table, then efficient bargaining may involve exchange of concession across different issues. Thus, an economist's idealized picture of the process of international cooperation consists of a single, all-embracing round of negotiation giving rise to a single, all-embracing cooperation agreement.

In reality, a number of practical considerations prevent this idealized outcome from materializing. To begin with, the number of channels through which countries can affect each other is limited, which in turn limits the possibility of finding mutually advantageous forms of exchange; direct cash payments could, in principle, overcome such limitations, but are, in practice, seldom used to settle international disputes. Negotiations also entail transaction costs, which become more significant the more complex negotiation is – with respect to both the number of parties involved and the number of issues covered; as a result, more limited forms of cooperation may be preferred to broader agreements. Finally, the issues over which international coordination is required cannot all be on the table before negotiations take place: new issues arise over time, and negotiation must therefore evolve through progressive steps, each adding to (and possibly modifying) previous arrangements.

The actual picture of the institutions through which international cooperation takes place bears little resemblance to the economist's idea of a single all-encompassing arrangement; rather, it is a patchwork of preferential agreements and issue-specific treaties. International cooperation is limited to certain areas: at the time of writing, for example, there is still no fully functioning international agreement for curbing greenhouse emissions. Even when simultaneous cooperation across different areas is achieved, it is usually dealt with by separate rules and agreements. Not only are certain areas excluded from existing agreements; countries also are excluded. Membership in the WTO, for example, is still limited, and there exist a number of preferential agreements (trade and otherwise) involving a small number of partners. In some cases, a country cooperates with a certain set of partners in a certain area, and with a different set of partners in another.

Nevertheless, even if in practice cooperation must be pursued in a piecewise fashion, the results achieved by a given web of arrangements can still be evaluated against the results that could be hypothetically achieved through a multilateral, multi-issue agreement: replicating such an outcome, within the constraints imposed by the given available instruments and institutions, remains desirable on grounds of economic efficiency. The question then naturally arises of how cooperation agreements in different areas should be interfaced so as to facilitate the establishment of multilateral cooperation across multiple areas, given that the possibilities of cooperation in one are not independent of cooperation in the other. For example, since the Uruguay Round of GATT/WTO trade negotiations there has been an intense debate over the potential overlaps between trade and non-trade objectives, especially with respect to environmental protection and workers' rights. Should trade sanctions be used to buttress environmental policy cooperation or promote more stringent labor standards? Or should the GATT/WTO forbid the use of trade sanctions to enforce non-trade agreements?

In this chapter, we examine the implications for the viability of multilateral cooperation of different legal principles governing how separate international laws and agreements relate to each other. We contrast three different possible legal regimes: conditionality – making cooperation in one area a condition for cooperation in another; separation – forbidding sanctions in one area to be used to enforce cooperation in others; and open rules, that is, the absence of any restrictions on the patterns of cross-issue cooperation arrangements and sanctions.

To analyze the impact of these different legal regimes on multilateral negotiations and the prospects for achieving cooperation over multiple issues, we focus on a scenario where countries are linked by international trade and transboundary pollution. Following the theoretical framework presented in Conconi and Perroni (2001), we describe a setting where countries can enter into selective and separate binding agreements with different partners along different policy dimensions, and examine the effects of conditionality/separation/open rules on the robustness of simultaneous multilateral cooperation arrangements on both trade and environmental policies.

We conclude that conditionality is more likely to facilitate multilateral, multi-issue cooperation in situations where the environmental policy stakes are small relative to the welfare effects of trade policies and when partial environmental coordination is preferred to no cooperation by all countries involved – implying that outsiders can free-ride effectively on partial environmental agreements. On the other hand, when

the costs of environmental compliance are high but the ability to free-ride on partial environmental agreements is limited, a conditionality rule can hinder multilateral cooperation by making it both attractive and viable for a single country to remain outside of any agreement. Separation can undermine cooperation because it limits the channels through which countries can "punish" defectors. However, for the same reason, separation can undermine the viability of coordinated responses to unilateral defections countries, and, through this channel, make defections less attractive. Again, the latter scenario requires that the incentives to free-ride on an environmental agreement be sufficiently large in comparison with the gains from trade policy cooperation. The broad message that emerges from this analysis is that one cannot draw general conclusions about the effects on multilateral negotiations of different legal linkage rules: whether the adoption of a certain rule is appropriate for certain areas depends on the specific structure of cooperation incentives in those areas.

The remainder of the chapter is organized as follows. Section 2 briefly reviews the economics literature on issue linkage. Section 3 examines the impact of different negotiation rules on the sustainability of multilateral cooperation. Section 4 looks at examples of how such negotiation rules are embodied in different treaties and institutional arrangements. Section 5 briefly discusses how taking into account the emergence of new issues, transaction costs, and the time horizon of policymakers could affect our analysis. Section 6 summarizes and concludes.

2 The Literature on Issue Linkage

The theoretical literature on cross-issue international cooperation has grown considerably in recent years, but the term "issue linkage" is used in this literature to denote a number of very different mechanisms.

A strand of this literature has stressed *direct strategic linkages*, in the sense that different policy instruments may be either close substitutes (meaning that commitments in one policy dimension can easily be undone by adjustments in the other) or complements (meaning that commitments in one policy dimension are ineffective unless accommodated by adjustments in the other). Here, there is clearly a need for joint negotiation: if two types of policies are perfect substitutes, then cooperation agreements reached over a single issue will be fully ineffective absent cooperation in the other. Copeland (1990), for example, examined the case of tariff negotiations when a secondary form of trade barrier (e.g., a non-tariff barrier) is set non-cooperatively. He shows that trade liberalization will induce substitution toward the less efficient, nonnegotiable instrument of protection, due to countries facing unilateral incentives to maintain trade protection. In this framework, the two types of trade instruments are imperfect substitutes for each other, and lack of cooperation with respect to one instrument directly undermines any cooperation achieved with respect to the other.

Issue linkage has also been proposed as a *means of handling asymmetries among countries*, when meaningful exchange of concessions can only take place if more that one issue is involved. As noted above, this idea stems from the more general notion of Coasian bargaining. It was first stressed by Raiffa (1982) and Sebenius (1983) in the political science literature, and by Folmer et al. (1993), Cesar and Zeeuw (1996), and

Abrego et al. (2001) in the economics literature. Although monetary side-payments across countries represent a more efficient means of facilitating international cooperation in the presence of countries' asymmetries, issue linkage will be beneficial when such side-payments are not available.

Another strand of literature has analyzed linkage in situations where agreements cannot be enforced by an external authority, and must be therefore self-enforcing through repeated interaction between countries by the threat of reversion to non-cooperation. In this context, a linked agreement is one in which a defection in any dimension of the agreement will trigger a retaliatory response over all dimensions. Spagnolo (1996) shows that removing the possibility of using *cross-issue punishment strategies* by delegating policies to independent bodies can facilitate cooperation if policy issues are complements in the governments' objective function, but can hinder cooperation if policy issues are substitutes. Limão (2001) finds that linked punishment under repeated interaction and lobbying can lead to a "reallocation of enforcement power" from one policy dimension to the other. With reference to trade policies, Ederington (2001) shows that, when countries can use domestic policies as a secondary trade barrier, a tariff is the most efficient instrument for supporting cooperation under repeated interaction.

What most of the contributions to this literature have in common is that they focus on bilateral relations rather than on *multilateral cooperation* – a notable exception being Maggi (1999). Literature on multilateral cooperation exists, but most of this literature does not examine cooperation across multiple issues. As we explain below, cross-issue cooperation amongst multiple countries involves forms of negotiation linkage that are not present in the case of bilateral agreements.

The question of how different legal regimes, such as conditionality and separation, affect cooperation has been only touched upon in the aforementioned theoretical literature. In particular, while the implications of separation have been examined with respect to self-enforcing bilateral agreements (by the Spangolo and the Limão papers), conditionality has little meaning in a bilateral world – since preferential cooperation is not an option in such a framework – and therefore has not been studied in literature on bilateral cooperation.

3 Conditionality, Separation, and Open Rules and Their Impact on Multilateralism

In this section, we analyze multilateral policy cooperation choices when international policy coordination is required across multiple issues. Following Conconi and Perroni (2001), the establishment of cooperation arrangements can formally be described as a two-stage process in which binding agreements are formed in the first stage through a sequence of proposals and counterproposals, and policies are selected in the second stage – cooperatively among countries participating in an agreement and non-cooperatively between countries belonging to separate agreements.

To accommodate the possibility of individual countries belonging to multiple agreements, we draw a distinction between *agreements*, by which members coordinate their choices along one or more policy dimensions, and *coalitions* of countries during the

negotiation phase, which can make coordinated objections to a proposed cooperation structure. The two concepts are distinct: agreement structures determine the actual policies chosen after cooperation choices are made; whereas coalitions of players can object to a proposed cooperation arrangement by proposing an alternative cooperation arrangement, affecting the structure of agreements – which, in turn, affects policy choices. A viable agreement structure must be immune from objections, and viable objections must be not only profitable for the objecting coalitions but also immune from further external or internal objections, i.e., they must involve arrangements that are themselves immune from objections.

Using this construct, we discuss how the viability of a multilateral, multi-issue agreement (an agreement where all countries jointly cooperate across all policy dimensions) is affected by three alternative legal regimes: conditionality, separation, and open rules. With open rules, countries can enter into selective and separate binding agreements with different partners along different policy dimensions. This means that single-issue agreements can both be proposed and used as counterproposals during negotiations. Conditionality makes participation in a cooperation agreement concerning a certain policy dimension a necessary condition for being admitted as a member to a cooperation agreement in another dimension; in other words, cooperation must take place simultaneously across multiple issues, ruling out single-issue agreements. Again, such restriction will affect both the proposals that can be made and the counterproposals that can be made as objections to a proposed cooperation structure. Separation concerns the use of sanctions along a certain policy dimension against countries that do not cooperate. In our abstract framework, this means that countries that unilaterally choose not to participate in a multilateral cooperation agreement along one policy dimension cannot be "punished" by being excluded from cooperation in another.

For concreteness, the discussion that follows focuses on a specific scenario where countries are linked by international trade and transboundary pollution. In this context, it has been suggested that multilateral cooperation could be enhanced by formally combining different issues with the aim of joint settlement.[1] This would correspond to our notion of conditionality. At the same time, there has been much discussion about the legality of using trade sanctions in response to noncompliance to international environmental treaties (Trebilcock and Howse, 2001). In our framework, a ban on the use of trade sanctions to pursue environmental cooperation goals – and vice versa – would correspond to a separation regime.

As noted earlier, most of the literature on multilateral policy cooperation has examined cooperation over trade policies and over environmental policies separately from one another. Riezman (1985), Krugman (1991), Bond and Syropoulos (1996), and Yi (1996), among others, have focused on the creation of Customs Unions, while Chander and Tulkens (1992), Carraro and Siniscalco (1993), and Barrett (1994), among others, have focused on International Environmental Agreements. The broad theme emerging from this literature is that the presence of spillovers between coalitions (positive in the case of environmental coalitions, negative in the case of trade coalitions) makes global cooperation difficult to sustain, and that partial cooperation, restricted to subsets of countries, is more likely to emerge. In our multi-issue cooperation framework, however, the effects of spillovers in different areas become linked, because cooperation incentives in one area can help support cooperation in the other.

Trade policies and environmental policies involve different cooperation incentives and generate different forms of "temptation" to deviate from a multilateral, multi-issue cooperation arrangement. When facing global environmental problems, a country may find it tempting to defect from multilateral environmental cooperation, free-riding on the abatement in global emissions that the remaining countries carry out. In the case of trade policies, a subset of countries may find it optimal to jointly defect from multilateral trade cooperation and form a preferential trade agreement, exploiting the resulting increase in international market power of the trade bloc to the advantage of its members.

In the first case, unilateral free-riding may be avoided if the remaining countries find it optimal, once a country has defected, to dissolve environmental cooperation amongst them, which effectively undermines the free-riding attempt. Alternatively, suppose that forming a preferential trade agreement is not attractive if all countries cooperate in environmental policies but becomes attractive if the excluded country also chooses not to participate in environmental policy cooperation; then, a unilateral defection from environmental cooperation could make a preferential trade agreement between the remaining countries viable, and a country defecting from environmental cooperation could thus be credibly excluded from trade cooperation. If the loss from such exclusion is large enough in comparison with the temptation to free-ride on a multilateral environmental agreement, then multilateral environmental policy cooperation could be sustained.

When the threat to multilateral cooperation comes from incentives to form a preferential trade agreement, environmental policy cooperation incentives could come into play as follows. Suppose that, in the presence of multilateral trade cooperation, a unilateral defection from environmental policy cooperation triggers a dissolution of the environmental agreement amongst the remaining countries, and is therefore not viable; but that, if one country is excluded from trade cooperation, the remaining countries no longer find it optimal to respond to a unilateral defection from an environmental agreement by fully dissolving the agreement. In such a scenario, excluding a country from trade cooperation induces the excluded country to abandon environmental policy cooperation. Then, the temptation to form a preferential trade agreement may be offset by the need to secure broad environmental policy compliance.

Which of these scenarios will prevail depends on the specific structure of environmental and trade policy cooperation incentives. For example, the presence of an environmental policy dimension can do little to help support trade cooperation if the costs and gains stemming from trade policies dominate the costs and gains stemming from environmental policies.

The preceding discussion of the linkages between trade and environment cooperation has focused on a situation where there are no constraints on cooperation patterns: unilateral defections are possible along a single policy dimension, and countries can respond to defections along one dimension by modifying their cooperation arrangements along the other. The presence of a conditionality rule or a separation rule, on the other hand, will limit the range of possible defections (objections) and responses (counterobjections); this will in turn determine which of the above negotiation linkages can be functional.

Consider, first, conditionality. Starting from a situation where all countries cooperate in both trade and environmental policies, a country will face a temptation to defect from

environmental cooperation, but it will lose from being left out of a preferential trade agreement. If the second effect dominates, then conditionality, by linking defection from an environmental agreement to defection from multilateral trade cooperation, can lock-in an otherwise reluctant partner within a multilateral, multi-issue cooperation arrangement. On the other hand, if the costs of environmental compliance are large in comparison with the gains from membership in a multilateral trade agreement, then conditionality may not be enough to prevent free-riding in environmental policies. It could, however, limit countries' ability to respond to single-country defections. Given that a country does not participate in the environmental agreement, it may be in the remaining countries' interest to limit their cooperation to trade policy only, dismantling their remaining environmental policy cooperation ties; this, in turn, could undermine the initial free-riding attempt and induce multilateral compliance. Conditionality, by tying cooperation in trade policies to cooperation in environmental policies, renders such a response nonfeasible, and can thus make unilateral defections viable.

Similarly, if two countries are tempted to form a preferential trade agreement, conditionality implies that by doing so they have to forgo environmental cooperation with a third country; this may induce them to remain in a multilateral agreement instead. However, suppose now that environmental cooperation incentives are not strong enough to lock-in the two countries within a multilateral arrangement in the presence of a conditionality rule. Then, conditionality could constrain the response of the remaining countries to the formation of a preferential trade bloc, and hence make it easier to be established.

A separation regime also limits the pattern of possible responses, as conditionality does, but in a different way: if a country defects from an environmental agreement, it cannot as a result be excluded from a preexisting trade agreement, and vice versa. Such separation of responses may undermine cooperation by making unilateral defections from an environmental treaty less costly for defectors; or, it may have no effect on cooperation outcomes if the temptation to defect along the environmental policy dimension is large in comparison to the trade sanctions that would be incurred under open rules. Similarly, if we consider the temptation for two countries to form a preferential trade bloc, separation could make it nonfeasible for the excluded country to respond by defecting from the environmental agreement, which in turn could make the formation of the trade bloc more viable. In the situations just described, separation can hinder multilateral cooperation.

However, it is also possible for a separation regime to produce the opposite effect. For example, suppose that a country is tempted to free-ride on an environmental agreement only if the remaining countries manage to maintain environmental policy cooperation amongst themselves (finding it optimal to do so), but not if a unilateral defection causes the environmental agreement to unravel. In this situation, a ban on trade sanctions could make it more difficult to maintain environmental cooperation amongst the remaining countries and may therefore discourage free-riding in the first place. Again, this can only be the case if free-riding incentives are large relative to trade policy cooperation incentives.

The overall conclusion we can draw from the above analysis is that rules must be assessed not just with respect to their effects on defections from multilateral cooperation (the objections), but also for their effect on the viability of coordinated responses to

such defections (counterobjections). In some cases, a rule may limit the temptation to defect – thus facilitating the establishment of multilateral cooperation; in others, the same rule may limit the range of possible responses to defectors – thus hindering the viability of multilateral cooperation arrangements.

Conditionality is more likely to facilitate joint multilateral cooperation over trade and environmental policies in cases where the environmental policy stakes are small relative to the effects of trade policies, and when partial environmental coordination is preferred to no cooperation by all countries involved – implying that outsiders can free-ride effectively on partial environmental agreements. On the other hand, when the costs of environmental compliance are high but the ability to free-ride on partial environmental agreements is limited, a negotiation tie-in restriction can hinder multilateral cooperation by making it both attractive and viable for a single country to remain outside of any agreement. Conversely, separation is more likely to be an obstacle to cooperation when the environmental policy stakes are small in comparison with the effects of trade policies; when the environmental policy stakes are comparatively large, on the other hand, separation may facilitate multilateral cooperation by making limited forms of environmental cooperation less viable.

If we apply these insights to the current debate about trade and environment, then we should conclude that a conditionality regime across trade and environmental agreements may be helpful to elicit cooperation for relatively "small" environmental issues; and a separation regime may be hurtful to cooperation in these cases. When larger costs and gains are involved in environmental protection – as in the case of policies for addressing global climate change – then conditionality could be a hindrance to cooperation, and a separation regime may actually be appropriate. These conclusions are consistent with what seems to be a widespread view in policy circles with respect to global climate treaties, namely that global trade and environmental institutions should remain separate rather than combined into a single WTO/World Environmental Organization (WEO) system.

4 Conditionality, Separation, and Open Rules in International Law

How can the three different linkage regimes discussed here be embodied in international laws and institutions? In this section we examine different possible approaches to the implementation of such rules, and provide examples of current practice.

4.1 Conditionality

Conditionality has often been used in the context of international aid policy. One of the problems that has affected foreign aid programs in the past is that it has been improperly targeted, with funds often ending in the hands of corrupt governments rather than being used for developmental purposes. For this reason, since the early 1990s, donors have increasingly made official development assistance conditional on political and administrative reforms, improved human rights regulation, environmental

protection, disarmament, and so forth (see Hopkins, 2000). This form of conditionality, however, does not correspond to the multilateral version of conditionality that has been discussed before, and is therefore more aptly described as "bilateral" conditionality.

Multilateral, cross-issue conditionality in international cooperation can be achieved by a negotiated rule restricting membership to a certain agreement to countries that also participate in another agreement or abide by certain international rules. This can also be implied by the agenda and format of international negotiations (see Jackson, 1997). In the Uruguay Round, for example, as against previous trade liberalization rounds, it was decided at the outset that the bulk of the multilateral agreements negotiated had to be accepted *as a whole*.[2] The "single undertaking" of the Uruguay Round includes, among others, the Agreement on Agriculture, the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), and the agreement on trade in services (GATS).[3] Similarly, certain political groups within the United States and Europe would like the next round of GATT/WTO trade negotiations in Doha, Qatar, to be a vehicle for promoting non-trade objectives – such as labor and environmental standards, gender equality, or social engineering – jointly with trade objectives.[4]

Conditionality can also take the form of rules *prescribing* that violations of international rules in a certain area be met by sanctions in other areas. Such conditionality rules can be found to be at work within certain multilateral environmental agreements (MEAs). For example, the Convention on International Trade in Endangered Species (CITES), the Basel Convention on the trade or transportation of hazardous waste across international borders, and the Nuclear Non-Proliferation Treaty do not merely permit, but *require* the use of trade restrictions as an instrument of environmental protection.

4.2 Separation

One way of achieving separation between different negotiation issues would be to delegate different policies to independent agents (see Spagnolo, 1996). Governments could thus "tie their hands" by not allowing agents to condition strategies in one game on the history of the other. It should be noted, however, that policy delegation is rarely used within international economic relations, and that international negotiations almost always involve government representatives rather than independent national agencies. One of the main problems with this separation mechanism is that delegation can only work if governments are able to *credibly* commit to it, since governments typically face ex post incentives to regain control of policies and renege on earlier promises if they can do so.

An alternative mechanism for achieving issue separation is to explicitly prescribe separation in international agreements. For example, Article XX of GATT does recognize the importance of a sovereign nation being able to pursue certain non-trade goals even when such action otherwise conflicts with various obligations relating to international trade. However, it establishes that the use of trade measures should not be "applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between countries where the same conditions prevail, or a disguised restriction to international trade." In the following, we discuss two examples of explicit

separation between trade and other policy areas – namely environmental protection and labor rights – in international treaties.

4.2.1 TRADE AND ENVIRONMENT

Among the classes of measures listed in Article XX of GATT, are those "necessary to protect human, animal or plant life or health" (XX(b)) and those "relating to the conservation of exhaustible resources if such measures are made effective in conjunction with restrictions on domestic production or consumption" (XX(g)). However, jurisprudence in this area has established a general principle of separation between trade measures and *global* – rather than domestic – environmental objectives.

The application of Article XX to the environment was addressed in several GATT and FTA panel decisions during the 1980s and early 1990s. The most influential panel ruling in this area relates to the so-called Tuna-Dolphin dispute between Mexico and Canada over the extraterritorial application of the US Marine Mammal Protection Act. This requires steps to be taken to curtail the incidental killing of marine mammals by commercial fishermen. In 1988 the US government introduced an import ban on tuna, harvested in a way that did not satisfy the standards for the protection of dolphins, applied to domestic fishermen. In 1991 a panel requested by Mexico ruled that the US ban was a violation of GATT Article III, and that the Article XX exceptions in GATT could not be applied on an extraterritorial basis, that is, GATT rules do not allow one country to take trade action for the purpose of attempting to enforce its own domestic laws in another country – even if it is to protect animal health or exhaustible natural resources. The reasoning behind this ruling was that, if the US arguments were accepted, any country could ban imports of a product from another country merely because the exporting country has different environmental, health, and social policies. This would create a virtually open-ended route for any country to apply trade restrictions unilaterally – and to do so not just to enforce its own laws domestically, but to impose its own standards on other countries. The door would be opened to a possible flood of protectionist abuses. This would conflict with the main purpose of the multilateral trading system, namely to achieve predictability through trade rules. The Panel also claimed that the language of "necessity" in Article XX(b) meant that a country would have to show that it had exhausted all options less restrictive than trade before resorting to import restrictions and noted that the United States had not exhausted the possibility of international cooperation with respect to dolphin conservation. Given this interpretation of Articles XX(b) and XX(g) by the jurisprudence, there is little textual basis for the reading that they allow trade measures to be adopted for enforcing international policy cooperation over global environmental issues.

There are also about 200 MEAs currently in force outside the WTO, dealing with various environmental issues. About 20 of these include provisions that can affect trade, that can ban trade in certain products, or can allow countries to restrict trade in certain circumstances. What is the relationship between the GATT/WTO agreement and these environmental agreements and conventions? The official position of the WTO Committee on Trade and Environment is that trade restrictions are not necessarily an effective instrument of environmental protection, and that the environmental agreements themselves, not trade sanctions, are the most appropriate

way to deal with international environmental problems, possibly in conjunction with helping countries acquire environmentally friendly technology, giving them financial assistance, providing training, and so forth.

4.2.2 TRADE AND LABOR RIGHTS

The interaction between trade policy and labor rights is one of the most contentious issues that the world trading system faces today.[5] Critics of free trade have argued that it is unfair that producers in industrialized countries should have to compete with imports from countries with much lower wage rates and poorer labor standards. In contrast, advocates of free trade often view differences in countries' labor regulations as a legitimate source of comparative advantage or disadvantage. Aside from competitiveness-based arguments, trade restrictions could also be justified as a means of inducing recalcitrant governments to meet a given set of labor standards. This may involve trade restrictions being imposed in the case of a country violating existing international agreements such as conventions of the ILO, or to induce a country to adopt new standards or norms.

GATT contains no explicit provision either permitting or requiring trade actions against labor rights violations. Article XX(e) permits otherwise GATT-inconsistent measures "relating to the product of prison labor." The possibility has been raised that Article XX(a), which permits the use of otherwise GATT-inconsistent measures "necessary to protect public morals," might be invoked to justify trade sanctions against products that involve the use of child labor or the denial of basic workers' rights. There is no GATT or WTO jurisprudence on the interpretation of Article XX(a), and on the reference to prison labor in XX(e). However, it has been suggested that the "fact that Article XX(e) is the only provision explicitly addressing production methods strongly indicates the other Article XX sections were not intended to include measures based on other production methods" (Feddersen, 1998, p. 76).

The conclusion that Article XX generally implies a separation regime between trade measures and labor rights issues seems to be in line with the official position of the WTO, as reflected in the declaration that emerged from the 1996 Singapore ministerial meeting. In general, the WTO agrees that certain core labor rights should be globally recognized and protected; however, its fundamental legal mandate is to regulate trade, and international protection of labor rights should be primarily the task of the ILO.

4.3 Open Rules

As discussed in Section 3, the third form of linkage regime – open rules – simply amounts to the absence of conditionality or separation rules, leaving countries *free* to enter into selective and separate agreements with different partners along different policy dimensions, and to use the threat of non-cooperation in one area to support cooperation in another.

This regime can either be implicit, when an international agreement contains no conditionality or separation rules, or it can be made explicit if an agreement formally

recognizes the possibility of cross-issue linkage. In the following, we provide two examples of the latter. We also briefly discuss an area of international law where the application of an open rules regime is still debated – namely the legal interface between trade and environmental agreements.

4.3.1 Trade and Security Policy

Article XXI of GATT explicitly allows trade measures to be imposed whenever a government considers this "necessary for the protection of its essential security interests," either in time of war or in the case of "other emergency in international relations." Sometimes the attempt to invoke a national security rationale is blatantly unjustified. For example, in 1975 Sweden argued that its import quota for footwear was motivated by national security concerns, because the survival of its domestic footwear was a guarantee against possible shortages of army boots in time of war.

GATT rules have been seldom used to contest economic sanctions imposed for foreign policy reasons. One exception is provided by the US embargo against Nicaragua in 1985, which was contested by the Nicaraguan government.[6] More recently, the introduction of the US Cuban Liberty and Democratic Solidarity (LIBERTAD) Act of 1996 – more commonly known as the Helms–Burton Act – introduced restrictions on goods of Cuban origin, as well as the possible refusal of entry visas and work permits for non-US nationals who have business interests in Cuba. This led the European Union to request the WTO to establish a Special Panel, which was approved on November 20, 1996; the panel suspended its work in April 1997, after the United States and Europe reached a bilateral resolution.

4.3.2 Trade and Monetary Policy

When GATT was drafted in 1947, balance-of-payments (BOP) problems were severe both for developing countries and for countries faced with postwar reconstruction. One of the purposes of the IBRD (the World Bank) and IMF was to assist in the postwar reconstruction of devastated economies.

Articles XII and XIV of GATT explicitly authorize certain trade restrictions when a need related to BOP conditions can be demonstrated.[7] A number of legal issues have arisen concerning the interaction between GATT rules on BOP-related trade restrictions and IMF rules (see Trebilcock and Howse, 2001).

4.3.3 Trade and Environment

Environmentalists often view trade restrictions as appropriate sanctions for noncompliance with international environmental standards; as a means of imposing such standards when there are none; or as a response to the failure of particular nations to engage in negotiations to develop such standards. On one hand, the protection of the global commons is clearly a legitimate concern that should be addressed. On the other hand, it has often been argued that the measures taken to elicit environmental cooperation might undermine trade cooperation.

There have been proposals to review Article XX of GATT in the sense of *allowing* the use of trade sanctions in defense of MEAs. For example, in his speech at the WTO High Symposium on Trade and Environment held in Geneva from March 15 to 16, 1999,[8] Sir Leon Brittan (then vice president of the European Commission) stressed the need to "be more sure than at present that WTO trade rules accommodate the aims of the parties to MEAs and therefore allow the necessary trade measures to be taken under such a MEA. If to achieve that confidence, we need a new interpretation of, or even a textual amendment to, WTO rules, I believe we should go down that route."

It should be noted that trade measures taken against signatories of MEAs pose no real legal issue for GATT, as long as the environmental agreement has come into force subsequent to the relevant GATT provisions. This is due to a general principle of public international law, whereby in the event of a conflict between international treaties, the provisions of the later treaty apply (Article 30(4b) of the Vienna Convention on the Law of Treaties). However, with respect to trade measures taken against nonsignatories of the environmental agreement who are members of GATT/WTO, the situation is unclear and is the subject of debate. Some environmental agreements explicitly say that countries that have signed the agreement should apply the rules of the agreement even to goods and services originating from countries that have not.[9] Whether this is in violation of the WTO agreement remains untested, because so far no dispute concerning the application of such provisions has been brought to the WTO.

5 Additional Considerations

The foregoing discussion has ignored a number of factors and problems that affect real-world multilateral negotiations. In this section we briefly comment on some of these additional aspects of international cooperation, speculating on their possible implications for the comparative effects of different legal cross-issue linkage regimes.

5.1 Emergence of New Issues

Our discussion so far has neglected the truly dynamic nature of multidimensional international cooperation (as distinct from the idea of repeated interaction on a single issue): changes occurring in the socioeconomic environment – technological innovation, institutional changes, demographic change – have a profound impact on international cooperation incentives. An example is provided by the relatively new international cooperation challenges arising from global climate change, a problem that has come to the fore only in the last three decades and which has deeply affected international negotiations in other areas. A more recent example is the "turnaround in US foreign policy," following the terrorist attacks of September 11. "The US, by necessity, is suddenly involved in the kind of furious diplomacy it abhorred and thought tedious as recently as two months ago. [...] In his first eight, balmy, months in office, Bush was happily tearing up the Kyoto Protocol, the 1972 Antiballistic Missile Treaty and Biological Weapons Convention, and so on. [...] Now he is constantly in talks with the leaders of other countries."[10]

The continuous emergence of new and unforeseen issues significantly alters the picture of multilateral negotiations we had painted earlier. When a new issue arises after agreement has been reached with respect to another, an exchange of concessions across the new and the old issues is only possible if the original agreement is renegotiated. Then, if agreements are binding, the previous agreement would come to represent the "fall-back" position for renegotiation, and thus affect the outcome of renegotiation. This means that countries, when entering into cooperation arrangements, also try to anticipate how current agreements might affect the outcome of any future agreements. Even if we think of international agreements as being not binding but self-supporting (i.e., continuously renegotiated), the *status quo* of cooperation may still be relevant. Theories of cooperation under repeated interaction predict that many different outcomes are possible; the *status quo* could then represent a "focal point" determining which particular outcome prevails.

Different legal linkage regimes may differently affect countries' strategic incentives to position themselves within current agreements so as to affect future negotiations in their favor. For example, one could conjecture that a blanket conditionality rule, forcing countries to cooperate simultaneously over new and old issues, may reduce incentives for entering into a preferential agreement on a certain issue in order to secure a "privileged" platform for negotiating an agreement over a new issue. Separation could have a similar effect, because it would remove countries' ability to threaten changes in existing cooperation arrangements with respect to old issues when negotiating agreements over new issues. More selective conditionality and separation rules, affecting certain dimensions of cooperation and not others, can be more difficult to impose when issues arise sequentially, simply because new issues are unforeseen and therefore cannot be selectively addressed by the provisions of earlier agreements.

5.2 Transaction Costs

Our analysis of multilateral negotiations in Section 3 did not account for the existence of transaction costs in international negotiations, costs that become more significant the broader the negotiation. For example, since the creation of GATT in 1947, trade negotiations have become increasingly complex. In the early GATT rounds, where the focus was mainly tariff cuts, and which involved representatives of a few nations only (23 in the first round), negotiations were concluded in a few months. The increase in the number of GATT member countries and in the number of issues to be negotiated upon has led to an increase in the complexity, and hence the duration, of more recent trade rounds. The Uruguay Round involved 117 participants, and its final document, comprising more than 26,000 pages, consisted of no less than 19 Agreements, 24 Decisions, 8 Understandings, and 3 Declarations.[11] The complexity of the negotiations explains why the Uruguay Round lasted from 1986 until 1994. Disagreement over the issues to be included in the negotiations has led to the collapse of the Seattle Round of GATT/WTO trade negotiations; disagreements with respect to the negotiation agenda have also started to emerge among the 142 countries participating in the new round of trade negotiations recently launched in Doha, Qatar.

Due to transaction costs, more limited forms of cooperation may actually be preferable to broader agreements on grounds of economic efficiency. This means that an inclusive, all-encompassing multilateral cooperation arrangement is not necessarily an ideal model of cooperation; which, in turn, changes how the effects of different linkage rules should be assessed.

5.3 Self-Enforcing Agreements and Myopic Policymakers

As mentioned earlier, one of the questions that has been stressed in the theoretical literature on international agreements is the problem of their enforcement. In the absence of external enforcement mechanisms, international agreements must be sustainable by the implicit threat of reciprocal punishments between the participating parties. This interpretation typically assumes domestic policymakers to be fully forward-looking. In reality, international policy interaction takes place between elected politicians who remain in office for a limited term. Theoretically, if these "finitely lived" policymakers behave as forward-looking welfare maximizers, political turnover should not affect international cooperation prospects: sharing a common objective, successive cohorts of policymakers would behave as a single infinitely lived player. However, if different policymakers have different policy stances – because of intrinsic differences in their policy preferences or because of political pressure or lobbying – then such equivalence breaks down, and the presence of limited terms of office may not be inconsequential for whether and how international policy cooperation takes place.

Specifically, political turnover implies that policymakers in different countries initially have only limited information about each other's incentives to cooperate, and must learn about each other through interaction. Infinitely lived policymakers would be able to fully infer any relevant information about their opponents from their policy choices within a finite period of time – after which they would interact with each other as fully informed opponents. Finitely lived policymakers, on the other hand, have only limited opportunities to learn about each other, and the need to learn arises continuously as new policymakers come on the scene. When multiple issues are involved, such learning may involve cooperation being established or renewed sequentially across different issues, with policymakers first cooperating in one area to "test each other" with respect to cooperation incentives in other areas. In this context, a conditionality or separation regime would constrain how such learning can take place, and, would, through this channel, affect cooperation prospects.

6 Summary and Conclusion

In this chapter, we have examined how legal regimes for the interface between agreements in different areas can affect the prospects to achieve multilateral cooperation. To compare the implications of different legal linkage regimes – conditionality,

separation, and open rules – we have focused on a scenario where countries can enter into selective and separate binding trade and environmental agreements with different partners. Our analysis suggests that conditionality is more likely to facilitate multilateral, multi-issue cooperation in situations where the environmental policy stakes are small relative to the welfare effects of trade policies; when the costs of environmental compliance are high, a conditionality rule can hinder multilateral cooperation. Separation can undermine cooperation by limiting punishment, but can also promote broad cooperation by making partial cooperation more difficult to sustain. Thus, how different linkage regimes affect multilateral negotiations depends on the structure of cooperation incentives for the countries involved.

We have provided some examples of how these linkage rules are translated into the letter of international law, and have touched upon certain real-world complications that affect the process of international cooperation, speculating on how they may affect negotiations. It should be stressed that what we have offered here are only conjectures: the literature on cross-issue cooperation in a multilateral environment is still in its infancy, and no theoretical analyses of these issues as yet exist.

Finally, it must be remarked that conditionality and separation rules could themselves be viewed as endogenous to the multilateral process, and hence require multilateral support and enforcement. We have noted that such rules are actually embodied in certain international laws and institutions, but our analysis has said nothing about their genesis. Further theoretical research is needed before this question can be addressed.

Acknowledgments

We gratefully acknowledge support from the University of Warwick and the ESRC.

Notes

1 For example, Carraro and Siniscalco (1994) point out that environmental free-riding incentives could be offset by making the signing of agreements entailing positive excludable externalities restricted to signatory countries (e.g., trade or R&D agreements) *conditional* on environmental cooperation.

2 Before the Uruguay Round, some agreements related to GATT 1947 had been negotiated during previous negotiating rounds. In particular, a number of agreements were negotiated during the Tokyo Round, in the area of non-tariff barriers. These agreements, however, were not adopted by all the contracting parties to GATT, and were only applied to those countries who accepted to be bound by them.

3 Four specific agreements are binding only on those Members who have accepted them. These agreements, referred to as the "purgatorial trade agreements," are the Agreement on Trade in Civil Aircraft, the Agreement on Government Procurement, the International Dairy Agreement and the International Bovine Meat Agreement (the latter two were terminated at the end of 1997).

4 The Bush administration has still not been able to win the "fast-track" trade negotiating authority from Congress due to the fact that the present bill "falls short in safeguarding labor rights and environmental standards" (*Financial Times*, November 7, 2001).

5 For a discussion, see Srinivasan (1998).
6 The Panel ruled that the embargo did not violate GATT, and noted that it was precluded from judging the validity of – or the motivation for – the US action (Jackson, 1997).
7 Article XII can be invoked by all Members, while Article XVIII(b) can be invoked only by the developing countries (defined as those in the early stages of development and with a low standard of living). The basic condition for invoking Article XII is to "safeguard the [Member's] external financial position and its balance-of-payments"; Article XVIII(b) mentions the need to "safeguard the [Member's] external financial position and ensure a level of reserves adequate for the implementation of its program of economic development." Both Articles refer to the need to "restore equilibrium on a sound and lasting basis." While Article XII mentions the objective of "avoiding the uneconomic employment of resources," Article XVIII(b) refers to "assuring an economic employment of production resources." Article XVIII(b) contains somewhat less stringent criteria than Article XII. Article XII states that import restrictions "shall not exceed those necessary (i) to forestall the imminent threat of, or to stop, a serious decline in its monetary reserves" or (ii) ". . . in the case of a contracting party with very low monetary reserves, to achieve a reasonable rate of increase in its reserves." Article XVIII(b) omits the word "imminent" from the first condition, and refers to an "adequate" level rather than a "very low" level of reserves. Both articles require Members to progressively relax the restrictions as conditions improve and eliminate them when conditions no longer justify such maintenance.

   The Uruguay Round Understanding of the Balance-of-Payments Provisions of the GATT states that contracting parties imposing restrictions for balance-of-payments purposes should do so in the least trade-disruptive manner and should favor price-based measures, such as import surcharges and import deposits, to quantitative restrictions. It also describes procedures for consultation by the GATT Balance-of-Payments Committee, as well as for notification of BOP measures.
8 See speech 99/47 on the WTO website.
9 This, however, would seem to be in contrast with Article 30(4b) of the Vienna Convention (see Hudec, 1996).
10 *New Statesman*, November 12, 2001.
11 Decisions were reached in four main areas: tariff reductions in agriculture, textiles, and apparel; institutional arrangements covering dispute settlement and the creation of the WTO; trade rules governing subsidies and countervailing duties, safeguards (VERs), and antidumping; and the new areas of services, investment, and intellectual property rights.

References

Abrego, Lisandro E., Carlo Perroni, John Whalley, and Randall M. Wigle 2001: Trade and environment: bargaining outcomes from linked negotiations. *Review of International Economics*, 9, 414–28.

Barrett, Scott 1994: Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46, 878–94.

Bond, Eric and Constantinos Syropoulos 1996: Trading blocs and the sustainability of interregional cooperation. In Matthew B. Canzoneri, Wilfred J. Ethier, and Vittorio Grilli (eds.), *The New Transatlantic Economy*, Cambridge University Press.

Carraro, Carlo and Domenico Siniscalco 1993: Strategies for the protection of the environment. *Journal of Public Economics*, 52, 309–28.

Carraro, Carlo and Domenico Siniscalco 1994: Policy coordination for sustainability: commitments, transfers, and linked negotiations. In Ian Goldin and Alan Winters (eds.), *The Economics of Sustainable Development*, Cambridge: Cambridge University Press.

Cesar, Herman and Aart de Zeeuw 1996: Issue linkage in global environmental problems. In Anastasios Xepapadeas (ed.), *Economic Policy for the Environment and Natural Resources: Techniques for the Management and Control of Pollution*, Cheltenham, U.K. and Lyme, N.H: Elgar.

Chander, Parkash and Henry Tulkens 1992: Theoretical foundations of negotiations and cost sharing in transfrontier pollution problems. *European Economic Review*, 36, 388–98.

Conconi, Paola and Carlo Perroni 2002: Issue linkage and issue tie-in in multilateral negotiations. *Journal of International Economics*, 57, 423–47.

Copeland, Brian R. 1990: Strategic interaction among nations: negotiable and non-negotiable trade barriers. *Canadian Journal of Economics*, 23, 84–108.

Ederington, Josh 2001: International coordination of trade and domestic policies. *American Economic Review*, 91, 1580–93.

Feddersen, Christoph T. 1998: Focusing on substantive law in international economic relations: the public morals of GATT's Article XX(a) and 'conventional' rules of interpretation. *Minnesota Journal of Global Trade*, 7, 75–122.

Folmer Henk, Pierre van Mouche, and Shannon Ragland 1993: Interconnected games and international environmental problems. *Environmental and Resource Economic*, 3, 313–35.

Hopkins, Raymond F. 2000: The political economy of foreign aid. In Finn Tarp (ed.), *Foreign Aid and Development: Lessons Learnt and Directions for the Future.* London and New York: Routledge.

Hudec, Robert 1996: GATT legal restraints in the use of trade measures against foreign environmental practices. In Jagdish Bhagwati and Robert Hudec (eds.), *Fair Trade and Harmonization: Pre-requisites for Free Trade?*, vol. 2. Cambridge, MA: MIT Press.

Jackson, John H. 1997: *The World Trading System: Law and Policy of International Economic Relations*. Cambridge, MA: MIT Press.

Krugman, P. 1991: Is bilateralism bad? In Elhanan Helpman and Assaf Rasin (eds.), *International Trade and Trade Policy*. Cambridge, MA, MIT Press.

Limão, Nuno 2001: Trade policy, cross-border externalities, and lobbies: do linked agreements enforce more cooperative outcomes? University of Maryland Center for International Economics Working Paper 02-01.

Maggi, Giovanni 1999: The role of multilateral institutions in international trade cooperation. *American Economic Review*, 89, 190–214.

Perroni, Carlo and John Whalley 2000: The new regionalism: trade liberalization or insurance? *Canadian Journal of Economics*, 33, 1–24.

Raiffa, Howard 1982: *The Art and Science of Negotiation*, Cambridge, MA: Harvard University Press.

Riezman, Raymond 1985: Customs unions and the core. *Journal of International Economics*, 19, 355–65.

Sebenius, James K. 1983: Negotiation arithmetic: adding and subtracting issues and parties. *International Organization*, 37, 281–316.

Spagnolo, Giancarlo 1996: Issue linkage, delegation, and international policy cooperation. *Fondazione Eni Enrico Mattei Working Paper* 49.96.

Srinivasan, T. N. 1998: Trade and human rights. In *Constituent Interests and U.S. Trade Policies*. Ann Arbor: University of Michigan Press, 225–53.

Trebilcock, Michael J. and Robert Howse 2001: *The Regulation of International Trade*. London and New York: Routledge.

Yi, Sang-Seung 1996: Endogenous formation of customs unions under imperfect competition: open regionalism is good. *Journal of International Economics*, 41, 151–75.

11

Antitrust Policy in Open Economies: Price Fixing and International Cartels

Eric W. Bond

Chapter Outline

This chapter surveys the literature on the impact of trade protection on the level of collusion between home and foreign firms, including the role of administered protection and multi-market collusion. It also examines conditions under which national competition laws will result in an efficient international outcome, and discusses the degree to which observed price-fixing rules are consistent with national welfare-maximization.

1 Introduction

This chapter examines the economics of international cartels and the incentive of countries in the setting of policy toward cartels and price-fixing arrangements. This issue is of interest because of evidence indicating that there continues to be substantial international cartel activity, despite increased enforcement by antitrust authorities worldwide.[1] A report by the World Trade Organization (WTO, 1997, Chapter 4) concluded, "… while the extent of cartel activities is intrinsically difficult to assess … there are some indications that a growing proportion of cartel agreements are international in scope." Prosecutions by the US Department of Justice (DOJ) in the 1990s also support the view that international cartels have been operating successfully in a number of international markets, and have maintained market sharing arrangements across countries

for prolonged periods of time. The DOJ (2000, Chapter 4) cites successful prosecutions of a number of cartels involving both US and foreign firms, including a food and feed additive cartel (operating from 1991 to 1995), a vitamin cartel (1990–1999), and a graphite electrodes cartel (1992–1997). These cartels had significant impacts on prices in the US market, with the feed additive cartel having been estimated to raise the price of citric acid by 30 percent and the price of lycine by 70 percent. The magnitude of these cartels is reflected in the fact that since 1997, international cartels have accounted for more than 90 percent of the fines imposed in criminal antitrust cases annually.

This chapter addresses three questions regarding cartel activity. The first is how the recent process of globalization affects international cartels. The last several decades have witnessed a steady reduction in trade barriers, as well as decreases in transport and communication costs. The objective of Section 2 of this chapter is to examine how these changes affect cartel behavior, and whether they are likely to increase the level of cartel activity. This question is addressed using a simple Cournot oligopoly model, which is used to characterize cartel agreements between firms that must be self-enforcing. This section emphasizes how models of tacit collusion can be used to understand the features of recent cartel agreements, and how trade barriers and the antidumping process affect the ability to sustain collusive agreements. In addition, it is shown that the patterns of collusive agreements and the response to trade liberalization may differ significantly in cases where collusion is organized on a worldwide basis, with firms forming joint agreements covering all markets.

The second question to be addressed is whether there is a potential role for international coordination regarding cartel policy. As trade barriers have fallen worldwide, the WTO has begun to pay more attention to how market rules affect market access and trade flows. Competition policy, which defines the rules of acceptable competition practices within a market, is one topic that is being considered by the WTO for future negotiations.[2] Section 3 examines the optimal policies for a country that is maximizing welfare, and examines the conditions under which the spillovers created by cartel policy give rise to a case for international agreements on cartel policy. The relationship among tariff policy, cartel policy, and other types of competition policy are also examined.

The final question concerns the interactions between political influence and the making of cartel policy. Although it is well understood that virtually all price-fixing agreements run counter to the public interest and efficient operation of markets, there is a wide variation across countries in the aggressiveness with which cartels have been prosecuted. The US DOJ has historically pursued price-fixing agreements with considerable zeal, while the United Kingdom followed a laissez-faire attitude to cartels in the interwar period, and oil-exporting countries participate in a cartel organized by the national governments. Economic theory can provide some insights on how these differences across countries might reflect differences in the underlying comparative advantage of the countries or differences in the political power of interest groups across countries. In addition, it can provide insights on how current international rules concerning the application of national competition policy can affect the nature of competition on a worldwide basis. Section 3 compares the predictions of a model that examines the role of political factors in determining price-fixing policy. Section 4 provides a brief comparison and contrast of the evolution of price-fixing policy in the United States and

the United Kingdom, in order to illustrate the role of political factors in shaping cartel policy.

2 Cartels and Price-Fixing Arrangements Among Firms

We begin by considering the economics of international price-fixing arrangements among firms. The analysis centers on a simple two-country model of a homogeneous product Cournot oligopoly, which is chosen because of its simplicity and its frequent use to analyze cartel problems. We first characterize the set of feasible cartel equilibria for collusion between home firms and foreign firms, assuming that agreements are enforceable and that firms are colluding on a market-by-market basis. We then examine how the feasible set will be affected if the agreements must be self-enforcing, and how the set of self-enforcing agreements is affected by trade policy. The analysis concludes with a discussion of how the results may differ if firms are colluding on a worldwide basis.

2.1 Cartels in Cournot Oligopoly

To examine the cartel problem, we utilize a two-country model of an industry that can be characterized as a Cournot oligopoly. The two countries will be referred to as home and foreign, with home (foreign) country variables being denoted by an h(f) superscript. The oligopoly is characterized by an exogenously given number of firms, n^k, in country $k \in \{h, f\}$, which implies the existence of barriers that prevent new entry in response to the formation of a cartel. Demand is linear and identical in each country, with the price received by firm i in market k given by $p_{ik} = A - \sum_{j=1}^{n^{\mathrm{h}}+n^{\mathrm{f}}} q_{jk}$, where q_{jk} is the quantity sold by firm j in market k. The cost function is assumed to have constant marginal cost that is identical for all firms, domestic and foreign, which we can normalize to zero for simplicity. The asymmetry between firms in market k is that firms from country $j \neq k$, must pay a trade cost of t^k for each unit sold in country k's market. The trade cost could include tariffs imposed by the government of k as well as transport costs between the countries. The assumption of identical marginal costs and symmetry in demand means that all firms from country j will face the same marginal cost and demand conditions in country k, so that we can conduct the analysis in terms of a representative home firm and a representative foreign firm in each market. We assume markets are segmented, with a firm from country j choosing a quantity q^j of sales for the domestic market and x^j of exports.

In this section we focus on the equilibrium when firms are colluding in a single market, which we will take to be the home market. The case in which $t^{\mathrm{h}} = 0$, so that home and foreign firms compete on an equal footing, is illustrated in Figure 11.1 for $n^{\mathrm{h}} = n^{\mathrm{f}} = 2$. When there is no collusion, each firm will choose its profit-maximizing output, taking the output of all other (home and foreign) firms as given. The R^{h} curve, which is derived in the Appendix, shows the output level that will be chosen by a representative home firm when the representative foreign firm exports x^{f}.[3] This curve

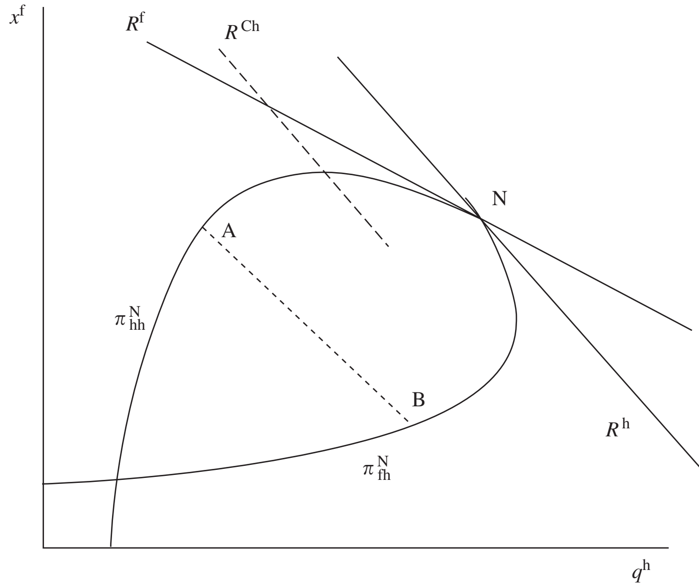

Figure 11.1 Profitable cartel outputs for $n^{\mathrm{h}} = n^{\mathrm{f}} = 2$

slopes downward because an increase in the output of foreign firms will reduce the demand for home products, leading to a lower output by home firms. Similarly, the R^{f} function shows the profit-maximizing export sales of the foreign firms, given that the representative home firm produces q^{h}. The Nash equilibrium output level in the home market, $(q^{\mathrm{hN}}, x^{\mathrm{fN}})$, is determined by the intersection of these two curves. With $t = 0$, the output levels are $x^{\mathrm{fN}}(0) = q^{\mathrm{hN}}(0) = A/(n^{\mathrm{h}} + n^{\mathrm{f}} + 1)$, which yields a profit to each firm of $(A/(n^{\mathrm{h}} + n^{\mathrm{f}} + 1))^2$.

With these demand and cost functions, the profit of a representative home in the home market is $\pi_{\mathrm{hh}} = (A - n^{\mathrm{h}}q^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}})q^{\mathrm{h}}$, and the profit of a foreign firm in the home market is $\pi_{\mathrm{fh}} = (A - n^{\mathrm{h}}q^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}} - t^{\mathrm{h}})x^{\mathrm{f}}$. The $\pi^{\mathrm{N}}_{\mathrm{hh}}$ curve in Figure 11.1 is the locus of values of q^{h} and x^{f} that yield a home firm the same profit as the Nash equilibrium. Profits of home firms will be decreasing in the output of foreign firms, so that all output combinations lying below the $\pi^{\mathrm{N}}_{\mathrm{hh}}$ locus will yield higher profits than the Nash equilibrium level. Since there are two home firms in the case illustrated in Figure 11.1, R^{h} will intersect the home iso-profit locus at an output level that exceeds the one that maximizes the profits of the home firms (i.e., the point at which the iso-profit locus is horizontal). A similar argument can be used to establish that the profits of the foreign firm will be decreasing in the output level of home firms, with profits of foreign firms exceeding the Nash equilibrium level for all output combinations to the left of the $\pi^{\mathrm{N}}_{\mathrm{fh}}$ locus. Therefore, output levels in the region lying below the $\pi^{\mathrm{N}}_{\mathrm{hh}}$ locus and to the left of the $\pi^{\mathrm{N}}_{\mathrm{fh}}$ locus will raise the profits of both types of firms.

Figure 11.1 can also be used to illustrate the importance of including foreign firms in a collusive agreement, because a cartel consisting of only domestic firms may be unprofitable. Suppose that the domestic firms were to organize and choose output levels

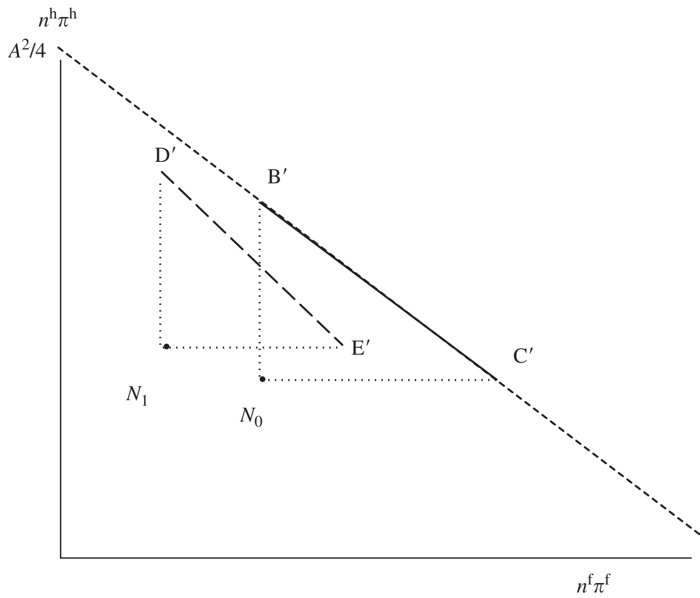

Figure 11.2 Efficient frontiers at $t^h = 0$ (B′C′) and $t^h > 0$ (D′E′)

to maximize the total profits of all domestic firms. The profit-maximizing choice of output for domestic firms would be illustrated by the dotted line R^{Ch} in Figure 11.1, which maximizes the profits of home firms for a given x^f. R^f is unaffected by this collusion, so the market equilibrium with home firm collusion is given by the intersection of the R^{Ch} and R^f curves. In the case illustrated in Figure 11.1, a cartel consisting of home firms alone would be unprofitable because it would result in an equilibrium profit level less than π^N_{hh}. When the home firms contract their output level, the market price will rise and the foreign firm responds with a larger output level. In this case, the market share of home firms falls sufficiently when they collude that the cartel is unprofitable. Whether a cartel of home firms alone is profitable will depend on the importance of home firms in the market. In the case where products are identical, it can be shown that a cartel of home firms will be profitable only if $n^h \geq (1 + n^f)^2$, which requires that at least 80 percent of the firms be home firms.[4]

Figure 11.2 illustrates the payoffs that can be attained by home and foreign firms by choosing agreements contained in the lens-shaped area between the π^N_{hh} and π^N_{fh} curves in Figure 11.1. With free trade, these payoffs will be illustrated by the triangle B′C′N$_0$. To see this, consider the outputs along the BC line in Figure 11.2. These outputs are combinations of q^h and x^f such that the industry output equals the monopoly output level of $A/2$. These output levels are efficient in that it is not possible to find outputs that yield higher profits for both types of firms, which is reflected by the fact that the iso-profit contours of the two firm types are tangential along this line. The outputs on BC in Figure 11.1 generate profits on line B′C′ in Figure 11.2, which satisfies $n^h\pi_{hh} + n^f\pi_{fh} = \pi^M$, where $\pi^M = A^2/4$ is the monopoly profit level. The point

N_0 in Figure 11.2 illustrates the payoffs to each firm in the Nash equilibrium, so that payoffs in the triangle $B'C'N_0$ are feasible profit levels that leave firms in each country at least as well off as in the Nash equilibrium. If firms are operating in a permissive legal environment where agreements to restrict output are not illegal and can be enforced by cartel members, then the firms would bargain over the payoffs on the efficient frontier $B'C'$. In the case where home and foreign firms have symmetric bargaining power, the outcome would be an equal allocation of an output of $A/(2(n^h + n^f))$ to each firm.

Figures 11.1 and 11.2 can also be used to illustrate how the efficient payoffs would be affected by the presence of trade barriers. If $t^h > 0$, foreign firms will have higher costs of serving the home country market and will produce a lower level of output, resulting in a leftward shift in R^f in Figure 11.1. The new Nash equilibrium will have a lower output and profit for foreign firms, and a higher output and profit for home firms. This is illustrated by the shift of the Nash equilibrium profit levels from N_0 to N_1 in Figure 11.2. Since foreign firms have higher costs when there is a tariff, the cost to the cartel of serving the home market will be higher the greater the market share allocated to the foreign firms. This is illustrated by the inward rotation of the efficient profit frontier to $D'E'$ in Figure 11.2. For $t^h \geq A/2$, production by foreign firms is unprofitable at the monopoly price and the efficient frontier becomes a single point with $n^h\pi_{hh} = A^2/4$ and $n^f\pi_{fh} = 0$.

The tariff has two impacts on the bargaining problem: it reduces the overall profit potential of the cartel and shifts the threat point for the bargaining problem in favor of the home firm. Clearly, the most profitable outcome for home firms would be $t^h \geq A/2$, so that home country firms could divide the monopoly profits among themselves. In light of the low overall level of tariffs that currently exist on manufactured goods, tariffs sufficiently high to allow the home firms to exclude foreign firms completely from a cartel are unlikely to be observed. However, lower levels of tariffs may still be beneficial to the home firms by shifting the bargaining power in favor of home country firms. The importance of tariffs in affecting the bargaining position of firms in a cartel is discussed by Hexner (1946, Chapter 7), who cites a number of cases from the interwar period in which tariffs were used by countries to increase their bargaining power in international cartels.

2.2 Self-enforcing Cartel Agreements

The discussion so far has assumed that a cartel can enforce an output restricting agreement among its members. However, cartel members will have an incentive to deviate from the outputs assigned because the assigned outputs are not profit maximizing given the outputs of other firms. The use of contracts to enforce cartel agreements is typically not an option because agreements to fix prices are illegal in most countries. Even in cases where cartels operate with explicit government approval, the agreements are typically not enforceable in courts.[5] Therefore, the cartel agreements will only be sustainable if there is a credible means of punishing a firm that deviates from the agreement.

If firms are interacting repeatedly in a market, a firm can be prevented from cheating on an agreement by the threat of a price war in the event that any firm cheats. For this

threat to be credible, it must be in the interest of the firms to actually carry out the punishments in the event of a deviation. A simple and credible threat that is utilized by Friedman (1971) is that if any firm deviates, the cartel members will revert permanently to the Nash equilibrium output levels. A cartel agreement that specifies home market outputs of $(q^{\mathrm{hC}}, x^{\mathrm{fC}})$ will be sustainable under this punishment scheme if it yields a higher payoff than can be earned by cheating and then suffering the punishment. Let $\pi_{i\mathrm{h}}^{\mathrm{D}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^i)$ be the payoff earned by firm i with trade cost t^i from deviating from the agreement and choosing the profit-maximizing output level, and let δ be the discount rate of the firm on the next period's payoffs. The average return earned by a deviating firm in the home market is $(1-\delta)\pi_{i\mathrm{h}}^{\mathrm{D}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^i) + \delta\pi_{i\mathrm{h}}^{\mathrm{N}}$. The deviation payoff of firm i will be decreasing in the output of all other firms, since the potential profits of a deviator grow the farther the output falls below the Nash equilibrium level. The agreement will be sustainable if this deviation payoff is no greater than the payoff under the agreement, which requires

$$\pi_{i\mathrm{h}}^{\mathrm{C}}(q^{\mathrm{hC}}, x^{\mathrm{hC}}, t^{\mathrm{h}}) \geq (1-\delta)\pi_{i\mathrm{h}}^{\mathrm{D}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}) + \delta\pi_{i\mathrm{h}}^{\mathrm{N}}(t^{\mathrm{h}}) \quad \text{for } i \in \{\mathrm{h, f}\}. \tag{1}$$

Since the deviation payoff is the profit-maximizing response to the outputs of all other firms, $\pi_{i\mathrm{h}}^{\mathrm{D}} > \pi_{i\mathrm{h}}^{\mathrm{C}}$ and $\pi_{i\mathrm{h}}^{\mathrm{C}} > \pi_{i\mathrm{h}}^{\mathrm{N}}$ in any profitable cartel agreement. It then follows that a given agreement is more likely to be sustainable the higher the δ, because higher values of δ make deviation less attractive by putting greater weight on the punishment phase. We will refer to the value of δ for which (1) holds with equality as the *minimum discount factor* for a firm from country i associated with the agreement $(q^{\mathrm{hC}}, x^{\mathrm{hC}})$ at tariff t^{h}. Denoting this minimum factor as $\delta_i^{\min}(q^{\mathrm{h}}, x^{\mathrm{f}}, t^{\mathrm{h}})$, an agreement will be sustainable for any $\delta \geq \delta_i^{\min}(q^{\mathrm{h}}, x^{\mathrm{f}}, t^{\mathrm{h}})$. Since the time period is defined as the length of time it takes the other firms to detect that a deviation has taken place, the value of δ will depend on how quickly firms can detect cheating as well as on the market rate of interest. Thus, globalization can lead to increased collusion among firms if it makes it easier for firms to detect cheating by other firms.

The following result, which is proven in the Appendix, illustrates some properties of the minimum discount factor.

Proposition 1. The minimum discount factor is less than one for any agreement with $\pi_{i\mathrm{h}}^{\mathrm{C}} > \pi_{i\mathrm{h}}^{\mathrm{N}}$

(a) With $t^{\mathrm{h}} = 0$, an equal division of the monopoly output level among all firms is easier to sustain the smaller the $n^{\mathrm{h}} + n^{\mathrm{f}}$. The minimum discount factor for a symmetric agreement with $q^{\mathrm{h}} = x^{\mathrm{f}} = q$ is decreasing in q.

(b) An increase in the probability of detection or level of fines for firms caught colluding will make collusion more difficult to sustain.

Proposition 1 reflects the usual "folk theorem" result of repeated games, in that any profitable agreement can be supported if sufficient weight is put on the future. Part (a) shows that collusion is easier to support when there are fewer firms, because firms have

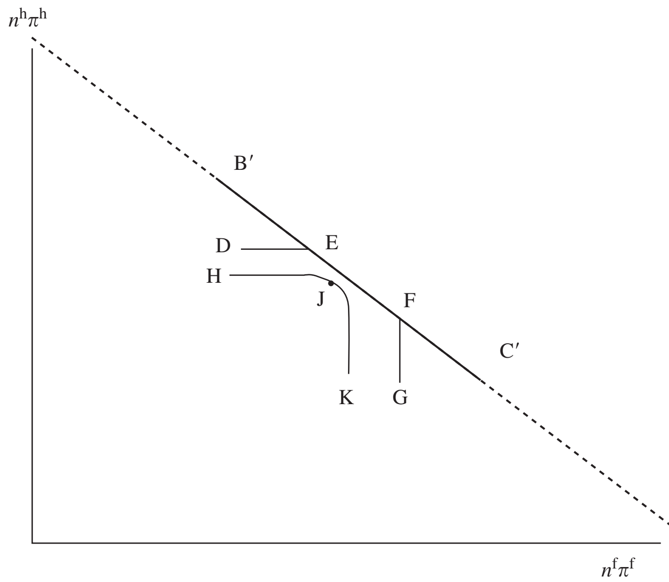

Figure 11.3 Sustainable profit levels with $\delta = 0.7$ (DEFG) and $\delta = 0.5$ (IJK)

a lower incentive to deviate when they have a larger market share. The latter half of (a) shows that if the monopoly output level is not sustainable, the no deviation constraint can be relaxed by increasing the output of each firm. For any $\delta > 0$, firms will be able to sustain some symmetric output $q < q^N(0)$ that yields higher profits than the Nash equilibrium. Part (b) shows that increased antitrust enforcement will make collusion harder to sustain, because it reduces the benefits from collusion.[6]

When agreements must be self-enforcing, cartel members must choose an agreement that satisfies (1). Figure 11.3 illustrates how the requirement that an agreement be self-enforcing shrinks the set of agreements that can be sustained as the discount parameter is reduced. The frontier B′C′ is the profit-maximizing frontier with $t^h = 0$ using the same parameter values as in Figure 11.2. Proposition 1 establishes that points on this frontier that yield payoffs to each firm type strictly greater than the Nash equilibrium can be sustained for δ sufficiently close to one. For a discount parameter equal to 0.7, the maximum payoffs that can be sustained are given by DEFG. Divisions of the monopoly output that are sufficiently close to symmetry (i.e., the interval EF) have minimum discount factors less than 0.7 and are sustainable. However, divisions of the monopoly output with highly asymmetric market shares (B′E and FC′) will be unsustainable because the firms receiving low market shares will have too large a deviation incentive. For a discount parameter equal to 0.5, the efficient frontier as given by HJK is sustainable. In this case the discount parameter is sufficiently low that no efficient agreement is sustainable. Point J is the highest symmetric payoff that is sustainable, and is the point that would be chosen if the bargaining power of home and foreign firms is symmetric. The set of sustainable agreements will shrink to the Nash point as $\delta \to 0$.

2.3 Trade Policy and the Sustainability of Agreements in Single Markets

Henry Havemeyer, president of the American Sugar Refining Company, is well known for his 1899 remark that "The mother of all trusts is the customs tariff bill." The results of the previous section imply that Havemeyer's assertion is correct in the case of a tariff sufficiently high to prevent entry of foreign firms at the monopoly price. With foreign firms excluded from the market, Proposition 1(a) shows that collusion will be easier to sustain in the home market. Thus, collusion will be easier to sustain with this prohibitive tariff than with free trade. However, the results in the literature suggest that for non-prohibitive tariffs, the relationship between trade collusion is subtler than suggested by Havemeyer's remark. In particular, results are sensitive to the level of the tariff, the form of protection, and whether firms compete in prices or quantities.

For the home firm, $\pi_{\text{hh}}^{\text{C}}$ and $\pi_{\text{hh}}^{\text{D}}$ are unaffected by increases in t^{h}, while $\pi_{\text{hh}}^{\text{N}}$ is increasing in t^{h}. Since this increases the right-hand side of (1), an increase in t^{h} must raise $\delta_{\text{h}}^{\min}(q^{\text{h}}, x^{\text{f}}, t^{\text{h}})$ because higher tariffs make the punishment less severe for a deviating home firm. For the foreign firms, an increase in t^{h} will reduce $\pi_{\text{fh}}^{\text{C}}$, $\pi_{\text{fh}}^{\text{D}}$ and $\pi_{\text{fh}}^{\text{N}}$, so the effect of an increase in t^{h} will depend on the relative magnitude of these effects. In particular, it can be shown that unambiguous results will not be attainable because the effect will depend on the particular agreement being considered. Davidson (1984) was the first to address this question. He used a Cournot model with $n^{\text{h}} = n^{\text{f}} = 1$ and assumed that the cartel was attempting to collude on the monopoly output level. He showed that there is a non-monotonic relationship between the tariff rate and the minimum discount factor in this case. An increase in the tariff will make a given agreement more difficult to sustain for the home country as noted above, so the output levels must be adjusted to compensate home firms. When the tariff rate is low, this compensation is possible and the minimum discount factor to sustain a monopoly output level will fall. However, when the tariff is high the foreign firms do not earn enough to compensate the home firms, and sustaining collusion will not be possible.

Fung (1992) analyzed the same question assuming that firms are Bertrand competitors. He found that trade liberalization makes collusion easier to sustain if the foreign firm is the low-cost producer, but could make it harder to sustain if the foreign firm is the high-cost producer. Rotemberg and Saloner (1989) compare the effects of tariffs and quotas in a model of Bertrand competition, and show that collusion is easier to sustain under a tariff than under a quota. This result is due to the fact that the quota makes it more difficult for the foreign firm to punish a deviation by the home firm, which makes the cartel harder to sustain.[7] These results indicate that policymakers should not automatically assume that trade liberalization will ensure that markets are competitive. Strict antitrust enforcement may be required along with trade liberalization, depending on the number of firms and type of competition.

A second way in which trade policy can affect collusion is that the protection process itself may have features that facilitate collusion between home and foreign firms. Prusa (1992) provides an example of this in his analysis of the antidumping process. Suppose that home firms file an antidumping petition against foreign firms, alleging that foreign

firms are selling at less than fair value in the home market. If the petition is upheld, it will result in the imposition of antidumping duties against the foreign firm. Prusa points out that one way for the foreign firms to avoid penalties is to reach an agreement with the home firms to raise prices, which would result in the withdrawal of the petition by the home firms. In the absence of an antidumping petition, such agreements between home and foreign firms to raise prices would be violations of antitrust laws. However, under the Noerr–Pennington doctrine these agreements are not subject to antitrust laws because they are part of an administrative process. Thus, the antidumping process may allow home and foreign firms to coordinate prices and market shares in a way that would otherwise be illegal. Prusa found that approximately 20 percent of the antidumping cases were withdrawn prior to a finding being reached.[8]

The antidumping process affects collusion in two ways. By eliminating the threat of antitrust enforcement, it will relax the no deviation constraint as indicated in Proposition 1(b) and shift the frontier of sustainable agreements outward. Second, it shifts the threat point in favor of the home firm because of the possibility of antidumping duties being imposed in the absence of an agreement (as in the shift from N_0 to N_1 rated in Figure 11.2). The home firms must be better off as the result of the antidumping process. This also raises the possibility that foreign firms could also benefit from the antidumping process if the outward shift in the frontier of sustainable agreements is large, relative to the shift in bargaining power.[9]

These results highlight the fact that the impact of trade policy on collusion will depend not just on the level of trade barriers, but also on the form of trade barriers. In particular, the rules of the protective process itself may facilitate collusion as in the case of the antidumping process.

2.4 Multi-market Collusion

The foregoing analysis has focused on collusion in the home market. Since it is assumed that the firms are also selling in the foreign market, the no deviation constraint (1) could also be applied to the home market to obtain the collusive agreements that would be sustainable in the foreign market. However, Bernheim and Whinston (1990) have pointed out that when firms compete in multiple markets, they will frequently be more profitable by linking the agreements across markets. For example, consider the symmetric case of $n^{\mathrm{h}} = n^{\mathrm{f}} = 2$ in Figure 11.2 and assume that both the home and foreign market are of equal sizes. If each country imposes a tariff of t, the cartel profit will be strictly below the monopoly profit level if both firms produce. However, the cartel can obtain the global monopoly profit if firms specialize geographically, with each firm type producing the monopoly output in their domestic market and selling nothing in the overseas market. This exchange of market shares across markets raises cartel profits by avoiding trade costs.

A second benefit of multi-market collusion is that it may make collusion easier to sustain by relaxing the no deviation constraint in (1). Let $\Pi_j^{\mathrm{C}} = \pi_{j\mathrm{h}}^{\mathrm{C}} + \pi_{j\mathrm{f}}^{\mathrm{C}}$ be the global profits of a firm of type j from a cartel agreement $(q^{\mathrm{h}}, x^{\mathrm{f}}, q^{\mathrm{f}}, x^{\mathrm{h}})$, with Π_j^{D} similarly defined to be the global profit from deviating and Π_j^{N} the global payoff on the Nash

equilibrium. The no deviation constraint will then require that

$$\Pi_i^{\mathrm{C}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, q^{\mathrm{fC}}, x^{\mathrm{hC}}, t) \geq (1-\delta)\Pi_i^{\mathrm{D}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, q^{\mathrm{fC}}, x^{\mathrm{hC}}, t) + \delta\Pi_i^{\mathrm{N}}(t), \quad \text{for } i = \mathrm{h}, \mathrm{f}. \tag{2}$$

It is clear that any cartel agreement that is sustainable with market by market collusion is sustainable with multi-market collusion, because an agreement satisfying (1) in the home and foreign markets for firms from country i will satisfy (2). However, suppose there is an agreement that satisfies (1) for firm i in the home market but not in the foreign market; this agreement could still satisfy (2), because it only requires that the sum of the individual market constraints be satisfied. With multi-market collusion, the firms can use extra enforcement power in one market to support the agreement in another market. Thus, the set of agreements sustainable with multi-market contact will be at least as large than can be satisfied with single market collusion in this case.[10]

In single market collusion, trade barriers affect the cartel members asymmetrically. With multi-market collusion, global changes in trade barriers will tend to have a more symmetric effect on firms because all firms act as import-competing firms in their home market but as exporters in foreign markets. The primary effect of trade barriers in this setting is to affect the profitability of geographic collusion in which each firm concentrates on its local market. Bond and Syropoulos (2002) show that this yields a non-monotonic relationship between the level of trade barriers and the profitability of collusion. At tariffs sufficiently high that firms are engaged in geographic collusion, trade liberalization will tend to make this geographic collusion more difficult to support. In this region, trade barriers make geographic collusion easier by making it less attractive for outside firms to deviate from the agreement. On the other hand, at low tariff rates firms will find it most attractive to choose agreements that assign all firms equal market shares in each market. This follows from the fact that deviation incentives are lowest when market shares are similar across markets. Low tariff rates make it easier to sustain agreements with symmetric market shares, because it reduces the tax on trade under symmetric agreements. This yields a U-shaped relationship in which profits initially decline with tariff protection, but eventually increase with tariffs as geographic collusion becomes sustainable. In addition, Bond and Syropoulos show that reciprocal trade liberalization may be welfare reducing in the neighborhood of free trade because it increases the sustainability of cartel agreements.

The evidence suggests that multi-market contact has been an important feature of international cartels, even from the earliest cases. For example, in US versus American Tobacco (1911) the DOJ prosecuted an agreement between US and British firms in which they agreed not to sell in each other's markets. Recent international cartels prosecuted by the United States, which are described in DOJ (2000, Chapter 4), show features of global division of markets. These agreements include a market-sharing agreement to fix prices and allocate market shares, an information-sharing agreement to verify market shares and prices, and a compensation scheme in case firms did not receive the allocated market share. The features of these cartels are consistent with the models of tacit collusion and multi-market contact presented in this section.

3 Optimal Antitrust Policy

We now turn to the optimal home country cartel and trade policies for this industry. We begin by examining the country's choice of policy toward cartels in the case where the country's antitrust instrument is the ability to prevent firms from engaging in collusion and the objective function is to maximize national welfare. This discussion highlights the role of the allocation of rights of countries to regulate price-fixing in international cartels in the absence of a supranational antitrust authority, because the interests of the importing and exporting countries will conflict. We then examine how international trade policy is affected by the enforcement of price-fixing policies, and conclude with an analysis of how results are changed if firms have more political power than consumer interests.

3.1 National Welfare and Cartel Policy

We first examine the case in which the objective of policymakers is to maximize national welfare, which is defined to be the sum of consumer surplus, producer profits, and tariff revenue. With the linear demand schedule assumed here, consumer surplus can be expressed as $S(Q) = (Q^i)^2/2$, where Q^i is the total supply of the good in country i. Assuming that all trade barriers take the form of tariff revenue, tariff revenue will be $t^{\mathrm{h}} n^{\mathrm{f}} x^{\mathrm{f}}$. Combining these terms with the profit functions derived previously yields an expression for welfare as a function of outputs in each market,

$$W^{\mathrm{h}}(q^{\mathrm{h}}, x^{\mathrm{f}}, q^{\mathrm{f}}, x^{\mathrm{h}}) = S(n^{\mathrm{h}} q^{\mathrm{h}} + n^{\mathrm{f}} x^{\mathrm{f}}) + n^{\mathrm{h}}(\pi_{\mathrm{hh}}(q^{\mathrm{h}}, x^{\mathrm{f}}) + \pi_{\mathrm{hf}}(q^{\mathrm{f}}, x^{\mathrm{h}})) + n^{\mathrm{f}} t^{\mathrm{h}} x^{\mathrm{f}}. \tag{3}$$

Equation (3) can be used to compare the national welfare at the Nash equilibrium with that at the cartel equilibria.

The following result is proven in the Appendix.

Proposition 2. National welfare in the Nash equilibrium in the home country market exceeds that in any profitable cartel between home and foreign firms in the home country market for all t^{h}. National welfare is higher in any cartel equilibrium in the foreign country market than in the Nash equilibrium at home.

This result implies that the optimal competition policy for the home country is one that prevents price-fixing of any type in the home market but allows any price-fixing arrangement in the foreign market. In the home country market, a reduction in output resulting from a cartel has three effects. First, it causes aggregate output in the home market to fall, which reduces the welfare of home country consumers. Second, there is an increase in the profits of home country firms, since any profitable cartel must give home firms a profit higher than in the Nash equilibrium. The final effect is that if $t^{\mathrm{h}} > 0$, a cartel will raise (lower) tariff revenue if it results in an increase (decrease)

in x^f. Although a cartel must lower overall industry output, it is clear from Figure 11.1 that there are some cartel outputs that result in higher outputs for foreign firms in the Nash equilibrium in cases where $n^h > 1$. The first part of the proposition establishes that the negative effects of a cartel on consumer surplus must dominate any positive effects on home firm profits or tariff revenues for *any* profitable cartel. This is quite a strong result, since it indicates that the optimal policy is to prevent cartels in its domestic market regardless of the level of tariffs.[11] Trade liberalization will not affect the home country's desire to pursue an aggressive policy against collusion in its home market. In contrast, the home country would be willing to allow its firms to participate in any cartel in export markets that raised the profits of its firms. Since all of the losses from the cartel in the foreign market fall on foreign consumers, cartels in the foreign market benefit the home firms as long as they raise the profits of home firms.

This yields a conflict between the exporting country and importing country over the efficient cartel policy if each government is attempting to maximize national welfare. If the home country has the right to control the behavior of any firm that is selling in the home country market, then price-fixing would not be allowed. For the world as a whole, the efficient policy will be to prevent cartels because the cartel arrangement creates a deadweight loss.[12] The losses to consumers from the cartel's output restriction exceed the gain in profits to the cartel members. Thus, we obtain the following corollary to proposition 2:

Corollary. If countries choose cartel policy to maximize national welfare in the model with constant marginal costs, then the assignment of the rights to the importing country to choose the market rules will result in an efficient international outcome.

This result suggests that a supranational authority to pursue cartels is unnecessary if the rights to set competition policy are allocated appropriately, because each government would have an incentive to prevent cartels in its own markets. In practice, international law assigns a limited right to regulate the actions of international cartels acting within its boundary. It is a well-established principle of international law that a country can impose civil or criminal sanctions on foreign conduct that causes substantial effects within its authority. The United States has used this principle, starting with US versus American Tobacco in 1911, to successfully pursue foreign firms that have engaged in price-fixing in the US market. On the other hand, a case against the Organization of Petroleum Exporting Countries for violation of US antitrust policy was dismissed because OPEC's actions were "acts of state" that were beyond the reach of the US judiciary. The courts reasoned that acts of state were issues that should be handled through international diplomacy by the executive and legislative branches of government. The ability of governments in importing countries to pursue cartels involving foreign firms thus depends on the extent of state sponsorship.[13]

It should be noted however that there are a number of caveats associated with the corollary. First, this assignment works in the case of seller cartels but not in the case of a cartel of buyers. If a small number of domestic producers are buying inputs from competitive foreign producers, the domestic producers have an incentive to collude in order to purchase the input at a lower price. Such buyer cartels would be inefficient from a world welfare perspective, but would be allowed to operate if the home country

set its market rules.[14] A second problem is that the prosecution of a worldwide cartel is a form of public good. Once the prosecution has been completed, other countries can free ride on the information collected by the first country to prosecute to collect damages for activities in their own market. This will tend to lead to underinvestment in the prosecution of international cartels. Both of these cases would argue for cooperation between governments and agreements that ensure information-sharing in international investigations of collusion. Finally, this discussion is based on the assumption that individual firms have constant marginal cost. With constant marginal costs, a country's decision about the level of competition to be allowed in its market has no impact on costs and outputs in other markets. If this assumption is relaxed, a country might have an incentive to allow cartels in its domestic markets if they had a favorable spillover to costs in other markets.

3.2 Tariff Policy and Multilateralism

Proposition 2 established that the optimal cartel policy is independent of the country's tariff policy. We now turn to an examination of the optimal tariff policy in this model, and how this optimal policy depends on the country's competition policy. The cross-country spillovers arising from tariff policy provide a useful contrast with those from cartel policy.

When governments set tariffs unilaterally and there is an effective competition policy in place, the optimal tariff for each country will be positive. The tariff has two effects: it increases consumer price and it shifts profits from foreign firms to home country firms. For small tariffs, the profit-shifting effect will dominate and welfare will be increased. A country can raise its welfare by using trade policy to give its firms a strategic advantage in oligopolistic industries, as noted by Brander and Spencer (1985). In contrast, raising profits of domestic firms in the home market through allowing cartels will not raise domestic welfare because the cartels will also raise the profits of foreign firms. Tariffs are a more attractive way to raise profits because they discriminate against foreign firms.

While the tariff is beneficial for the imposing country, the loss it imposes on the other country is larger than the gain of the imposing country. World welfare would be increased by the elimination of trade barriers. However, no country would reduce tariffs unilaterally. Thus, the setting of tariff policy is a Prisoner's Dilemma, so attaining the efficient world equilibrium will require an international agreement in which countries cooperate in tariff reduction. This contrasts with the case of cartel policy, where unilateral policy may lead to an optimal outcome if the rights to set policy are assigned appropriately and there are no cost spillovers between markets. The cross-country spillovers of cartel policy are subtler than those that arise with tariff policy.

If there is no competition policy in place, a country's optimal tariff policy will take into account its impact on the degree of collusion in the market. Davidson (1984) calculates the optimal tariff for the case in which the firms collude at the monopoly level if it is sustainable or compete at the Nash level if monopoly is unsustainable. Under this assumption, it will always be optimal to set the tariff at least high enough to break up the cartel. Syropoulos (1992) derives the optimal tariff in a model with $n^{h} = 0$

and a more flexible form of collusion in which foreign firms choose the minimum sustainable level of output. He shows that when the discount parameter is higher, which facilitates collusion, the optimal tariff will also be higher. These results establish that in the absence of a separate antitrust policy to prevent collusion, the optimal tariff of a country will be used to affect the level of collusion as well as to shift rents from the foreign firms.

3.3 Political Economy Considerations

The analysis in this section is based on the assumption that domestic consumers and producers carry equal weight in the government's objective function. An alternative assumption that is frequently made in the literature on tariff policy is that if interest groups form organized lobbies, they will be able to influence legislation through lobbying activities and campaign contributions to politicians. This can be captured by assuming that policymakers maximize a weighted social welfare function, where a greater weight is placed on the payoff to organized interest groups.[15] Since cartels will be most common in industries with a small number of firms, these industries are more likely to be able to overcome the free rider problems associated with organizing lobbying activities. Therefore, producers involved in cartels are likely to be politically powerful and receive greater weight in government decisions.

If the government objective function places sufficient weight on producer interests, then there will be some cartel outcomes that will be welfare improving for the importing countries. This would suggest that price-fixing arrangements might be allowed in countries where producer interests are strong, or that industries in which producers wield considerable political influence may be exempted from price-fixing legislation. Note however that firms are likely to find lobbying for tariff protection more attractive than lobbying for the ability to collude with foreign firms. The tariff hurts consumers, but raises government revenue and reduces foreign firm profits. As long as the government values tariff revenue more heavily than profits earned by foreign firms, the tariff will be a much less costly policy for transferring income to domestic producers.

On the other hand, international cooperation on cartel policy could be quite anticompetitive when producer interests are weighed heavily in all countries. In this case producer interests in both countries would lobby their governments for lax international cartel policies. One example is the agreement on aluminum production negotiated in 1994, under which Russia and other major aluminum producers agreed to cut production at a time of depressed prices. Such an agreement would have been in violation of domestic antitrust policies if negotiated by the firms without government support.

3.4 Other Forms of Competition Policy

In addition to using antitrust policy to affect the behavior of firms, governments can also use competition policy to affect the number of firms in an industry. One tool for controlling the number of firms is the ability to control mergers. Merger policy is

related to cartel policy in that if firms are prevented from behaving collusively, they might choose to coordinate their output levels by combining into a single firm. For example, the passage of the Sherman Antitrust Act was followed by a wave of mergers. Since there was no national authority or legislation to prevent mergers at that time, some attribute this merger wave to an attempt by firms to evade restrictions against price-fixing. However, there are two important differences between merger and cartel policies. The first is that mergers may result in cost reductions for the firms, and thus could potentially be justified on efficiency grounds. Thus, merger policy will require more discretion on the part of government authorities.

A second difference is that a merger simultaneously affects competitiveness in all markets, whereas cartel policy could in principle be evaluated on a market-by-market basis. The international spillovers in this case will be stronger than those associated with cartels, and would lead to a much stronger argument for international agreements in the case of mergers. Suppose that $n^{\mathrm{h}} = 2$ and $n^{\mathrm{f}} = 0$. The home country would prefer to allow the firms to compete in the home country market and behave as monopolists in the foreign market, but the foreign government is likely to impose a strict competition policy to prevent such cartel behavior in its market. If the home country is given the right to control mergers, it might choose to allow a merger between the two firms to take place if the gains from monopolizing the foreign market exceeded the costs in terms of higher prices in the home country market.

Governments may also have the ability to increase the number of firms in the market by using policies that encourage entry or by forcing the break-up of existing firms. In the previous example, the foreign government would be able to capture some of the profits in the local market by encouraging entry of a foreign competitor. This suggests that merger policy will have a profit-shifting role. Richardson (1999) and Horn and Levinsohn (2001) have examined the optimal competition policy by examining the Nash equilibrium in which governments choose the number of firms optimally. In the case where there are no trade barriers, the Nash equilibrium of the game in which countries choose the number of firms from their own country optimally will have more firms than is socially optimal because each country is using the entry of local firms to capture a larger share of profits in each market.[16] This creates a Prisoner's Dilemma with entry policy, similar to that noted in the case of tariff policy.

It should be noted that actual merger policy is more complex than indicated by a unilateral choice of the number of firms. In particular, the merger between two large US firms that operate in Europe would have to be approved by competition authorities in Europe as well as in the United States. The existence of overlapping jurisdictions will make the standard for mergers more stringent than the efficient standard for the world, since it requires gains to both countries from the merger. The issues of overlapping jurisdictions introduce issues into the approval of mergers that are similar to those that arise in the cartel case.

4 National Cartel Laws

In this section we examine the policies toward price-fixing in the United States and the United Kingdom. These countries provide an interesting contrast. Both countries started

from a common law tradition that treated cartels quite liberally, but the policies diverged in the twentieth century with the United States following a much more aggressive policy toward preventing price-fixing agreements.

Under the common law, the public interest was not considered in cases involving price-fixing and restraint of trade. This policy changed dramatically in the United States with the passage of the Sherman Antitrust Act in 1890, which declared illegal "every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several states, or with foreign nations." According to Peritz (2000), this language is generally interpreted as being a congressional mandate for the courts to develop a common law for competition policy. The question that needed to be resolved in subsequent Supreme Court decisions was whether this language made price-fixing illegal *per se*, which means that there is no allowable defense for it, or whether a *rule of reason* approach should be followed. The latter approach would require courts to evaluate the public costs and benefits of each particular agreement. Although a few early court decisions allowed some cartels to operate because they prevented "ruinous competition" among firms, the rulings eventually evolved to a *per se* standard.[17] This *per se* standard is what would be expected to be applied in the domestic market to maximize national welfare, although it arose more out of court interpretation than explicit legislation. In contrast, the decision of whether to break up a trust was based on a rule of reason approach. This approach is also consistent with national welfare maximization, since having large firms could be socially efficient if the large firms have lower costs than would smaller firms.

A second way in which US policy has been consistent with the national welfare model is the exemption of export trade associations, which occurred as a result of the Webb Pomerene Act of 1918. This act allowed cooperation among exporters as long as their actions did not have an adverse effect on the US market, which is consistent with Proposition 1. However, there are several other types of collusive activity that are exempt from antitrust law. The Clayton Antitrust Act exempted labor unions, and agricultural cooperatives have also been ruled to be exempt from antitrust law. Since agricultural lobbies and labor unions have historically been powerful political interest groups, it could be argued that these exemptions reflect the influence of organized lobbies on legislation. However, in each of these cases it could be argued that these organizations provide public goods that are socially desirable.

US policy on cartels seems broadly consistent with an objective of national welfare maximization, since cartels are prosecuted vigorously and there are relatively few sectors that are excluded. This policy has been the result of legislation that emphasizes the benefits of competition, and courts that have interpreted this as making cartels illegal *per se*. This strict interpretation eliminates the ability to use price-fixing arrangements as a means of maintaining prices in declining industries, a practice that has been followed in many countries where the *rule of reason* has been applied to cartel agreements.

Historically, Britain has been less aggressive in the prosecution of price-fixing arrangements than has the United States. There was no significant antitrust legislation in the United Kingdom in the first half of the twentieth century, and price-fixing agreements became quite common. In the United States, the support for antitrust policy came as a result of the widespread activity of large trusts. The support for the Sherman Act came primarily from midwestern states, where farmers and small businessmen

were faced with large trusts. In the United Kingdom, there was no such strong support for strict action against trusts. One explanation that has been proposed is that the United Kingdom was following a much more liberal trade policy at the time, so that trusts may have been a less severe problem. However, this lax policy remained even as tariff policy became more protectionist. Furthermore, the link between tariffs and trust activity is ambiguous as noted before. Scherer (1994) argues that the scale of the US market is what made the wave of trusts in the late nineteenth century so much larger than in other countries.

UK policy on cartels became stricter following World War II, when a government white paper on employment suggested a policy be implemented. The Restrictive Trade Practices Act of 1956 addressed this problem by outlawing restrictive practices that were harmful to the public interest. Restrictive agreements had to be registered with the Restrictive Practices Court, which would then evaluate the agreements. In contrast to the *per se* approach that had emerged in the United States by this time, the United Kingdom took a *rule of reason* approach. Although there was a clear presumption in the legislation that price-fixing agreements were against the public interest, the legislation specified six “gateways” that allowed courts to uphold the agreements. These exceptions allowed the court to weigh the cost of the agreement to consumers against benefits that might accrue from it to export earnings, employment, the public health, or other benefits to the public as purchasers of the product.[18]

The Restrictive Practices Court had the effect of substantially reducing the number of cartels operating in the United Kingdom. Once firms recognized that most agreements would be struck down, a number of agreements were abandoned. Subsequent modification of the legislation has expanded the range of products covered under the agreement to include commercial services, but has not significantly modified the structure of the agreement. However, this policy is still less aggressive than the US policy in pursuing cartels, since it relies to a large degree on voluntary revelation of agreements and does not impose severe fines for agreements that harm the public interest. Also, the government’s power to investigate tacit agreements that have not been reported is quite weak.[19] For industries that are of substantial size, EU cartel policy is also relevant for UK firms. Article 85 of the EU treaty identifies agreements that directly or indirectly fix the buying or selling price of products as being inconsistent with a common market and are to be considered automatically void. However, an exemption is also granted to an agreement that “contributes to improving the production or distribution of goods or promoting technical or economic progress.” Cini and McGowan (1998) report that this exemption has led the European Commission to allow “crisis cartels” for industries such as shipbuilding, coal, and steel that are declining or in recession.

In summary, UK policy (as well as the policy of most industrialized countries) has evolved over time to a position that is closer to that of the United States, although the policy continues to have some “rule of reason” components. The political demand for agreements to fix prices to protect firms in certain sectors in the United States has primarily been met using instruments of trade protection, rather than antitrust instruments. The antidumping process is likely to play the role of providing protection to declining industries, which is allowed through “crisis cartels” in the European Union. In addition, the US government has in the past negotiated voluntary export restraints in certain industries. Since these processes involve restrictions in quantities by foreign

producers (with the foreign producers being allowed to keep the proceeds of the trade restriction), they have an effect similar to a cartel. Thus, the political demand for cartels is met through channels outside the reach of antitrust law in the United States.

5 Conclusions

The first question addressed in this chapter was the effect of trade policy on collusion between home and foreign firms. We identified three ways in which trade policy can affect collusion: through its effect on the relative bargaining power of firms that are colluding, on the types of agreements that can be sustained, and on the ability to avoid antitrust policy. The results showed that trade liberalization will reduce the bargaining power of home firms relative to foreign firms involved in the cartel. The impact of trade liberalization on the sustainability of collusion was shown to be ambiguous, and depends on the type of trade instrument used and on whether firms are Cournot or Bertrand competitors. The multi-market contact model suggests that trade liberalization will change the types of agreements that firms reach. For high tariff levels, firms will choose agreements that involve geographic specialization of markets to reduce trade costs, whereas with low tariffs they will tend to choose similar market shares to reduce the deviation incentives.

Overall, it seems unlikely that reductions in trade barriers would be responsible for a significant increase in cartel activity. However, other changes associated with globalization could make collusion easier to sustain. For example, if globalization results in increased concentration in an industry or reduced costs of monitoring cartel agreements it can facilitate collusion.

The second question examined was the potential gains from international harmonization of competition policy. The results indicated that the cross-country spillovers from cartel policy tend to differ from those of tariff policy, because governments that are maximizing national welfare have the correct incentives to monitor seller cartels operating within their own borders. The main roadblock to national policy is the presence of state-sponsored cartels, which are currently beyond the reach of national governments attempting to prevent collusion in their own market. Thus, further gains in reducing collusion would require international agreements to limit state sponsorship. The second type of gain would be from agreements that facilitate cooperation between antitrust authorities in different countries, which would make it easier for countries to pursue international cartels and to reduce the free rider problem associated with cartel enforcement.

The final question concerned the political economy of cartel policy. Economic theory would suggest that price-fixing agreements should be prevented under virtually all circumstances, but political influence could well lead to a policy that allows cartels to operate among firms that are politically powerful. The historical development of US policy since the Sherman Antitrust Act has resulted in making price-fixing illegal *per se*. Since this takes place through a judicial process, the primary potential for political influence on the antitrust process is through the enthusiasm with which such cases are pursued by the justice department. However, the trade protection process has facilitated

output restrictions among home and foreign firms through the antidumping process and through the negotiation of voluntary export restraints. In contrast, the United Kingdom and the European Union follow a *rule of reason* approach, which allows more discretion in allowing cartels to operate. This suggests that the political pressure for collusion in declining industries in the United States may be realized through the antidumping process, rather than direct exemptions from antitrust enforcement. The linkage between the use of antitrust exemptions and trade protection to raise profits in declining industries is an area that merits further research.

Acknowledgments

I thank Costas Syropoulos, Kala Krishna, and seminar participants at Penn State and the University of Copenhagen for helpful comments and discussion.

Appendix

Derivation of Cournot Equilibrium. Firm i will choose its output level, q^i, to maximize $\pi(X, c^i) = (A - X - q^i - c^i)q^i$, where X is the aggregate output of all other firms and c^i is firm i's marginal cost. Firm i's reaction function and profit function will then be given by

$$\tilde{q}(X, c^i) = \frac{(A - X - c^i)}{2} \qquad \tilde{\pi}(X, c^i) = \frac{(A - X - c^i)^2}{4}. \tag{A.1}$$

Using the symmetry assumption, the reaction function of all home (foreign) firms in the home market is derived from (A.1) using $c^i = 0$ ($c^i = t^{\mathrm{h}}$). For home firms, the Nash equilibrium output must satisfy $q^{\mathrm{h}} = \tilde{q}((n^{\mathrm{h}} - 1)q^{\mathrm{h}} + n^{\mathrm{f}}x^{\mathrm{f}}, 0)$. The locus of values satisfying this condition is plotted as the R^{h} locus in Figure 11.1. Similarly, for foreign firms, the Nash equilibrium output must satisfy $q^{\mathrm{f}} = \tilde{q}(n^{\mathrm{h}}q^{\mathrm{h}} + (n^{\mathrm{f}} - 1)x^{\mathrm{f}}, t^{\mathrm{h}})$, which is plotted as the R^{f} locus. The Nash equilibrium output is obtained by solving these equations simultaneously, yielding

$$q^{\mathrm{hN}}(t^{\mathrm{h}}) = \frac{A + n^{\mathrm{f}}t^{\mathrm{h}}}{n^{\mathrm{f}} + n^{\mathrm{h}} + 1}, \qquad x^{\mathrm{fN}}(t^{\mathrm{h}}) = \frac{A - (n^{\mathrm{h}} + 1)t^{\mathrm{h}}}{n^{\mathrm{f}} + n^{\mathrm{h}} + 1}, \qquad \pi^{i\mathrm{N}}(t^{\mathrm{h}}) = q^{i\mathrm{N}}(t^{\mathrm{h}})^2$$
$$\text{for } i = \mathrm{h}, \mathrm{f}. \tag{A.2}$$

Proof of Proposition 1. The profit functions for the home market for a representative firm with an agreement $(q^{\mathrm{h}}, x^{\mathrm{f}})$ will be

$$\pi_{\mathrm{hh}}(q^{\mathrm{h}}, x^{\mathrm{f}}) = (A - n^{\mathrm{h}}q^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}})q^{\mathrm{h}} \qquad \pi_{\mathrm{fh}}(q^{\mathrm{h}}, x^{\mathrm{f}}, t^{\mathrm{h}}) = (A - n^{\mathrm{h}}q^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}} - t^{\mathrm{h}})x^{\mathrm{f}}. \tag{A.3}$$

The deviation payoffs obtained from (A.1) are

$$\pi^{\mathrm{D}}_{\mathrm{hh}}(q^{\mathrm{h}}, x^{\mathrm{f}}) = \frac{(A - (n^{\mathrm{h}} - 1)q^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}})^2}{4}$$
$$\pi^{\mathrm{D}}_{\mathrm{fh}}(q^{\mathrm{h}}, x^{\mathrm{f}}, t^{\mathrm{h}}) = \frac{(A - (n^{\mathrm{f}} - 1)x^{\mathrm{f}} - n^{\mathrm{h}}q^{\mathrm{h}} - t)^2}{4}. \tag{A.4}$$

From (1), $(q^{\mathrm{h}}, x^{\mathrm{f}})$ will be sustainable for firm i if $Z_{i\mathrm{h}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}, \delta) \equiv \pi_{i\mathrm{h}}^{\mathrm{C}}(q^{\mathrm{hC}}, x^{\mathrm{hC}}, t^{\mathrm{h}}) - (1 - \delta)\pi_{i\mathrm{h}}^{\mathrm{D}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}) - \delta\pi_{i\mathrm{h}}^{\mathrm{N}}(t^{\mathrm{h}}) \geqslant 0$. Since $\partial Z_{i\mathrm{h}}/\partial\delta = \pi_{i\mathrm{h}}^{\mathrm{D}} - \pi_{i\mathrm{h}}^{\mathrm{N}} > 0$ for profitable cartels, $\delta^{\min}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}})$ is the solution to $Z_{i\mathrm{h}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}, \delta) = 0$. Note also that since $\pi_{i\mathrm{h}}^{\mathrm{D}}$ is bounded above, $\delta^{\min}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}) < 1$ for any agreement with $\pi_{i\mathrm{h}}^{\mathrm{C}} > \pi_{i\mathrm{h}}^{\mathrm{N}}$.

To establish part (a), note that from (A.2) to (A.4) a symmetric agreement with $q^{\mathrm{h}} = x^{\mathrm{f}} = q$ will be sustainable at free trade if

$$H(q, \delta) \equiv Z_{i\mathrm{h}}(q, q, 0, \delta) = q(A - Nq) - \frac{(1-\delta)(A + q(1-N))^2}{4} - \frac{A^2\delta}{(1+N)^2} \tag{A.5}$$

where $N = n^{\mathrm{h}} + n^{\mathrm{f}}$. $H(q, \delta)$ is strictly concave in q, with $H(q^{\mathrm{N}}(0), \delta) = 0$ and $H_q(q^{\mathrm{N}}(0), \delta) < 0$ for all $\delta \in (0, 1)$. This means that for $\delta > 0$, there will exist an interval $[\tilde{q}(\delta), q^{\mathrm{N}}(0)]$ of outputs for which $H(q, \delta) \geqslant 0$. $\tilde{q}$ will be decreasing in δ, because H must be increasing in δ when evaluated at $H(q, \delta) = 0$. The most profitable, sustainable symmetric agreement will be $q = \max(\tilde{q}(\delta)A/(2N))$. The minimum discount factor to sustain q is then given by $\delta^{\min} = \tilde{q}^{-1}(q)$ for $q \geqslant A/(2N)q$. The first part of (a) holds because $H(A/(2N))$ is decreasing in N. The second part follows from the fact that $\tilde{q}$ is decreasing in δ and $H(q^{\mathrm{N}}(0)) = 0$ for all δ.

To establish (b), assume that if antitrust authorities discover a cartel, the members must pay a fine F and are prevented from engaging in any future collusion. Letting ϕ be the probability of being caught, the expected present value of a cartel is $V^{\mathrm{C}} = (\pi_{i\mathrm{h}}^{\mathrm{C}} - \phi F + \delta\phi\pi_{i\mathrm{h}}^{\mathrm{N}}/(1-\delta))/(1-(1-\phi)\delta)$. The agreement will be sustainable if the average payoff under the cartel, $(1-\delta)V^{\mathrm{C}}$, exceeds the payoff in a deviation. This yields the no deviation requirement that

$$Z_{i\mathrm{h}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}, \delta, \phi, F) = \pi_{ij}^{\mathrm{C}} - \phi F - (1 - (1-\phi)\delta)\pi_{ih}^{\mathrm{D}} - \delta\pi_{ih}^{\mathrm{N}}(1-\phi) \geqslant 0. \tag{A.6}$$

As in the previous case, $Z_{i\mathrm{h}}(q^{\mathrm{hC}}, x^{\mathrm{fC}}, t^{\mathrm{h}}, \delta^{\min}, \phi, F) = 0$. Differentiation of (A.6) gives $\partial\delta^{\min}/\partial\phi = -(\partial Z_{i\mathrm{h}}/\partial\phi)/(\partial Z_{ij}/\partial\delta) > 0$ and $\partial\delta^{\min}/\partial F = -(\partial Z_{ih}/\partial F)/(\partial Z_{ij}/\partial\delta) > 0$. □

Proof of Proposition 2. Differentiating (3) yields

$$dW^{\mathrm{h}} = n^{\mathrm{h}}(p^{\mathrm{h}} + n^{\mathrm{f}}x^{\mathrm{f}})dq^{\mathrm{h}} + n^{\mathrm{f}}(t^{\mathrm{h}} + n^{\mathrm{f}}x^{\mathrm{f}})dx^{\mathrm{f}} + n^{\mathrm{h}}[(p^{\mathrm{f}} - t^{\mathrm{f}} - n^{\mathrm{h}}x^{\mathrm{h}})dx^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{h}}dq^{\mathrm{f}}]. \tag{A.7}$$

The first term in parentheses in (A.7) is positive for $p^{\mathrm{h}} \geqslant 0$ and the second term is positive for $t^{\mathrm{h}} \geqslant 0$. This establishes that iso-welfare contours will be negatively sloped in Figure 11.1 at all the agreements that raise cartel profits. To establish the result, it is sufficient to show that home welfare is increasing in q^{h} on the negatively sloped portion of $\pi_{\mathrm{hh}}^{\mathrm{N}}$ and is increasing in x^{f} on the negatively sloped portion of $\pi_{\mathrm{fh}}^{\mathrm{N}}$. This will ensure that the Nash equilibrium yields higher profits than any point on the boundary of the profitable cartel agreements.

Along the $\pi_{\mathrm{fh}}^{\mathrm{N}}$ curve, $dq^{\mathrm{h}}/dx^{\mathrm{f}} = ((p^{\mathrm{h}} - t^{\mathrm{h}})b^{\mathrm{h}} - n^{\mathrm{f}}x^{\mathrm{f}})/(n^{\mathrm{h}}x^{\mathrm{f}})$. Substituting this result in (A.7) yields $dW^{\mathrm{h}}/dx^{\mathrm{f}} = (p^{\mathrm{h}}(p^{\mathrm{h}} - t^{\mathrm{h}})b^{\mathrm{h}})/x^{\mathrm{f}} > 0$, so home welfare is increasing in x^{f} along the $\pi_{\mathrm{fh}}^{\mathrm{N}}$ locus. Using a similar argument, the slope of the $\pi_{\mathrm{hh}}^{\mathrm{N}}$ locus is $dx^{\mathrm{f}}/dq^{\mathrm{h}} = (b^{\mathrm{h}}p^{\mathrm{h}} - n^{\mathrm{h}}q^{\mathrm{h}})/(n^{\mathrm{f}}q^{\mathrm{h}})$.

Using (A.7), $dW^{\rm h}/dx^{\rm f} = (p^{\rm h}(n^{\rm h}q^{\rm h} + n^{\rm f}x^{\rm f}) + t^{\rm h}(p^{\rm h}b^{\rm h} - n^{\rm h}q^{\rm h}))/q^{\rm h}$. This expression must be positive, since $t^{\rm h} < p^{\rm h}$ is required for $x^{\rm f} > 0$. This yields the result that the home iso-welfare contour will be flatter than the $\pi^{\rm N}_{\rm hh}$ locus at the Nash equilibrium point.

The policy that is efficient from a world welfare perspective can be obtained by summing home and foreign welfare, $dW^{\rm h}+dW^{\rm f} = p^{\rm h}(n^{\rm h}dq^{\rm h}+n^{\rm f}dx^{\rm f})+p^{\rm f}(n^{\rm f}dq^{\rm f}+n^{\rm h}dx^{\rm h})$. World welfare contours will have a slope of -1 in Figure 11.1, so welfare at the Nash equilibrium point must exceed that at any profitable cartel agreement.

Notes

1 International cartels and market-sharing agreements were a common feature of the early twentieth century, with cartels often operating with tacit (if not explicit) government approval. A 1944 report by the Kilgore Committee identified 105 products that had been identified in antitrust proceedings as being part of international cartels (Nussbaum, 1986). Government policies (particularly in Europe) became less favorable toward cartels in the second half of the twentieth century.

2 Competition policy includes government policies that regulate actions of firms (e.g., price-fixing, information exchange) as well as the structure of industry (merger policy, subsidies to entry). This chapter focuses on price-fixing policy. Broader surveys of competition policy issues include Scherer (1994), Bliss (1996), and Levinsohn (1996).

3 Note that this line is not a reaction function in the usual sense. Along this locus, $q^{\rm h}$ is a best response for a home firm when foreign firms choose $x^{\rm f}$ and all other home firms choose $q^{\rm h}$.

4 This point is made by Salant et al. (1983) in the context of mergers. If firms merge and there is no cost-saving effect of the merger, then the effect of a merger is identical to that of the formation of a cartel. Mergers in Cournot oligopoly that do not involve cost savings will then be profitable only in cases where the merging firms have very large market shares. This market share requirement is less restrictive in the case where the products are imperfect substitutes and firms are Cournot competitors, but it continues to be the case that nonmembers will benefit from the formation of the cartel. In contrast, Deneckere and Davidson (1985) show that when firms are Bertrand competitors who choose prices to maximize profits (treating the prices of all other firms as given) with products that are imperfect substitutes, collusion among any subset of the firms will be profitable for the colluding firms. Note however that the non-colluding firms will be free riders in the Bertrand case as well, because they will earn higher profits than the colluding firms in the new equilibrium. They also show that the profits of the coalition members will be increasing in the number of participants. Thus, the result that firms forming a collusive agreement will want to include as many of the market participants as possible holds in both the price-setting and quantity-setting cases.

5 Under the common law, price-fixing agreements were allowed but would not be enforced by courts. In international price-fixing agreements involving national governments, such as the Organization of Petroleum Exporting Countries (OPEC), no courts are available to enforce an agreement if a participating country chooses to deviate.

6 This result is derived under the assumption that the probability of detection is independent of the type of agreement chosen. The effect of the choice of output levels on the probability of detection is an area that has received relatively little attention. This could be particularly important in the case of multi-market collusion to be considered later.

7 Syropoulos (1992) examines the problem of sustaining collusion among quantity-setting oligopolists when products are differentiated and there are no domestic firms. In this case higher tariffs make it easier to support a cartel with a given output level. He also considers

the effect of quantitative restrictions on foreign firms and shows a non-monotonic relationship between the size of the quota and the minimum discount factor. Quotas that are less restrictive than the Nash output level make collusion easier to support because they deter deviation without reducing the threat of punishment. Tariffs that are more restrictive than the Nash level make collusion harder to support by reducing the threat of punishment.

8 Prusa's theoretical model would predict that all petitions should be withdrawn. Zanardi (2004) extends the model to include bargaining costs, so that collusion may not result when bargaining costs are high. He also finds some support for the role of bargaining costs in his empirical analysis. Hartigan (1995) raises the possibility that when the firm's costs are private information, the antidumping process can affect collusion by revealing the costs of the firms during the fact finding process.

9 Prusa (1992) shows that this can arise in the case where firms are Bertrand competitors and the discount parameter is sufficiently high that efficient agreements are sustainable.

10 This result is due to the fact that agreement and deviation payoffs are additive across markets in this case. Spagnola (2001) shows that linked agreements could be less attractive in cases where deviation in one market makes deviation in the other market more profitable.

11 As in most analyses of cartel behavior, this result assumes identical products and a fixed number of firms. An exception to this approach is Fershtman and Pakes (2001), who analyze a model of collusion in which product quality and entry are endogenous. They use a simulation analysis to show that collusion results in both higher prices and higher product quality, and in some cases welfare may be higher under collusion.

12 The Nash equilibrium would be dominated by an outcome that imposed marginal cost pricing through the use of price ceilings and allowed only one firm (assuming that there are fixed costs associated with each firm). However, such a policy would require substantial information on the part of the government and could be quite costly if mistakes are made. As a result, we consider only second-best policies such as bans on collusion, tariffs, and control of the number of firms.

13 For example, US firms that were part of a cartel registered under the Webb Pomerene Act used the defense that this cartel was an act of state when they were prosecuted by the European Union. The argument was rejected by the European Union on the basis that the US law allowed, but did not require, the export cartel activity. Scherer (1994) provides a detailed discussion on this point.

14 In practice, collusion among purchasers is much less common than collusion among sellers. This probably reflects the fact that buyers are typically less concentrated than sellers, so collusion may be more difficult to sustain.

15 Grossman and Helpman (1994) provide a formal justification for this approach in a model where lobbies use campaign contributions to influence the decisions of politicians.

16 Richardson (1999) and Horn and Levinsohn (2001) both examine the question of how the level of trade policy affects the optimal number of firms. One might anticipate that trade liberalization would result in an incentive to increase the number of firms, since both trade liberalization and entry are being used to shift profits. However, both papers find that the effect of trade liberalization on the Nash equilibrium number of firms is ambiguous. Richardson analyzes the case where governments impose tariffs, and shows that a reduction in the tariff will result in a larger number of firms if the fixed costs of entry are sufficiently large when the government maximizes national welfare. If the home country cares only about the profits of firms, trade liberalization will result in an increase in the optimal number of firms for the home country. Horn and Levinsohn consider the use of export subsidies and show that the effect of a reduction in export subsidies on the marginal benefit of entry is ambiguous.

17 An important case in establishing this was US versus Trenton Potteries, in which Trenton Potteries argued that its agreement to fix prices broke down frequently and had little effect on the market and therefore did not result in an unreasonable price. A lower court had ruled that the issue of whether the price was reasonable was not an element to be considered in the case. In the majority opinion upholding the lower court ruling, Justice Stone said, "The aim and result of every price-fixing agreement is … the elimination of one form of competition. The power to fix prices, whether exercised reasonably or not, involves the power to control the market."

18 Stevens and Yamey (1965) report that of the first 21 cases brought before the court that involved price-fixing agreement, 7 were upheld by the court. The agreements that were upheld were generally viewed as having benefits for consumers or in export markets. Some of the justifications seem questionable. For example, a price-fixing agreement on nuts and bolts was upheld because the agreement saved consumers the cost of shopping around.

19 Symeonidis (1998) provides a discussion of the weaknesses of UK policy in pursuing cartels.

References

Bernheim, B. Douglas and Michael D. Whinston 1990: Multimarket contact and collusive behavior. *RAND Journal of Economics*, 21, 1–26.

Bliss, Christopher 1996: Trade and competition control. In Jagdish Bhagwati and R. Hudec (eds.), *Fair Trade and Harmonization: Economic Analysis*. Cambridge: MIT Press.

Bond, Eric W. and Costas Syropoulos 2002: Reciprocal trade liberalization and multimarket collusion. Manuscript.

Brander, James and Barbara Spencer 1985: Export subsidies and international market share rivalry. *Journal of International Economics*, 18, 83–100.

Cini, Michelle and Lee McGowan 1998: *Competition Policy in the European Union*. New York: St Martin's Press.

Davidson, Carl 1984: Cartel stability and tariff policy. *Journal of International Economics*, 17, 219–37.

Deneckere, Raymond and Carl Davidson 1985: Incentives to form coalitions with Bertrand competition. *RAND Journal of Economics*, 16, 473–86.

Fershtman, Chaim and Ariel Pakes 2001: A dynamic oligopoly with collusion and price wars. *RAND Journal of Economics*, 31, 207–36.

Friedman, James 1971: A non-cooperative equilibrium for supergames. *Review of Economic Studies*, 38, 1–12.

Fung, K. C. 1992: Economic integration as competitive discipline. *International Economic Review*, 33, 837–47.

Grossman, Gene M. and Elhanan Helpman 1994: Protection for sale. *American Economic Review*, 84, 833–50.

Hartigan, James 1995: Collusive aspects of cost revelation through antidumping complaints. *Journal of Institutional and Theoretical Economics*, 151, 478–89.

Hexner, Erving 1946: *International Cartels*. Durham: University of North Carolina Press.

Horn, Hendrik and James Levinsohn 2001: Merger policy and trade liberalization. *Economic Journal*, 111, 244–76.

Levinsohn, James 1996: Competition policy and international trade. In Jagdish Bhagwati and R. Hudec (eds.), *Fair Trade and Harmonization: Economic Analysis*. Cambridge: MIT Press.

Nussbaum, Helga 1986: International cartels and multinational enterprises. In Alice Teichova, Maurice Levy-Leboyer, and Helga Nussbaum (eds.), *Multinational Enterprise in Historical Perspective*. Cambridge University Press.

Peritz, Rudolph J. 2000: *Competition Policy in America: History, Rhetoric, and Law*. Oxford: Oxford University Press.

Prusa, Thomas 1992: Why are so many anti-dumping petitions withdrawn? *Journal of International Economics*, 33, 1–20.

Richardson, Martin 1999: Trade and competition policy: concordia discors? *Oxford Economic Papers*, 51, 649–64.

Rotemberg, Julio and Garth Saloner 1989: Tariffs vs. quotas with implicit collusion. *Canadian Journal of Economics*, 89, 237–44.

Salant, Stephen, Switzer, Sheldon, and Reynolds, Robert 1983: Losses due to merger: the effects of exogenous change in industry structure on Cournot–Nash equilibrium. *Quarterly Journal of Economics*, 98, 185–99.

Scherer, Frederick M. 1994: *Competition Policies for an Integrated World Economy*. Washington, DC: Brookings Institution.

Spagnolo, Giancarlo 2001: Issue linkage, credible delegation, and policy coordination. CEPR discussion paper #2778.

Stevens, Robert B. and Basil S. Yamey 1965: *The Restrictive Practices Court: A Study of the Judicial Process and Economic Policy*. London: Weidenfeld and Nicolson.

Symeonidis, George 1998: The evolution of UK cartel policy and its impact on market conduct and structure. In Stephen Martin (ed.), *Competition Policy in Europe*. Amsterdam: Elsevier.

Syropoulos, Constantinos 1992: Quantitative restrictions and tariffs with endogenous firm behavior. *European Economic Review*, 36, 1627–46.

U.S. Department of Justice 2000: *International Competition Policy Advisory Committee, Final Report* (http://www.usdoj.gov/atr/icpac/finalreport.htm).

World Trade Organization 1997: *Annual Report for 1997*.

Zanardi, Maurizio 2004: Antidumping law as a collusive device. *Canadian Journal of Economics*, 37:1, 95–122.

12

Modern Commercial Policy: Managed Trade or Retaliation?

Thomas J. Prusa and Susan Skeath

Chapter Outline

In this chapter, we study worldwide antidumping (AD) case filing patterns in order to determine the extent to which economic versus strategic incentives explain the behavior of countries employing AD actions. We compare four main hypotheses: big supplier, big change in imports, tit-for-tat (TFT), and the club effect. We find strong support that AD actions are directed at big suppliers, but far weaker evidence that AD actions are filed against suppliers whose imports have surged. We also find very strong evidence that AD actions are used strategically to deter further use of AD and/or to punish trading partners who have used AD. Our findings reject the notion that the rise in AD activity is solely explained by an increase in unfair trading.

1 Introduction

Modern commercial policy has been influenced significantly by the global movement toward freer trade and more open borders. Where tariffs and quotas once held sway over policymakers, antidumping (AD) and countervailing duty cases are now the path to World Trade Organization (WTO)-legal protection. Antidumping in particular has become increasingly prevalent worldwide.

To say that AD is now the most popular form of international trade protection is an understatement. In terms of the quantity of trade litigation, AD has lapped the field – several times over. In 1999 for instance, WTO members reported 360 AD disputes. By contrast, they reported only 42 countervailing duty disputes. More astounding perhaps is the fact that countervailing duty law has long been the *second* most commonly

used trade statute. Countervailing duty law takes the silver medal, but it is a far distant second.

The preeminence of AD is neither an entirely recent phenomenon nor simply a one-year anomaly. In the United States, for instance, over the past 25 years there has been more than twice as many AD disputes as countervailing duty allegations. In fact, there have been more disputes filed under the US statute than under *all other* US trade statutes put together. The same is true for the European Union. Antidumping is simply the 900-pound gorilla of trade laws.

The rise in AD activity was not entirely unexpected. Ethier (1982, p. 488) warned that AD would be "the principal battleground for the 'new protectionism' as concerns trade in manufactures among the developed economies." But throughout most of the postwar era, it was easy for General Agreement on Tariffs and Trade (GATT) negotiators to look the other way when it came to AD. After all, during that period, AD was really only a significant trade impediment for a handful of countries. Most countries either did not have an AD statute or chose not to use the law.

However, Ethier's words now ring true and low levels of AD use are no longer the norm. As a number of more recent authors have pointed out, the dramatic proliferation in the use of AD since 1980 is a particularly alarming development (Miranda, Torres, and Ruiz 1998; Prusa, 2001). More than three times the number of countries have been involved in AD disputes – both in terms of those filing cases and those being filed against – during the 1990s as compared with the 1980s.

One of the most disconcerting aspects of the growth, and certainly the most crucial for the issues examined in this chapter, is the fact that the "traditional" users of AD (Australia, Canada, the European Union, New Zealand, and the United States) have been joined by an ever-growing group of "new" users, including South Africa, Brazil, and Mexico among others. Most of the new users did not even have their own AD statutes until the late 1980s.

As the number of users and cases filed annually grow, it is increasingly difficult to identify the motives of the users of AD and to argue that increased usage signals merely an increase in unfair trade. We are concerned, for example, that AD is increasingly being used as pure protectionism rather than as a valid trade remedy. It appears that AD's seductive lure of "WTO legitimate" protectionism is hard to resist and that an increasingly large number of countries are beginning to abuse the rules.

In Prusa and Skeath (2002), we provide some nonparametric evidence to suggest that the recent rise in AD activity cannot be explained solely by an increase in unfair trading practices. Here, we probe more deeply and examine whether there is any more significant empirical support for the proposition that AD filings are motivated by factors other than pure economics – by factors other than dollars and pesos so to speak. For the purposes of this chapter, economic motivations include, say, filing AD complaints after a surge in imports, or perhaps even simpler, filing AD complaints against countries who hold the biggest import market share. In other words, we are willing to be agnostic as to whether such possibly benign patterns imply any unfair behavior. It is certainly possible that these patterns are consistent with unfairly low priced imports.

By contrast, "non-economic" factors behind AD use include what we refer to as strategic or retaliatory motivations. For instance, when a country tends to file AD complaints against precisely those countries that have previously filed cases against it,

we argue that its AD use is at least partially motivated by retaliation. Such behavior is completely inconsistent with the notion of AD as a way to combat unfair import practices. In fact, it suggests that countries view AD as a way to combat unfair governmental policies rather than a private firm's unfair practices.

In this chapter we study worldwide AD filing patterns in order to determine whether we can identify the motivations for the increased use of AD disputes. In order to do so, we have compiled a database of all AD actions filed worldwide between 1980 and 1997 supplemented by trade statistics for each country filing cases.

Our analysis provides strong support for the view that strategic incentives influence AD filings. In particular, we find that, all else equal, AD users are far more likely to investigate suppliers who are also AD users. In other words, there appears to be a "club" effect, where countries that have used the law are also more likely to be subject to investigations. Further, we find some evidence that countries are also more likely to use AD specifically against those countries that investigated them in the past. This is the type of pattern one might expect from game players employing tit-for-tat (TFT) strategies. All things considered, we believe our findings reject the notion that the rise in AD activity is solely explained by an increase in unfair trading.

The chapter proceeds as follows. In Section 2 we present a brief history of AD legislation along with a discussion of the growth in AD activity over the past two decades. In Section 3 we review the theoretical models motivating our empirical approach. In Section 4 we present parametric results. Concluding comments are made in Section 5.

2 History and Background

2.1 Short History of Antidumping Legislation

Although opinions regarding AD vary, the facts regarding the laws and their use are clear. The focus of recent attention on AD has been the upsurge in use in the past two decades. However, AD laws have been in place for considerably longer and have an interesting history.

The first modern AD law (Act to Amend the Customs Tariff, 1897) was passed in Canada in 1904. AD provisions were formally passed soon after in New Zealand (Agricultural Implement Manufacture, Importation and Sale Act, 1905) and Australia (Australian Industries Preservation Act, 1906). Both the United States and Britain passed their own statutes within the next 15 years (the US Revenue Act of 1916, and the Safeguarding of Industries Act of 1921, respectively) as did several European countries (e.g., France, in 1908) and most of the British Commonwealth (Finger,1993, p. 16).

Some of the original AD laws were industry-specific, as in New Zealand, or based loosely on earlier competition law, as in the United States. The 1916 American law used the "intent to restrain competition" language of the Sherman Act of 1890. Such laws were relatively quickly replaced with statutes more similar to the Canadian one in which a more general "injury from imports" standard prevails (Finger, 1993, p. 21). The New Zealand Customs Amendment Act (1921) and the US Antidumping Act

(1921) both allowed broader administrative authority in the imposition of AD penalties with evidence of injury than did earlier acts.

Over the intervening years, these statutes have been variously amended, repealed, and replaced with more modern laws up to and through the period of negotiation of GATT. In the United States, the most major changes occurred with the Tariff Act of 1930 and the Trade Agreements Acts of 1974 and 1979, although there were additional changes made in both 1984 (Trade and Tariff Act) and 1988 (Omnibus Trade Act). Many of the later changes made it easier for domestic firms both to prove the existence of dumping and to obtain relief under the law, including extensions of the definition of "less than fair value" to include both international price discrimination and sales below cost (Hufbauer and Elliott, 1994, p. 113). Elsewhere, in Europe, Canada, and the British Commonwealth countries in particular, similar changes were being made to AD statutes throughout the century.

As individual nations tinkered with existing AD statutes or passed new ones during the latter half of the twentieth century, AD also became institutionalized in Article VI of the original formulation of GATT in 1947. Article VI does not, however, specify how signatory countries should identify dumping behavior or how they should make injury determinations. As a result, GATT member countries negotiated an additional agreement aimed at providing pertinent information on the implementation of Article VI. This "Agreement on the Implementation of Article VI" (more commonly called the Antidumping Agreement or the Antidumping Code), most recently revised during the 1994 Uruguay Round, regulates the use and application of Article VI by member countries. In particular, it details procedures for much of the AD process, including the determination of injury, collection of evidence, and imposition of duties.

The original version of the Code, negotiated during the Kennedy Round (1967), was of limited significance due in part to the fact that the United States did not sign that document. The Tokyo Round (1979) then produced some amendments to the Code that helped to transform what had been a little-used trade statute into the workhorse of international trade protection that is modern AD law. First, the definition of "less than fair value" sales was broadened to incorporate both price discrimination and below-cost sales, as had been done in the United States some years earlier. In addition, the Tokyo Round Agreement removed the provision that dumped imports be demonstrated as the "principal cause of material injury" before duties could be imposed.

The Tokyo Round "improvements" in the AD Code were significant in opening the door to AD case filings. They also provided an important and useful general framework for investigations. Even the Tokyo Round agreement, however, bound only the 27 GATT member countries. It was not until the Uruguay Round (1994) that the AD Agreement came into force for a broader set of countries worldwide.

2.2 Trends in Antidumping Filing Behavior

Despite the amount of legislative time and paperwork devoted to AD in the United States, Europe, and the British Commonwealth over the first six decades of the twentieth century, and by GATT signatories after 1947, AD disputes were not very common until the 1970s. Hufbauer (1999) notes that the United States filed a total of only 246 cases

between 1934 and 1974, an average of only 6 per year. Worldwide, fewer than a dozen cases were filed by GATT member countries each year between 1948 and the early 1960s (Finger and Fung, 1994).

This relatively low level of activity began to change in the 1970s, due in part to the amendments made to the GATT AD Code in 1979 and to country-specific changes in AD legislation. In the United States, for example, an average of 35 cases were filed each year in the second half of the 1970s (Finger and Fung, 1994, p. 199). Looking more broadly at the AD case filing history, we see that AD actions went on to become even more common during the 1980s; an average of 139 cases were filed each year from 1980 through 1989.

Despite the increase in its use in the two decades between 1970 and 1990, many did not view AD as a trading-system problem because it was an instrument still actually wielded by only a handful of countries. Prior to the late 1980s, essentially all AD actions were initiated by the five "traditional" users: the United States, Canada, the European Union, Australia, and New Zealand. Specifically, traditional users were responsible for a total of 99.4 percent of all of the AD cases filed between 1980 and 1985 as well as for more than 95 percent of the cases filed during the entire decade of the 1980s.

By the late 1980s, however, the traditional users found they were no longer the only users of AD. In Figure 12.1 we detail the number of cases filed annually from 1980 through 1998, and we distinguish cases filed by traditional and new users. Looking at Figure 12.1, we see that by the late 1980s the traditional users were joined by a group of "new" users. These new users were responsible for no more than 5 percent of case filings in any year through 1986, but filed 20 percent of all AD cases in 1987 and no less than 11 percent in any year thereafter.

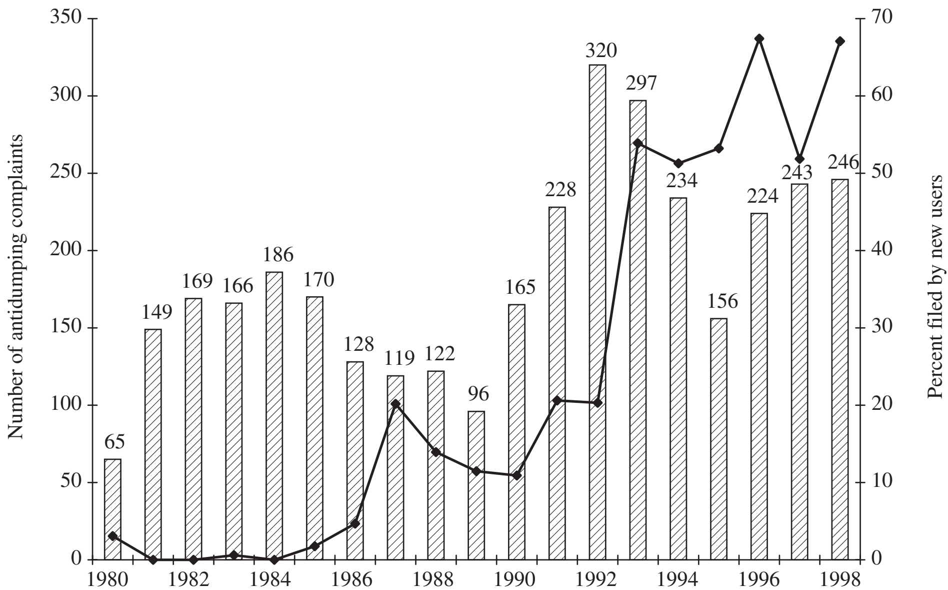

Figure 12.1 Worldwide use of antidumping, 1980–98

Table 12.1 Antidumping activity: 1980s versus 1990s

	Percent of total AD cases (worldwide) (%)		**Year country began using AD (GATT/WTO reports)**	**No. AD actions against country prior to the adoption of own AD statute**
	1980–9	**1990–8**		
USA	28.2	16.4	–	–
Australia	25.5	16.0	–	–
Canada	21.3	6.5	–	–
EU	19.8	14.6	–	–
New Zealand	0.7	2.4	–	–
Finland	1.4	0.1	–	–
Mexico	2.6	7.8	1987	10
South Korea	0.4	2.6	1985	39
Brazil	0.1	5.5	1988	55
Argentina		6.3	1991	16
South Africa		6.1	1994	20
India		4.2	1992	16
Turkey		1.6	1993	15
Colombia		1.2	1991	4
Poland		1.2	1991	43
Israel		1.1	1993	13
Indonesia		1.1	1996	31
Venezuela		1.0	1993	18
Peru		0.9	1994	1
Malaysia		0.7	1995	32
Philippines		0.7	1993	9
Chile		0.5	1993	5
Egypt		0.3	1997	7
Thailand		0.3	1993	35
Costa Rica		0.3	1996	1
Trin–Tobago		0.2	1996	3
Japan		0.2	1991	164
Czechoslovakia		0.1	1998	69
Nicaragua		0.1	1998	2
Panama		0.1	1998	0
Singapore		0.1	1994	34
Ecuador		0.1	1998	2
Guatemala		0.1	1996	0
Total #	1,401	2,113		

Table 12.1 provides some additional information on AD activity in the past two decades. In Table 12.1, we see that only 9 GATT countries filed AD cases during the 1980s while 39 did so during the 1990s. In addition, we can see that the group of new users first included countries such as Mexico and South Korea, and has since grown to include South Africa, Argentina, Indonesia, and Poland, among others. In fact, over

the last 10 years the set of new user countries has steadily grown to encompass nearly 30 countries, a third of whom have begun to use AD only in the second half of the 1990s.

Together, Figure 12.1 and Table 12.1 show that these new users are now eclipsing the traditional users in terms of the number of cases filed annually. In every year since 1993, new users filed more AD actions than traditional users. In 1996, for instance, new users accounted for nearly 70 percent of all AD cases. Overall, during the entire 1990s new users accounted for more than 40 percent of the total number of cases filed.

To a large extent, the data suggest that the table has turned for the traditional users. The United States, the world's most prolific user of AD law, is now the second most *investigated* country, trailing only the People's Republic of China. Countries such as Mexico, Argentina, South Africa, Brazil, and India are now among the heaviest users of AD.

Table 12.1 also shows that virtually none of the new users were strangers to AD actions before filing their own cases. Only 2 of the 27 countries that began using AD consistently after 1985 had never had AD cases filed against any domestic interests. Well over half of the countries had been named in at least 13 AD cases before adopting their own statutes. Such striking trends certainly raise the specter that countries are using AD law for reasons other than punishing unfair trade.

2.3 Schools of Thought Regarding Antidumping

Broadly speaking, positions toward AD fall into three main camps. In the first camp are those who believe that AD simply protects against unfair trade and hence that the increase in AD actions reflects an increase in unfair trade. For example, McDonald (1998) and Mastel (1998) claim that the increased use of AD actions is largely due to liberalization and not an indication that AD is "out of control." Their argument is that as traditional trade barriers are lowered previously benign or irrelevant unfair pricing behavior becomes injurious and must be disciplined.

In the second camp are those who argue that AD facilitates trade liberalization by allowing countries to raise tariffs for those industries that are more significantly injured by imports than trade negotiators anticipated (Corden, 1974; Bagwell and Staiger, 1990). From this perspective, without the ability to resort to "special protection" countries could never have negotiated the wide-ranging reductions in tariff rates such as those achieved during the Tokyo and Uruguay Rounds of GATT. In other words, AD acts as a de facto safeguard mechanism.[1] Given the greater degree of economic interdependence and the increased volatility of exchange rates and trade volumes now observed worldwide, the increase in AD activity may simply reflect the greater need for *ex post* flexibility with respect to tariffs (Fischer and Prusa, 2003).

Finally, in the third camp are those who believe that AD is essentially a strategic or protectionist policy. For instance, it has been shown that AD promotes collusion by serving as a punishment device (Staiger and Wolak, 1991) and by coordinating pricing and quantity decisions (Prusa, 1992). There is also ample evidence that AD protection is most often granted to politically powerful industries (Finger, Hall, and Nelson 1982; Hansen and Prusa, 1995, 1996).

Whether the growth of AD activity raises or lowers welfare depends on what perspective is most accurate. From the perspective of the first two camps AD measures raise welfare; therefore, the growth of AD activity reflects how vital it is to the liberal trading system. On the other hand, from the perspective of the third camp AD is a loophole in the GATT/WTO rules; consequently, the growth of AD activity is a threat to free trade.

In Section 3, we present a more comprehensive review of the various theories used to explain AD behavior. We then proceed, in Section 4, to test the hypotheses generated by these theories.

3 Theory

In its purest form, an AD action should occur only after an incident of dumping. There are problems with this conclusion, however, due in large part to the manner in which AD statutes are written. Countries generally have significant discretion in the use of AD law, and implementation by the authorities is often poor. Thus, countries and individual industries within countries have learned that they can use the laws to their advantage in a variety of ways. As discussed above, we can classify explanations for AD actions into three broad schools of thought: (i) AD as a response to unfair trade; (ii) AD as special or safeguard protection; and (iii) AD as a strategic weapon, to punish those who have used it or to deter others from abusing it. These three “camps” serve to motivate the hypotheses to be tested.

3.1 Unfair Trade

AD actions may be filed in response to unfair trade practices by importers. The standard theoretical explanation for AD is based on the existence of dumped imports, goods that are sold either at a price below that set in the importer’s domestic market or at a price below the importer’s cost of production, implying that goods are not sold at “fair value.”

Brander and Krugman (1983) provide a formal theoretical treatment of dumping as international price discrimination.[2] Their model shows that oligopolistic firms in segmented markets have profit-maximizing incentives to sell into each other’s markets. Further, equilibrium export prices for each firm can fall below their domestic prices making their foreign sales qualify as dumped.

The first formal model of dumping as pricing below production cost appears in Ethier (1982). His model shows that factor market considerations in unfavorable periods can lead to below-cost export pricing in the international equilibrium. In both this case and that analyzed by Brander and Krugman (1983), domestic producers may experience losses due to the existence of dumped imports. Arguably, these are the losses that are the concern of international AD statutes. The GATT Code provides for the imposition of AD duties in cases in which dumped imports are found to have caused such material injury to domestic firms.

Given the level of detail at which we were able to construct our data set, we are constrained to use country-level data. Therefore, we cannot determine whether there

have been actual instances of dumping or unfair trade practices. Instead, we will use indirect measures of unfair practices. For example, one could expect that exporting at unfairly low prices would result in large import volumes and/or large increases in imports and thus that AD cases would be directed at countries with such trends.[3]

3.2 Special Protection

It has long been argued that overall trade liberalization can only be achieved if countries have the ability to suspend their obligations for those industries that are more significantly injured by imports than trade negotiators anticipated. This view was formalized by Bagwell and Staiger (1990), who developed a game-theoretic model in which "special protection" arises as a short-term cooperative remedy for changes in underlying trade flows.

Their basic structure is that of a repeated Prisoners' Dilemma game in which cooperative behavior is defined as imposing low tariff rates, while high tariff rates represent defection. In a repeated game, players could solve the dilemma by using contingent strategies, like TFT, that incorporate punishment schemes to handle instances of defection from the cooperative outcome, or by the creation of formal agreements that include third-party punishment mechanisms, such as GATT.[4] However, as Bagwell and Staiger (1990) argue, these schemes cannot account for the growing incidence of the use of "special protection," such as escape clause actions, VERs and AD.

To better explain the existence of special protection, Bagwell and Staiger use a simple two-country partial equilibrium model of trade in one sector with a Prisoners' Dilemma payoff structure. They then assume that each country's output is subject to (observed) random shocks where positive shocks lead to higher volumes of trade. In the repeated game, cooperation can be sustained by means of a credible threat to revert forever to the static Nash equilibrium in the event of any defection. Crucially, Bagwell and Staiger show that the level of cooperation, that is, the value of the cooperative tariff, depends on the volume of trade and on the variance of the shocks (i.e., the size of the import "surge"). They show that the cooperative tariff rate is increasing in both import size and variance. In other words, Bagwell and Staiger formalize Corden's (1974) argument that special protection may be a tacitly agreed method to maintain cooperation in volatile trade periods.

From an empirical perspective, the "special protection" hypothesis implies two testable outcomes. First, AD cases are more likely to be filed against a country's largest trading partners. And, second, AD cases should be more likely to be filed against those trading partners with recent surges in import volume.

3.3 Strategic Use of Antidumping – Punishment and/or Deterrence

According to Bagwell and Staiger, special protection is a (symmetric) cooperative response to a changing trade environment that allows countries to maintain cooperation in times of volatile trade volumes. Within their model, the changes in tariff levels are

equilibrium responses to changes in the size or variance of imports. Yet, there are reasons to believe that there may be additional factors that should be considered when studying AD filing behavior. For example, Finger (1993) argues that the countries that use AD form a type of "club," in that they tend to apply AD against one another rather than against non-club members. Similarly, Prusa (2001) argues that many countries appear to file AD actions against countries that have previously investigated them, suggesting a type of TFT behavior.[5]

Both of these arguments suggest that countries may be using AD in a strategic fashion to punish defectors (or perhaps to deter defection). Without any further elaboration, however, both patterns are inconsistent with AD as protection against unfair trade and are outside the Bagwell–Staiger special protection framework. Here we suggest two extensions to the Bagwell–Staiger model that would generate the club effect and/or TFT filing in equilibrium: (i) imperfect monitoring of demand shocks and (ii) political economy incentives to misuse AD. A third possibility discussed below is that the punishment strategies are in fact evidence that countries are "out of equilibrium" (in the Bagwell–Staiger notion of cooperative equilibrium).

One extension to the Bagwell–Staiger framework would broaden the basic Prisoners' Dilemma game to allow for the demand shock to be unobserved. As shown by Green and Porter (1984) this type of unobservable shock can induce a cartel to enter into a "punishment" phase even if no player has deviated from the cooperative outcome. Applying the Green and Porter logic to our application generates filing patterns consistent with punishment strategies. In particular, we could consider AD users as members of a type of cartel or *club*, in which monitoring the behavior of each club member is not possible. Then, when club members experience trade-related economic difficulties, they respond by filing AD cases, whether or not the named countries have actually "dumped."

A pattern of Green and Porter reversion to high duties could give rise to either direct punishment or indirect punishment. If, for example, country A's use of AD against Country B leads Country B to file its own AD case against Country A, then we will observe direct punishment. For convenience we will refer to this type of behavior as "TFT" punishment even though we are intending a somewhat broader than usual interpretation of the term.[6] On the other hand, if country A's use of AD against Country B leads other countries (country C, D, E, etc.) to file AD cases against Country A, then we will observe indirect punishment, or what we will refer to as "club effect" punishment. Finger (1993) was the first (to our knowledge) to suggest this type of behavior and we follow his lead in referring to it as a club effect.

A second potential extension to the Bagwell–Staiger model would allow political motivations for offering protection. For instance, in the Grossman and Helpman (1994) framework politicians care not just about national welfare but also about the financial contributions they raise. They show that politicians will offer more protection than is optimal (from a national point of view) because doing so allows them to raise additional funds to finance future campaigns.

Applied to our model, Grossman–Helpman's political economy story would suggest that politics would lead governments to apply AD more aggressively than simple economics would predict. In order to deter politicians from abusing this flexibility, trading partners may want to raise the cost to applying AD. They do so by filing their

own AD actions in response to seeing others using AD. In equilibrium, one would expect that politicians would design AD so only industries with a high willingness to pay for protection will file AD actions, since doing so induces other countries to respond with their own AD actions. As with the Green and Porter extension, this extension would give rise to observing AD cases being filed as indirect and/or direct punishment.

Finally, it may be the case that worldwide AD use is not in "equilibrium." That is, the heavy use of AD law during the late 1970s and early 1980s may have been too aggressive, and rather than maintaining the cooperative tariff level as implied by Bagwell and Staiger, such heavy use may have been perceived as a defection to the high protection regime. If this is the case, then recent AD actions might be "out of equilibrium" responses to the earlier defection. As with the above two scenarios, depending on the punishment strategy such defections could give rise to either direct or indirect punishment.

4 Estimating the Strategic Motives for AD Filings

In order to investigate the motives for filing AD cases, we have compiled a database of all AD cases filed from 1980 through 1998 by GATT/WTO members. The filing trends were depicted in Figure 12.1 and reported in Table 12.1 as well as discussed briefly in Section 2.2. The most salient facts are as follows: (i) over 3,500 AD cases were filed worldwide over the 19-year period; (ii) two-third of the cases were filed by traditional users, although the percentage has been falling in recent years; (iii) only 9 countries filed AD cases during the 1980s, but 29 countries filed AD cases during the 1990s, suggesting that many countries are beginning to catch the AD bug; and (iv) about one-fifth of the cases were filed against nonmarket economies.

This final point is noteworthy since the rules for determining the existence of dumping are quite different when the affected country is a nonmarket economy (Boltuck and Litan, 1991). As a result, the motivation for filing against a nonmarket economy is likely to be quite different than that for filing against a market economy. For this reason, we drop cases against nonmarket economies from our analysis.[7]

Another issue that is worth mentioning is that our data is based on aggregate filing data. In most cases and in most countries, petitions are initiated by firms, trade unions, or trade associations. Our discussion of clubs and retaliation was couched in language that suggested that governments make the decisions to file, and effectively, it may be that they are, in the sense that governments communicate with the pertinent firms about the possibility of filing. On the other hand, it may be that firms themselves retaliate. See Maur (1998) for evidence of retaliation at the industry level. Regardless of one's preferred interpretation, we are interested in whether there is any evidence of retaliatory filings.

As discussed in the preceding section, there are four main hypotheses that we examine in our analysis of AD filing behavior. These hypotheses are: big supplier, big change in imports, TFT, and club effect. The first two are motivated by the Bagwell–Staiger model of special protection and are consistent with the view that AD actions are used to prevent unfair trade. The last two are outside the basic Bagwell–Staiger model and are consistent with the belief that strategic considerations influence AD actions.

The "big supplier" hypothesis is that AD cases are filed against a country's largest import suppliers (based on aggregate country data). The "big change" hypothesis is that AD cases are filed against suppliers who have the largest percentage change in imports. The TFT hypothesis is that a country is more likely to file AD cases against those suppliers who previously have filed an AD case against it. For example, under the TFT hypothesis if the United States and Australia have filed cases against Mexico in the past, then Mexico will be more likely to file cases against the United States and Australia than against, say, South Korea and Venezuela. The club hypothesis is that a country is more likely to file AD cases against countries that also use AD, regardless of whether it has been directly named in the past. Continuing with the above example, if South Korea has used AD against countries other than Mexico, under the club hypothesis Mexico would be more likely to name South Korea than Uruguay (i.e., a country that had not used AD at all).

Figure 12.2 gives a broad picture of how the four hypotheses are reflected in the data. Overall, 47 percent of AD cases are filed in a manner consistent with TFT and 62 percent are consistent with the club effect. Thus, it appears that these two strategic hypotheses find substantial support in the data.

Economic incentives also receive strong support. For instance, overall 99 percent of AD cases are filed against big suppliers, where we have defined big suppliers as those countries with import shares greater than that of the median supplier. The big change hypothesis also seems reasonably consistent with the data, as 54 percent of AD cases are filed against suppliers whose percentage change in imports is greater than that of the supplier with the median percentage change.

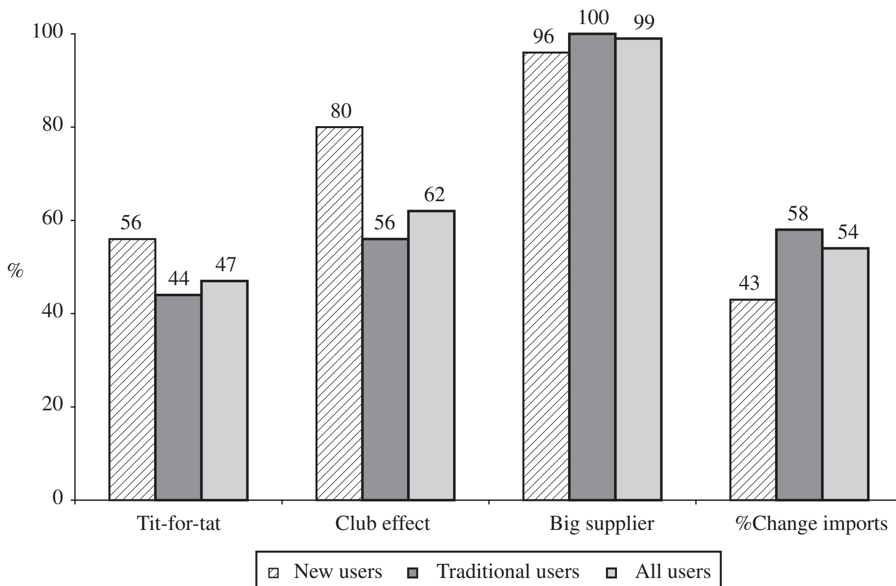

Figure 12.2 Antidumping actions consistent with alternative hypotheses

There also appear to be substantial differences in the patterns between new and traditional users. For instance, the filing patterns of new users seem to more strongly support the strategic hypotheses, TFT and club, than those of traditional users. While new users may indeed use AD more strategically than do traditional users, this may be due to the fact that a different proportion of countries supplying new and traditional users were AD users. For instance, if all of a new user's suppliers were AD users, then we would report 100 percent of this country's cases were consistent with the club effect.

In order to identify the individual effects of each hypothesis, we proceed with parametric estimation of the AD filing decision.[8] To do so, we model each country's decision to file an AD action against each trading partner in each year as a function of the economic and strategic incentives. Let F_{ijt} be a binary variable that denotes whether country i files an AD case against country j in year t. We can then express the filing decision as

$$F_{ijt} = f(\text{TFT}_{ijt}, \text{Club}_{jt}, \text{Imports}_{ijt}, \%\Delta\text{Imports}_{ijt}, \text{All_Imports}_{it}),$$

where TFT_{ijt} denotes whether country j has ever filed an AD case against country i prior to year t, Club_{jt} denotes whether country j is a member of the "AD club" prior to year t, Imports_{ijt} denotes imports from j to i at time t, and $\%\Delta\text{Imports}_{ijt}$ denotes the percentage change in imports from j to i between $t-1$ and t. The final variable, All_Imports_{it}, is not directly related to the proposed hypotheses but rather captures the intuition that overall filing patterns might systematically vary across users. That is, everything else equal, some users may use AD more aggressively than other users. A natural way countries might differ in their use of AD is in their exposure to imports. Thus, the variable All_Imports_{it} denotes country i's import in year t from all sources.[9]

We estimate this decision function using probit.[10] In Tables 12.2–12.5 we present the parameter estimates for the model. A different specification is presented in each table. We begin with the simplest specification (Table 12.2) and move to the most general specification (Table 12.5). In all specifications we measure the strategic hypotheses (TFT and club) using simple dummy variables. In addition, we allow the strategic incentives to vary by the level of imports; both linear and quadratic specifications are investigated. Moreover, in specifications B and D (Tables 12.3 and 12.5) we allow the hypotheses to have different impacts for new or traditional users. In all specifications we estimate (but do not report) year effects.

Finally, we note that in each table we report two sets of results, one based on data for all years and the other based on data only for each country's first three years of AD use. The restricted dataset, including information only on each country's first three years of AD activity, was constructed in order to control for potential endogeneity problems with respect to our TFT and club effect tests. The following example clarifies our concern.

Suppose that in 1981 the United States filed an AD case against Mexico. Next assume that Mexico filed an AD case against the United States in 1983. According to our definitions we would say Mexico's filing was consistent with TFT (and also the club effect). Now, suppose that the United States filed a case against Mexico in 1986. Once again, according to our definitions we would say this filing is consistent with TFT (and the club effect). But, this TFT finding was the result of Mexico responding

Table 12.2 Probit regression: Specification A

	Linear All years	**Quadratic All years**	**Linear First 3 years**	**Quadratic First 3 years**
TFT	−0.103	−0.185	−0.268	−0.323
	[0.069]	[0.073]*	[0.134]*	[0.153]*
Club effect	1.422	1.433	1.501	1.491
	[0.058]**	[0.060]**	[0.107]**	[0.116]**
TFT × imports	−15.362	−17.262	−5.48	−31.604
	[5.163]**	[7.883]*	[21.675]	[54.846]
Club effect × imports	−32.722	−92.264	−63.186	−139.455
	[5.878]**	[9.596]**	[23.165]**	[55.480]*
TFT × imports_{ijt}		82.054		1,889.74
		[40.986]*		[2,015.892]
Club effect × imports_{ijt}		1,177.68		2,071.59
		[101.478]**		[2,089.293]
Imports	60.577	139.025	103.205	253.893
	[3.418]**	[6.570]**	[10.952]**	[22.824]**
Imports_{ijt}		−1,336.18		−4,904.17
		[94.662]**		[694.949]**
%Δimports	−0.013	−0.012	−0.044	−0.041
	[0.005]*	[0.005]*	[0.020]*	[0.020]*
%$\Delta\text{imports}_{ijt}$		0		0
		[0.000]*		[0.000]*
Imports: all suppliers	−0.037	−0.419	−1.13	−3.236
	[0.031]	[0.075]**	[0.245]**	[0.791]**
Imports: all suppliers_{ijt}		0.106		2.482
		[0.022]**		[1.095]*
Constant	−1.829	−1.846	−1.629	−1.553
	[0.106]**	[0.111]**	[0.124]**	[0.131]**
Observations	20,684	20,684	7,088	7,088

*Standard errors in brackets; *, significant at 5%; **, significant at 1%.*

to the earlier filing of the United States, and hence is endogenous. In this scenario, our measures of the TFT and club effects would both be distorted by such endogeneity. In an attempt to mitigate this potential problem, we chose to report statistics on only the first three years of each country's use of AD in addition to reporting results for the full dataset.

Given that the results do differ somewhat across the two datasets we will need to keep in mind that the results based on the data from the "first three years" are probably best interpreted as capturing the incentives during a country's initial years of using AD. The results based on the "all years" data should be interpreted as capturing a country's long-run strategic use of AD.

Our most straightforward specification is presented in Table 12.2. In this specification we constrain the strategic and economic motives to be the same for new and traditional users. The nonlinear nature of the probit estimates makes it difficult to easily interpret the

Table 12.3 Probit regression: Specification B

	Linear All years	**Quadratic All years**	**Linear First 3 years**	**Quadratic First 3 years**
Traditional user × TFT	0.092	0.024	0.868	0.893
	[0.110]	[0.115]	[0.403]*	[0.409]*
Traditional user × club effect	1.633	1.672	1.121	1.136
	[0.080]**	[0.082]**	[0.300]**	[0.304]**
New user × TFT	−0.114	−0.175	−0.369	−0.467
	[0.087]	[0.090]	[0.141]**	[0.161]**
New user × club effect	1.224	1.221	1.527	1.515
	[0.073]**	[0.075]**	[0.113]**	[0.122]**
TFT × imports	−16.547	−19.121	−7.943	−21.276
	[5.248]**	[7.994]*	[21.968]	[55.040]
Club effect × imports	−35.93	−99.2	−66.537	−147.309
	[5.932]**	[9.667]**	[23.211]**	[55.532]**
TFT × $imports_{ijt}$		85.269		1,536.12
		[41.247]*		[2,018.725]
Club effect × $imports_{ijt}$		1,236.86		2,227.65
		[102.304]**		[2,089.915]
Imports	62.923	144.571	104.597	257.372
	[3.452]**	[6.657]**	[10.991]**	[22.946]**
$Imports_{ijt}$		−1,391.50		−4,978.62
		[95.583]**		[699.143]**
%Δimports	−0.013	−0.012	−0.044	−0.042
	[0.005]**	[0.005]*	[0.020]*	[0.020]*
%$\Delta imports_{ijt}$		0		0
		[0.000]*		[0.000]*
Imports: all suppliers	−0.111	−0.555	−1.202	−3.114
	[0.034]**	[0.079]**	[0.251]**	[0.794]**
Imports: all $suppliers_{ijt}$		0.124		2.142
		[0.022]**		[1.107]
Constant	−1.825	−1.833	−1.631	−1.565
	[0.107]**	[0.113]**	[0.125]**	[0.132]**
Observations	20,684	20,684	7,088	7,088

*Standard errors in brackets; *, significant at 5%; **, significant at 1%.*

economic significance of the coefficients. Thus, to address the economic significance, we also calculate the "change in probability" implied by the estimates and report these results within our discussion below.[11]

In reviewing the results from our simplest specification, Specification A, as reported in Table 12.2, we see first that almost all included variables are statistically significant. This result holds true for the tests run on both datasets and for both the linear and quadratic forms of our model. Importantly, there is support here for all four of our hypotheses, both strategic and economic. The largest observed effect is from the "big supplier" hypothesis but both strategic hypotheses receive considerable support in the

Table 12.4 Probit regression: Specification C

	Linear All years	**Quadratic All years**	**Linear First 3 years**	**Quadratic First 3 years**
TFT	−0.105	−0.169	−0.239	−0.531
	[0.069]	[0.074]*	[0.136]	[0.167]**
Club effect	1.422	1.422	1.466	1.444
	[0.058]**	[0.060]**	[0.107]**	[0.117]**
TFT × imports	−15.377	−17.089	−16.553	77.365
	[5.179]**	[8.064]*	[22.458]	[63.465]
Club effect × imports	−32.605	−91.968	−32.839	−222.225
	[5.894]**	[9.679]**	[26.418]	[83.002]**
TFT × imports$_{ijt}$		91.023		−2,900.63
		[40.954]*		[2,573.625]
Club effect × imports$_{ijt}$		1,221.62		13,050.37
		[102.898]**		[6,149.566]*
Traditional user × imports	60.54	145.988	125.591	308.348
	[3.446]**	[6.936]**	[14.159]**	[32.491]**
Traditional user × imports$_{ijt}$		−1,411.83		−5,955.52
		[97.939]**		[915.737]**
Traditional user × %Δimports	−0.051	−0.052	−0.086	0.68
	[0.022]*	[0.024]*	[0.060]	[0.229]**
Traditional user × %Δimports$_{ijt}$		0		−0.966
		[0.000]		[0.295]**
New user × imports	60.309	128.926	72.991	336.01
	[4.303]**	[7.678]**	[16.780]**	[67.138]**
New user × imports$_{ijt}$		−1,392.98		−15,861.96
		[103.353]**		[5,857.828]**
New user × %Δimports	−0.008	−0.008	−0.036	−0.034
	[0.005]	[0.005]	[0.021]	[0.039]
New user × %Δimports$_{ijt}$		0		0
		[0.000]		[0.002]
Imports: all suppliers	−0.032	−0.461	−1.28	−3.002
	[0.032]	[0.077]**	[0.264]**	[0.817]**
Imports: all suppliers$_{ijt}$		0.108		1.536
		[0.022]**		[1.180]
Constant	−1.825	−1.848	−1.656	−1.554
	[0.106]**	[0.112]**	[0.128]**	[0.143]**
Observations	20,684	20,684	7,088	7,088

*Standard errors in brackets; *, significant at 5%; **, significant at 1%.*

data. Curiously, the "import surge" hypothesis, measured by the impact of changes in import volumes, appears to have a significant negative (albeit small) effect on AD use. This result appears in all specifications of our model and suggests that our hypothesis related to import surges is not upheld in the data.[12]

Looking more closely at the impact of the strategic incentives, we note that the club effect is relatively large. Across the four regressions reported in Table 12.2, AD club

Table 12.5 Probit regression: Specification D

	Linear All years	**Quadratic All years**	**Linear First 3 years**	**Quadratic First 3 years**
Traditional user × TFT	0.101	0.038	0.506	0.442
	[0.111]	[0.122]	[0.714]	[0.774]
Traditional user × club effect	1.667	1.718	0.727	0.634
	[0.080]**	[0.084]**	[0.405]	[0.470]
Traditional user × TFT × imports	−12.293	−15.335	−450.106	−623.843
	[5.401]*	[9.526]	[457.990]	[865.514]
Traditional user × club effect × imports	−43.87	−126.551	515.157	600.45
	[6.124]**	[11.174]**	[441.018]	[852.858]
Traditional user × TFT × $imports_{ijt}$		66.349		71,705.43
		[47.814]		[206,287.278]
Traditional user × club effect × $imports_{ijt}$		1,504.10		−69,209.65
		[112.670]**		[206,253.146]
Traditional user × imports	64.928	166.598	120.615	293.154
	[3.656]**	[7.711]**	[14.718]**	[32.954]**
Traditional user × $imports_{ijt}$		−1,633.29		−5,549.37
		[105.370]**		[921.847]**
Traditional user × %Δimports	−0.061	−0.064	−0.085	0.696
	[0.025]*	[0.027]*	[0.059]	[0.230]**
Traditional user × %Δ$imports_{ijt}$		0		−0.969
		[0.000]*		[0.294]**
New user × TFT	−0.113	−0.209	−0.332	−0.611
	[0.094]	[0.102]*	[0.143]*	[0.173]**
New user × club effect	1.154	1.172	1.514	1.526
	[0.078]**	[0.082]**	[0.113]**	[0.124]**
New user × TFT × imports	−35.552	−21.154	−12.28	82.456
	[12.791]**	[27.388]	[22.665]	[63.964]
New user × club effect × imports	11.861	−152.594	−48.917	−334.968
	[17.031]	[53.528]**	[26.963]	[93.383]**
New user × TFT × $imports_{ijt}$		13.03		−2,873.52
		[769.768]		[2,534.934]
New user × club effect × $imports_{ijt}$		10,005.76		20,795.18
		[3,687.251]**		[7,283.504]**
New user × imports	41.592	215.72	84.816	431.932
	[11.747]**	[46.484]**	[17.377]**	[78.343]**
New user × $imports_{ijt}$		−10,244.97		−23,079.17
		[3,606.287]**		[7,009.086]**
New user × %Δimports	−0.008	−0.008	−0.037	−0.034
	[0.005]	[0.005]	[0.021]	[0.039]
New user × %Δ$imports_{ijt}$		0		0
		[0.000]		[0.002]
Imports: all suppliers	−0.107	−0.622	−1.331	−3.043
	[0.035]**	[0.082]**	[0.271]**	[0.823]**
Imports: all $suppliers_{ijt}$		0.135		1.559
		[0.023]**		[1.193]
Constant	−1.824	−1.849	−1.643	−1.532
	[0.107]**	[0.114]**	[0.128]**	[0.142]**
Observations	20,684	20,684	7,088	7,088

*Standard errors in brackets; *, significant at 5%; **, significant at 1%.*

members are about 15–20 percent more likely than non-club members to be investigated by other countries.[13] Second, and surprisingly, the TFT effect appears with the wrong sign in every regression in this specification. Our hypotheses above suggested that TFT retaliation incentives would increase the likelihood of AD actions and we therefore expected a positive coefficient on this variable. We note, however, that the estimated TFT impact is quite small, lowering the probability of being named by only about 1 percent.

The negative sign on the TFT variable is sufficiently curious to warrant further investigation. In particular, it is possible that the unanticipated sign results from the fact that Specification A does not allow new and traditional users to differ in their strategic and economic incentives to use AD. If the two types of users do differ systematically in their filing patterns, the parameters reported in Table 12.2 would reflect the weighted average of the traditional and new users impact. Thus, the results from Specification A may be significantly skewed and may not accurately capture either the strategic or economic incentives for AD use.

In Specifications B, C, and D we address this issue by allowing for separate new and traditional user estimates. The three specifications incorporate different levels of variation across the user types. In particular, in Specification B (Table 12.3) we allow the strategic effects to vary by type of user but we hold the Bagwell–Staiger economic effects common across the users. In Specification C we do the opposite, allowing the Bagwell–Staiger economic effects to vary by type of user but holding the strategic effects common across users (Table 12.4). Finally, in Specification D we allow all the parameters to vary by type of user (Table 12.5).

Given our interest in strategic filing issues, specification B (reported in Table 12.3) is our preferred specification. As in Specification A, almost all of the parameters are statistically significant. In particular, there is statistical support for all four AD use hypotheses. We also see, as in Specification A, that there is strong support for the "big supplier" hypothesis and that the effect of this parameter on one's chance of being named in an AD case is quite large. At the same time, the "import surge" hypothesis receives much weaker support and appears with the wrong sign, as discussed above.

Importantly, distinguishing between new and traditional users in this specification allows us to refine our understanding of TFT and club incentives. In particular, the estimates indicate that the TFT incentive for traditional users ranges from moderate (in the all years dataset) to strong (in the first three years dataset). Traditional users are between 0.5 percent and 5 percent more likely to file AD cases against countries that have filed against them in the past. By contrast, we find that the TFT effect of new users is consistently negative and strongly significant in the dataset for the first three years. The change in probability in this case is small, but the results indicate that new users are about 0.5 percent *less* likely to file against countries that had filed against them (in the first three years or overall). These contrasting results suggest that our concerns about the results in Specification A were justified and that it is important to estimate separate effects for traditional and new users.

In order to investigate further this variation across user types in the incentives to use AD, we look at the extent to which strategic incentives vary both by user type and by level of imports. We find that the TFT and club effects vary considerably with the size

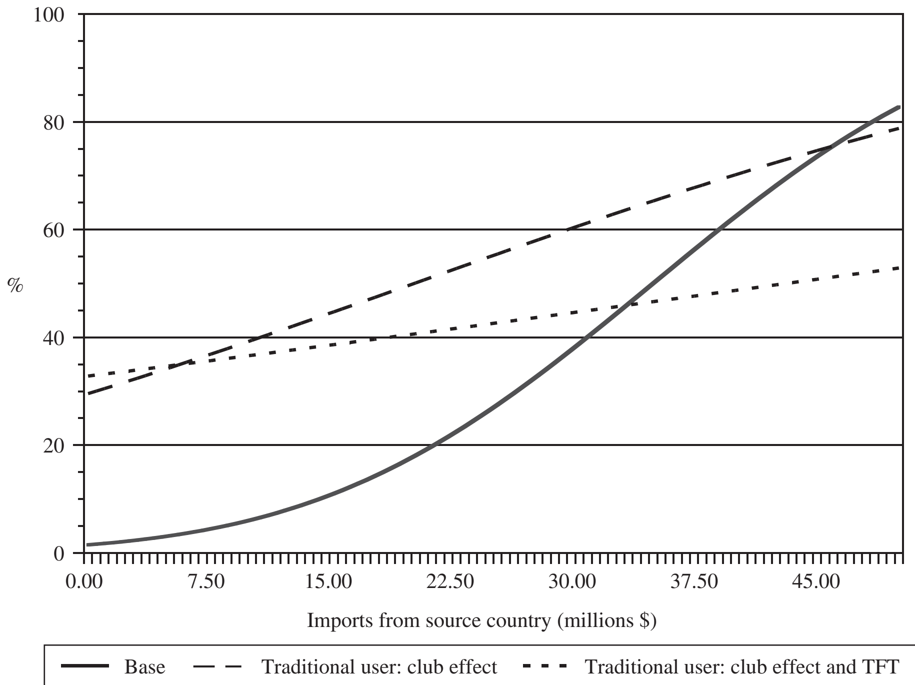

Figure 12.3 Traditional users (Specification B, all years)

of the supplier, for each type of user. To illustrate this finding, we plot the probability of naming a country in an AD case against the dollar value of imports from the source country. We do this for both traditional and new users for the linear version of our model for both the "all years" and "first three years" datasets in Figures 12.3–12.6. For example, Figure 12.3 shows the estimated filing behavior of traditional users using the linear specification and the "all years" dataset. In Figure 12.3, we plot separately the incentive to name a country that is not a club member ("base" curve), a country that is a club member, and a country that is a club member that has also used AD against a traditional user.

Precisely, as one would expect, we see in all of the plots in Figures 12.3–12.6 that the probability of naming a country is increasing in imports. Exactly how the probability varies depends, however, on the different strategic incentives. It also depends on the type of user under consideration.

For traditional users over all years (Figure 12.3), we see that the probability of naming a non-club member (the base curve) is essentially zero for low import values, but increases quite sharply as import values rise. Being a club member, however, significantly increases the likelihood of being named at small import values. Interestingly, the club effect impact dissipates and eventually becomes negative at large import values. In other words, it appears that being a club member is especially costly if the country is small. This finding suggests that the smaller a country is, the more likely it is that *other* AD users punish it with AD protection. The TFT effect is larger than the club effect

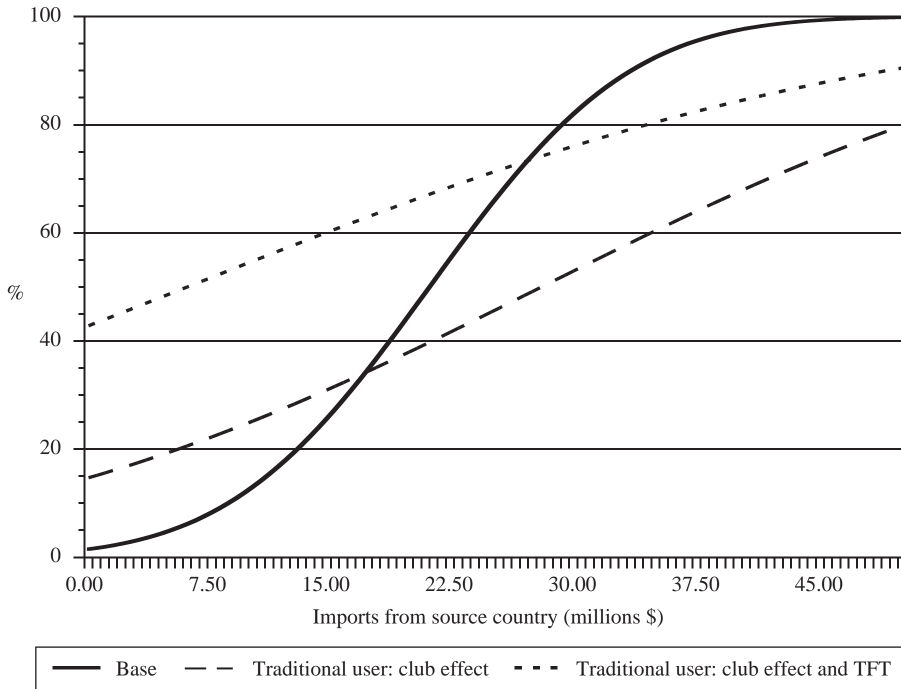

Figure 12.4 Traditional users (Specification B, first three years)

only for very small import values and in general reduces the chance of being named. We interpret this result as likely indicating that TFT incentives may be correlated with country-specific variables that on average reduce the chance of being named.[14] Overall, the "all years" results indicate reasonably strong support for the club hypothesis but only modest support for the TFT hypothesis for traditional users.

By contrast, when we look at the behavior of traditional users using the linear specification based on the "first three years" dataset we find quite strong support for both the club and TFT hypotheses (Figure 12.4). For small import volumes, club and TFT each increase the likelihood of being named by about 20 percentage points. To be precise, the TFT effect is economically significant, increasing the likelihood of being named upward of 30–40 percent (at low import values). As we saw in Figure 12.3, club and TFT both dampen the import size effect for large importers. Ultimately, beyond a critical size (about $27 million), TFT incentives do not appear important; similar club incentives become unimportant beyond about $17 million. Nevertheless, the estimates indicate that traditional users' filing patterns in this restricted sample are heavily influenced by strategic considerations.

In Figures 12.5 and 12.6 we look at the behavior of new users using the linear specification based on the "all years" and "first three years" datasets, respectively. For new users, the two figures tell very similar stories. For both datasets the club effect is economically large and significant, increasing the likelihood of being investigated by about 20 percent at low import volumes. As we saw was true for traditional users, this

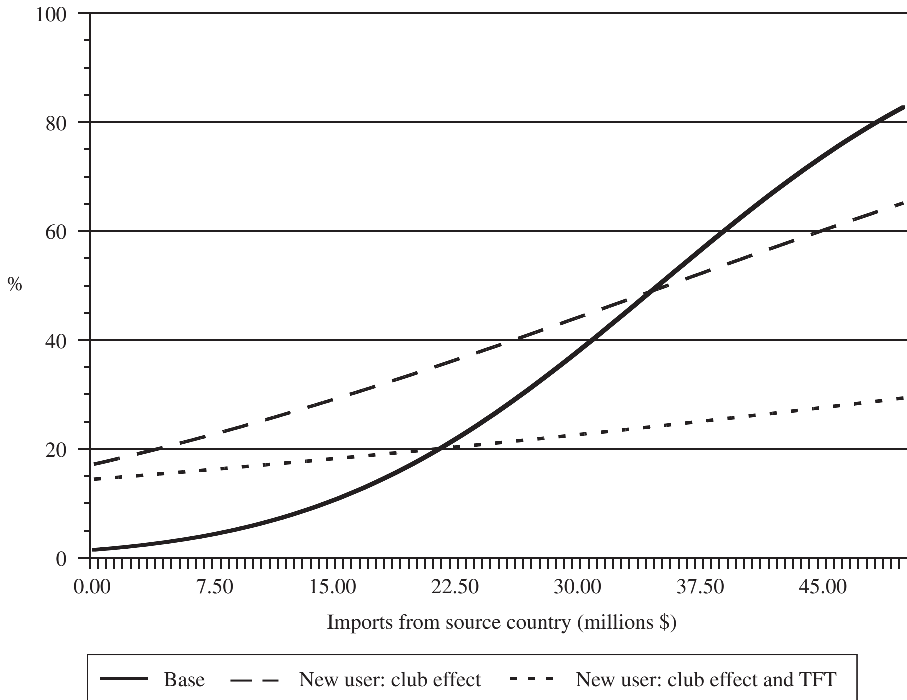

Figure 12.5 New users (Specification B, all years)

effect dissipates as import volume grows. For both datasets we also find TFT incentives to be negative and small for the new users.

Combining the insights of all four figures, we conclude that traditional and new users do have significantly different patterns of AD use. Thus, it appears that the probit results based on Specification A are misleading because they constrain a single parameter to capture the incentives of these two very different sets of users. New users appear to have a more pronounced club effect and traditional users appear to be more influenced by TFT (at least in the restricted sample). Finally, given the consistency of the results across new and traditional users and across datasets, we believe that strategic considerations are outweighed by economic factors in motivating AD cases filed against big suppliers.

Our third specification of the model, Specification C (reported in Table 12.4), allows the parameters associated with the economic hypotheses to vary across users while constraining the parameters associated with the strategic hypotheses to be the same for both new and traditional users. As in the earlier specifications, we find considerable statistical significance for most variables across both types of users. The exception is that the percentage change in import parameters is insignificant for new users (and relatively small in magnitude for traditional users). Notice, however, that unlike the Specification B results in which we saw a clear difference between new and traditional users in the strategic incentives, there is very little difference here between the estimated impact of economic factors for the two types of users.

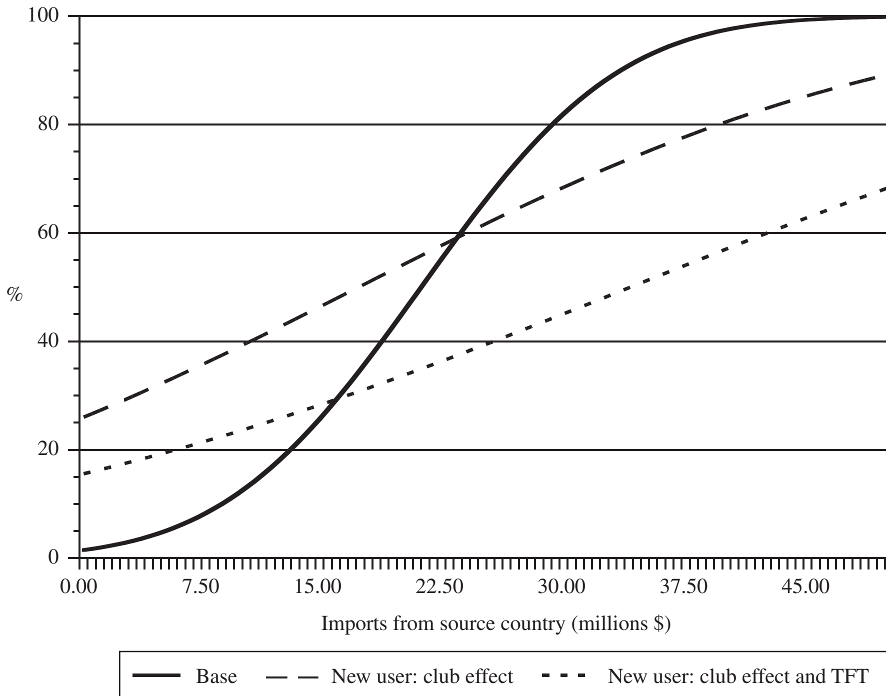

Figure 12.6 New users (Specification B, first three years)

Finally, in Specification D (Table 12.5) we allow both strategic and economic hypotheses to vary by user. This is the most flexible specification, but we are a bit concerned that there is insufficient variation among the exogenous variables. Thus, we may not be meaningfully identifying the separate strategic and economic incentives. We note that the TFT impact estimated here is very similar to that estimated in Specification B (Table 12.3). Namely, traditional users' filing patterns are consistently estimated to indicate the presence of positive club and TFT incentives. Both effects increase the likelihood that traditional users will file AD actions. New users demonstrate a strong club effect, but a very small TFT incentive. In other words, the club effect is more pronounced for new users, but offsetting this stronger club effect is a *negative* TFT effect.

5 Concluding Comments

The worldwide upsurge in AD use in the past two decades has begun to receive considerable attention. Concerns have arisen in part because the increased AD activity is evident not only in the number of cases filed annually but also in the range of countries now actively using AD policies. With the intensified use of this WTO-legal protection comes a need to identify the motives underlying the changes in patterns of behavior.

In particular, it is increasingly difficult to argue that the changes in AD use reflect only an increase in unfair trading behavior that has taken hold since 1980.

Our findings suggest that AD use is *not* just about protecting domestic industries from the unfair trade behavior of foreign entities. Although we find strong support for the economic hypothesis that AD actions are directed at big suppliers, we find far weaker evidence that AD actions are filed against importers with recent surges in their import volume. If AD filings were primarily aimed at combating unfair trading practices, we would expect that suppliers with large import surges would generally be the targets of those filings. Our evidence implies that this is not the case and that there are other incentives that play a significant role in national use of AD policy.[15]

Following recent literature that suggests that strategic, particularly retaliatory, incentives may play a role in AD use, we test for evidence of such behavior in case filings from the last two decades. In our results, we find evidence of a strong strategic component in the use of AD worldwide. Further, there seems to be ample evidence from our analysis that the traditional and new users of AD employ the AD laws in different ways. We observe strong strategic incentives for AD use particularly (but not exclusively) for traditional users, who show evidence of using AD for both club effect and TFT reasons. New users have a more significant club effect, but a negative TFT effect that mutes the club effect somewhat. In addition, our results suggest that strategic considerations are outweighed by economic factors in motivating AD cases filed against big suppliers.

Overall, our evidence is consistent with results beginning to be reported elsewhere in the literature. The econometric evidence provided here bolsters the results of the non-parametric tests performed in our earlier work (Prusa and Skeath, 2002), for example, and we find evidence of the type of retaliatory behavior that Blonigen and Bown (2001) argue may reduce the likelihood of US AD cases being filed against certain countries. Indeed, we provide evidence to support Messerlin and Reed's (1995) contention that AD measures may be used as a form of "disguised retaliation." Our findings reject the notion that the rise in AD activity is solely explained by an increase in unfair trading and support the view that strategic incentives play a critical role in the motivation of individual countries to file AD complaints against their suppliers.

Acknowledgments

We would like to thank Kyle Bagwell, Tom Pugel, and Bernie Yeung for helpful conversations. We also thank seminar participants at the University of Otago, Australia National University, Hong Kong University of Science and Technology, and NYU for comments and suggestions on an earlier draft. Work on this project was started while Skeath was a Commerce Divisional Visiting Fellow at the University of Otago. Financial support from the Class of '32 Social Sciences Fund at Wellesley College is gratefully acknowledged. We would also like to thank WTO Rules Division, and Jorge Miranda in particular, for making the WTO AD Measures Database available.

Notes

1 Given that the injury standard under AD is much lower than under safeguard protection and that most countries do not have a national interest clause in their AD statute, the fact that AD acts as de facto safeguard protection raises concerns. Nevertheless, this observation is consistent with the view that AD facilitates trade liberalization. A subtler question is whether the *abuse* of AD is necessary for trade liberalization.

2 This definition of dumping is that traditionally found in the literature. It can be traced at least to Viner who argued "for confining the term dumping to *price-discrimination between national markets.*" (1923, p. 3)
3 The injury requirement also makes it more likely that one would name countries with large imports or large changes in imports since such trends are usually taken as evidence of economic harm.
4 Papers by Dixit (1987), Jensen and Thursby (1984), Mayer (1981), and Riezman (1982) explore these issues.
5 Messerlin and Reed (1995, p. 1573) suggest a similar effect by which the AD Code, in encouraging the standardization and spread of AD law across WTO-member countries "may tempt countries to take AD measures as disguised retaliation."
6 True TFT strategies punish the original defector for only so long as she continues to defect. Instances of cooperation by the original defector are rewarded by cooperative play on the part of a true TFT player. Our TFT player reverts to the Nash equilibrium, and therefore defects, for all remaining periods following a defection.
7 We also debated whether to drop countries that are members of a preferential trade agreement (e.g., drop Mexico and Canada from US filing data). Although we were concerned about preferential treatment, we were unable to gather complete data on which FTA have special rules pertaining to AD actions against member countries. As a result, we do not control for FTA.
8 Note that Prusa and Skeath (2002) present nonparametric tests of the hypotheses presented above. The results there show variation in effects across user types and suggest that strategic motives for AD use cannot be dismissed.
9 Miranda, Torres, and Ruiz (1998) suggest that small importers use AD more aggressively.
10 We also estimated using logit and found that the results are not significantly different.
11 The "economic" impact estimates are not reported in the text for all regressions or variables. They are available from the authors upon request.
12 In later specifications (C and D, see below) we allow the economic impact parameters to vary according to type of user (traditional versus new). In those specifications, the "percentage change in imports" variable is still consistently estimated with a negative coefficient although the measured effect is smaller and less significant for new users than for traditional users.
13 As is standard practice, evaluation of the club and TFT effects are done at the mean values of all other exogenous variables. As we will discuss below, this average impact may not reflect the impact at other points in the distribution.
14 This result may also be capturing some of the "fear of retaliation" incentives identified in Blonigen and Bown (2001), although their work is based solely on US AD filing behavior.
15 In future research we are also interested in analyzing whether the patterns we find based on filing data are confirmed, if based on duties levied or on measures taken (i.e., an outcome-based measure of protection).

References

Bagwell, Kyle and Robert Staiger 1990: A theory of managed trade. *American Economic Review*, 80, 779–95.

Blonigen, Bruce A. and Chad P. Bown 2001: Antidumping and retaliation threats. NBER Working Paper No. 8576, October.

Boltuck, Richard and Robert E. Litan (eds.) 1991: *Down in the Dumps: Administration of the Unfair Trade Laws*. Washington, DC: Brookings Institute.

Brander, James A. and Paul Krugman 1983: A 'reciprocal dumping' model of international trade. *Journal of International Economics*, 15, 313–21.

Corden, W. Max 1974: *Trade Policy and Economic Welfare*. Oxford: Clarendon Press.

Dixit, Avinash K. 1987: Strategic aspects of trade policy. In Truman F. Bewley (ed.), *Advances in Economic Theory: Fifth World Congress*. New York: Cambridge University Press, 329–62.

Ethier, Wilfred J. 1982: Dumping. *Journal of Political Economy*, 90, 487–506.

Finger, J. Michael (ed.) 1993: *Antidumping*. Ann Arbor, MI: University of Michigan Press.

Finger, J. Michael and Kwok-Chiu Fung 1994: Will GATT enforcement control antidumping? *Journal of Economic Integration*, 9, 198–213.

Finger, J. Michael, H. Keith Hall, and Douglas R. Nelson 1982: The political economy of administered protection. *American Economic Review*, 72, 452–66.

Fischer, Ronald D. and Thomas J. Prusa 2003: Contingent protection as better insurance. *Review of International Economics* (forthcoming).

Green, Edward J. and Robert H. Porter 1984: Noncooperative collusion under imperfect price information. *Econometrica*, 52, 87–100.

Grossman, Gene and Elhanan Helpman 1994: Protection for sale. *American Economic Review*, 84, 833–50.

Hansen, Wendy and Thomas J. Prusa 1995: The road most traveled: the rise of title VII protection. *The World Economy*, 18, 295–313.

Hansen, Wendy and Thomas J. Prusa 1996: Cumulation and ITC decision-making: the sum of the parts is greater than the whole. *Economic Inquiry*, 34, 746–69.

Hufbauer, Gary 1999: Antidumping: a look at U.S. experience – lessons for Indonesia. Report to the Ministry of Industry and Trade, Republic of Indonesia, http://www.iie.com/papers/hufbauer0899-1.htm, August.

Hufbauer, Gary and Kimberly Elliott 1994: *Measuring the Costs of Protection in the United States*. Washington, DC: Institute for International Economics.

Jensen, Richard and Marie Thursby 1984: Free trade: two noncooperative approaches. Ohio State University Working Paper.

Mastel, Greg 1998: *Antidumping Laws and the U.S. Economy*. Armonk, NY: M.E. Sharpe.

Maur, Jean-Christophe 1998: Echoing antidumping cases: regulatory competitors, imitation and cascading protection. *World Competition*, 21, 51–84.

Mayer, Wolfgang 1981: Theoretical considerations on negotiated tariff adjustments. *Oxford Economic Papers*, 33, 135–53.

McDonald, Brian 1998: *The World Trading System: The Uruguay Round and Beyond*. New York: St. Martin's Press.

Messerlin, Patrick A. and Geoffrey Reed 1995: Antidumping policies in the United States and the European community. *The Economic Journal*, 105, 1565–75.

Miranda, Jorge, Raul A. Torres, and Mario Ruiz 1998: The international use of antidumping: 1987–1997. *Journal of World Trade*, 32, 5–71.

Prusa, Thomas J. 1992: Why are so many antidumping petitions withdrawn? *Journal of International Economics*, 33, 1–20.

Prusa, Thomas J. 2001: On the spread and impact of antidumping. *Canadian Journal of Economics*, 34, 591–611.

Prusa, Thomas J. and Susan Skeath 2002: The economic and strategic motives for antidumping filings. *Weltwirtschaftliches Archiv* (forthcoming).

Riezman, Raymond 1982: Tariff retaliation from a strategic viewpoint. *Southern Economic Journal*, 48, 583–93.

Staiger, Robert W. and Frank A. Wolak 1991: The effect of domestic antidumping law in the presence of foreign monopoly. *Journal of International Economics*, 32, 265–87.

13

Antidumping versus Antitrust: Trade and Competition Policy

Ian Wooton and Maurizio Zanardi

Chapter Outline

This chapter surveys the use of antidumping (AD) and antitrust (AT) policies in encouraging and maintaining competition in the international marketplace. The recent surge in the use of AD is documented and the debate on AD versus AT is analyzed, with a particular focus on the experiences of various Regional Trade Agreements. Some of the possible solutions to the debate are evaluated, taking as granted the political constraint that AD cannot be abandoned without an alternative policy being put in its place.

1 Introduction

Our interest is in the policies used by countries to discipline the behavior of imperfectly competitive firms in the international marketplace. In particular, we wish to examine the competition policy options available to national governments as their economies become more integrated with the world trading system. We are especially interested in the different opportunities that regional trade agreements (RTAs) offer relative to those under the multilateral trading regime of the World Trade Organization (WTO).

The international dimension is crucial. In a world of self-sufficient countries, firms would be producing goods entirely for domestic consumption. If firms behave competitively and there are no externalities, it has been long acknowledged that the free market is efficient. Were any firm (or group of firms) perceived to be exercising an inappropriate degree of market power, however, a government could use its arsenal of competitive policies to address this anticompetitive behavior. We shall label these as "antitrust" (AT) policies.

As trade barriers are lowered, firms will be exposed to foreign competition, leading to penetration of domestic markets by imports from abroad and opportunities for domestic firms to sell in new export markets. Traditional trade theory predicts that such trade will yield aggregate gains for all participating nations, though there can be long-term implications for industrial structure and income distribution. However, if firms behave competitively, global efficiency is achieved through a complete elimination of trade barriers (i.e., free trade) and without the need for any government intervention in product markets.

Difficulties arise when firms can and do exercise market power. In such instances, government intervention can be justified. The problem is that the instruments that are available to a government to discipline a domestic firm are not accessible in the case of a foreign-based enterprise. The exporting country could, of course, also use AT policies to force its firms to behave more competitively in foreign markets, but it frequently is not in the interests of the nation to do so (Levinsohn, 1996).[1] In the absence of AT instruments, the government must find alternatives. A resurrection of trade barriers is a possibility. Another is the use of "antidumping" (AD) policies, which allow governments to impose duties whenever goods are sold in export markets at less than their "fair value." These policies, in essence, prevent firms from price discriminating between markets.

Given the international dimension, the advantage of AD over AT is that it does not require any supranational powers. The disadvantage of AD policies is that they do not fully address the problem of inappropriate use of market power and, indeed, often induce more distortions in the market than they resolve and are often captured by special interests. As it is generally believed that the negative aspects outweigh the positive ones, a major part of this chapter is an investigation of the motives behind AD.

If the debate on AD versus AT hinges on the issue of supranational powers and coordination of local provisions, the recent wave of globalization can offer a solution. Depending on the degree of integration of the participating countries, supranational authority might be possible. Indeed with an RTA, AT may be rehabilitated. Therefore, another component of this chapter is the investigation of whether AT should replace AD under an RTA (and why it frequently has not), with an emphasis on the interplay between economics and legal provisions.

In summary, this chapter focuses on the most appropriate and feasible means of encouraging and maintaining competition in the international marketplace, using AD and AT instruments. However, it is not intended to be a survey of dumping, nor a survey of AD, as the long literature would take us too far afield and we would unnecessarily replicate recent contributions (Blonigen and Prusa, 2003; Niels, 2000). In the same way, it is not a survey of AT. In the next section, we focus on the issues surrounding the implementation and use of AD. Section 3 considers the justifications for the protection of competition and how AD and AT, especially, address it. The multinational dimension is analyzed in Section 4 and some conclusions are drawn in Section 5.

2 Antidumping

A discussion of AD would not be of value, at least from a practical point of view, unless the phenomenon were quantitatively relevant. There are various ways to assess

the (mis)use of this largely protectionist tool. In the following, we present some data on the adoption of AD laws and their use. Even though usually ignored, the time pattern of worldwide implementations of AD laws is interesting in itself, and particularly in relation to the trade liberalization experience of many countries. More generally, we argue that the surge in the use of AD can be linked to the parallel processes of trade liberalization and legal developments in the General Agreement on Tariffs and Trade (GATT) and WTO dispositions.

2.1 Some Numbers

Figure 13.1 shows the worldwide implementation of AD laws over time. The history of AD starts at the beginning of the twentieth century, with Canada being the first country to adopt an AD legislation in 1904, followed by Australia in 1906 and several other nations in the years up to 1920. Finger (1993) explains this first wave of AD legislation referring to a general hostility toward Germany, which, toward the end of World War I, had a huge production surplus ready to be dumped,[2] and the general climate of trust busting in the political debate.[3] Following this initial burst of AD legislation, Viner (1923) provides the first comprehensive analysis of this protectionist tool and of its economic rationale as a remedy for discriminatory pricing and cartel behavior.

The subsequent 30 years were surprisingly quiet. Beginning in the 1950s, however, AD started to acquire a worldwide status, with more and more countries adopting AD laws at an increasing pace, only interrupted in the 1970s, that reaches a total

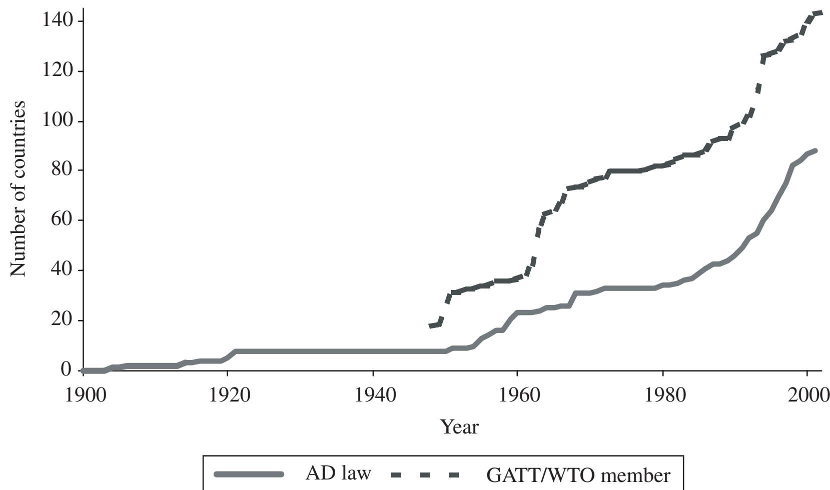

Figure 13.1 Countries with AD laws and membership of GATT/WTO (as of December 31, 2001). Source: Zanardi (2004b) for AD laws and http://www.wto.org for GATT/WTO membership

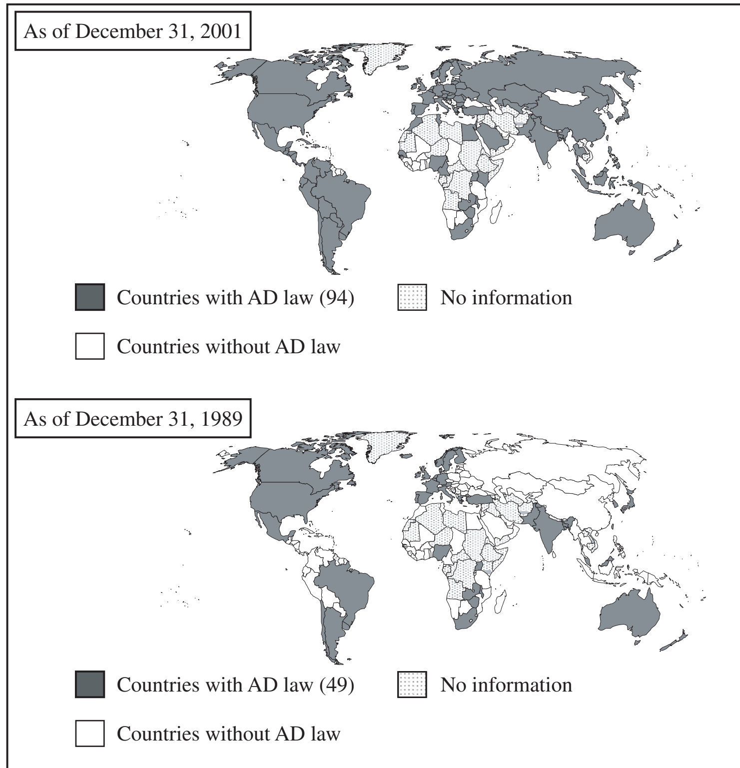

Figure 13.2 The geography of AD laws. Source: Zanardi (2004b)

of 94 countries having an AD law by the end of 2001.[4] In the meantime, the rationale for AD as a remedy for predatory behavior (as suggested by Viner) has changed dramatically. This is reflected in a recent communication from the United States to the WTO (1998): "The antidumping rules are not intended as a remedy for predatory pricing or any other private anti-competitive practices with which competition laws are concerned."

The number of countries with AD laws only gives a partial picture. Figure 13.2 places these numbers in a geographical context and it shows that, in the last decade, most of the activity in introducing AD legislation has been in developing countries. Much of this is catch-up, as most of the industrialized world already had AD laws prior to 1990. The least change has taken place in central Africa, where few countries have any AD laws. These countries are also amongst the least industrialized with low trade volumes. Thus their need for AD policies is not great, nor would any such policies have much impact on world trade. Despite this, the increased adoption of AD laws worldwide means that more than 90 percent of worldwide imports in 2000 were potentially subject to AD actions![5]

Despite the fact that the economic literature has shown that AD laws can have pervasive effects even when duties are not imposed,[6] numbers relating to the application

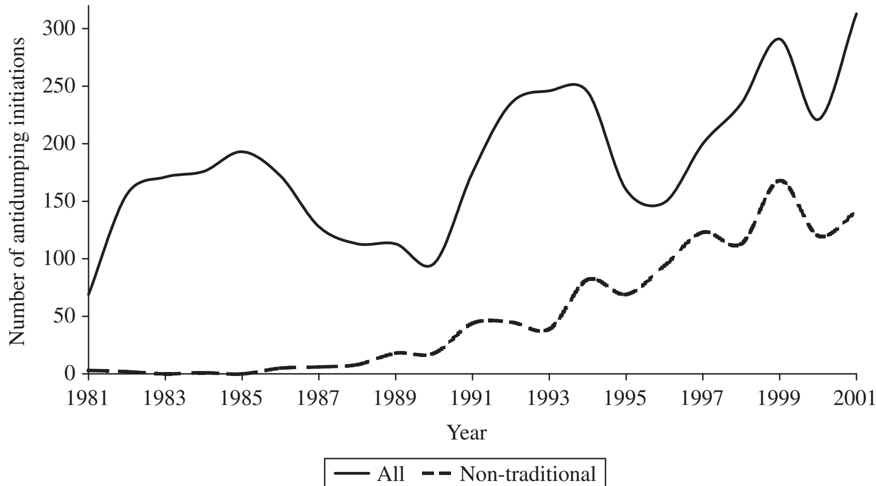

Figure 13.3 AD initiations from 1980, distinguishing nontraditional users. Note: "Nontraditional" refers to all countries except Australia, Canada, the European Union, New Zealand, and the United States. Source: WTO reports on AD activity

of AD are natural counterparts to the data on the adoption of AD laws. Figure 13.3 illustrates an upward trend in the number of petitions filed since the beginning of the 1980s.[7] Moreover, the ups and downs seem to correlate with business cycle fluctuation; a hypothesis confirmed by Knetter and Prusa (2003) for Australia, Canada, the European Union (EU), and the United States. This same graph also shows that the growing number of countries adopting an AD law is consistent with larger and larger shares of petitions being filed by nontraditional users (i.e., countries other than Australia, Canada, the European Union, New Zealand, and the United States).

Simply looking at filing statistics can be misleading, since the number of petitions must be related to the total volume of imports for each country. In other words, five petitions in five years by Trinidad–Tobago may be relatively more important than 136 investigations by the United States in the same period, given the total value of imports for these countries. Finger et al. (2001) carry out an analysis of this issue, constructing an index of initiations (impositions) relative to imports (exports) over the period 1995–9.[8] Tables 13.1 and 13.2 summarize some of their results. In relative terms, developing countries are the most intense users of AD, while the United States and the European Union (and industrial economies in general) are among the least intense users, contrary to the usual presumption of blame. A similar picture obtains when looking at the imposition of AD duties, with transition economies being the most targeted countries.

Reviewing these numbers can only lead to the conclusion that the importance of AD has greatly increased over time. However, the players have changed. Developing countries are now using this protectionist tool as much, if not more, than the traditional countries. Why is this the case? Has the occurrence of dumping increased? In the following, we suggest some possible answers.

Table 13.1 Antidumping initiations, 1995–9

Initiating country	Initiations	Initiations per $US imports
Argentina	89	2,125
South Africa	89	2,014
Peru	21	1,634
India	83	1,382
New Zealand	28	1,292
Trinidad & Tobago	5	1,257
Venezuela	22	1,174
Nicaragua	2	988
Australia	89	941
Colombia	15	659
Brazil	56	596
Panama	2	431
Israel	19	418
Chile	10	376
Indonesia	20	330
Mexico	46	290
Egypt	6	278
Turkey	14	204
Korea	37	185
Canada	50	172
Guatemala	1	168
Costa Rica	1	144
Ecuador	1	140
European Union	160	130
Philippines	6	113
United States	136	100
Malaysia	11	97
Slovenia	1	66
Poland	4	65
Czech Republic	2	45
Singapore	2	10
Thailand	1	10

Source: *Finger et al. (2001).*

2.2 Some Possible Explanations

Short of being a complete answer, Figures 13.1 and 13.4 illustrate a strong correlation between trade liberalization and AD. Membership in GATT/WTO is accompanied by implementation of AD laws. Even if an AD law is not a necessary condition for accession to the WTO, in recent years the two aspects have been closely related. In this case, it is likely that the solid line in Figure 13.1 will slope upward in the near future, as more than 20 countries without such a law are negotiating access to the WTO.

Table 13.2 Targets of antidumping initiations, 1995–9

Targeted country	Initiations	Initiations per $US exports	Targeted country	Initiations	Initiations per $US exports
Armenia	1	6,454	Slovak Republic	5	488
Georgia	1	3,723	Thailand	30	485
Kyrgyzstan	1	3,559	Turkey	13	478
Tajikistan	1	3,003	Hungary	9	475
Azerbaijan	1	2,970	Macau, China	1	451
Yugoslavia	5	2,913	Poland	12	427
Kazakhstan	11	2,465	Pakistan	4	417
Macedonia	3	2,203	Bahrain	1	404
Ukraine	25	1,995	Taiwan	47	368
Bosnia-Herzegovina	1	1,790	Hong Kong, China	11	363
Latvia	3	1,731			
			Uruguay	1	361
Egypt	6	1,531	Argentina	8	298
Bulgaria	6	1,226	Colombia	3	252
Uzbekistan	3	1,213	Czech Republic	6	229
Belarus	6	1,195	Venezuela	5	220
Cuba	2	1,188	Vietnam	2	212
Romania	10	1,099	Slovenia	2	211
Lithuania	4	1,054	Malaysia	16	190
India	38	1,028	Mexico	20	179
Honduras	1	1,026	Peru	1	173
Paraguay	1	873	Israel	4	169
Zimbabwe	2	847	Iran	3	151
Moldova	1	811	New Zealand	2	137
Bolivia	1	797	European Union	179	101
South Africa	20	770	United States	66	100
Brazil	41	750	Japan	44	95
Trinidad & Tobago	2	746	Philippines	2	78
China	137	739	Saudi Arabia	4	63
Estonia	2	713	Algeria	1	61
Indonesia	36	658	Australia	3	53
Croatia	3	598	Singapore	6	46
Korea	75	537	Canada	10	44
Turkmenistan	1	535	Switzerland	3	34
Russian Federation	41	531	Norway	1	20
Chile	9	528	Liechtenstein	1	Not available
Costa Rica	2	489			

Source: *Finger et al. (2001).*

Similarly, the International Monetary Fund (IMF) and the World Bank have, in the past, included the enactment or drafting of AD legislation as a condition for the release of loans.[9]

Figure 13.4 suggests a process of substitution of definitive AD measures in place of tariffs as nontraditional countries move along the trade liberalization path. Even though

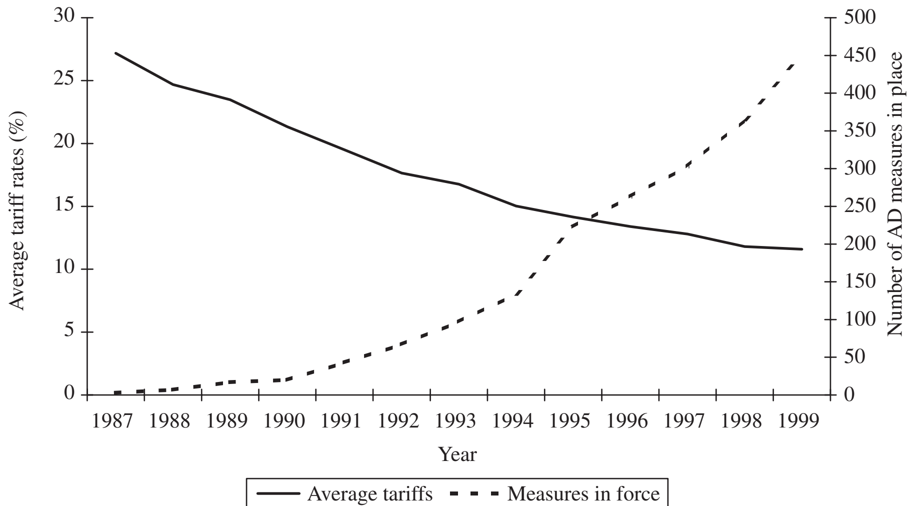

Figure 13.4 Average tariffs and definitive AD measures for nontraditional users. Note: "Nontraditional" refers to all countries except Australia, Canada, the European Union, New Zealand, and the United States. Source: Lindsey and Ikenson (2001). Reproduced with permission.

correlation does not imply causation, there is a generalized presumption that

> . . . antidumping laws administered in strict conformity with the Antidumping Agreement actually assist governments in their efforts to continue trade-liberalizing measures by providing relief to domestic industries injured by foreign firms that engage in unfair trade practices, even as international trade liberalizes. From this perspective, the antidumping rules are a critical factor in obtaining and sustaining necessary public support for the shared multilateral goal of trade liberalization.
>
> ***Communication of the US to the WTO (1998)***

In the very same document, the United States goes even further in stating that AD is "necessary to the maintenance of the multilateral trading system."

But the hypothesis of AD as a pressure valve is not fully able to explain the scenario outlined in Figures 13.1 and 13.4. In fact, many countries had AD laws on their books for many years but did not use them until recently. Bureaucratic limitations certainly constitute a reasonable explanation for some countries, but a retaliatory motive is another worrying possibility. Indeed, some countries have gone as far as including a reference to the latter possibility in their law. For example, Article 56 of the new Chinese law that entered into force on January 1, 2002 states that "Where a country (region) discriminatorily imposes anti-dumping measures on the exports from the People's Republic of China, China may, on the basis of the actual situations, take corresponding measures against that country (region)."[10] A formal test of this motive is provided by Prusa and Skeath (2002), who find some support for a retaliation motive for both traditional and new users of AD, while excluding an increase in unfair practice as an explanation of the increased use of AD.

Another possible explanation is that the increased use of AD by developing countries is related to a change in the composition of traded goods. Obviously, this is associated with their increasing role in world trade arising from their pursuit of trade liberalization. Still, this hypothesis does not explain why nontraditional users apply the AD law as heavily as reported in Table 13.1.

If AD as either a pressure valve or means of retaliation goes a great way in explaining the surge of initiations, the time pattern of implementation of AD laws in Figure 13.1 also suggests a legal element to complement these explanations. Once GATT entered into force, the first wave of implementations took place in the late 1950s, coinciding with the decision of the contracting parties to GATT to have a "systematic study" of AD legislation (GATT, 1958). The first Antidumping Code was negotiated during the Kennedy Round and entered into force in 1967, with five countries then implementing an AD law in 1968 and a couple more following shortly thereafter.[11] The Tokyo Round Code dates to 1980 and again Figure 13.1 shows an increase in implementations in the 1980s after a period of inaction. The Uruguay Round Antidumping Code entered into force on January 1, 1995 and represented a major leap forward. Apart from the new rules that it includes,[12] this Code is an integral part of the WTO Agreement and, as such, it applies to all members whereas previous codes had to be signed by each GATT member.[13] This requirement can explain the steep slope in Figure 13.1 in the second half of the 1990s. It may be that the WTO requirements standardize AD laws thereby limiting the protectionist bias in some countries' legal dispositions. Nonetheless, we think that the temptation to (mis)use an AD law is strong once a country implements it.

Whether the link runs from the legal requirements to the implementation and use of AD laws, or vice versa, cannot be said with the data available. However, it is reasonable to conclude that legal aspects, trade liberalization, and retaliatory motives are part of the same (vicious) circle that surrounds the use of AD.

2.3 Antidumping and Regional Trade Agreements

The economic literature that discusses trade liberalization always acknowledges the importance of RTAs. Preferential trading has flourished recently, with around 80 percent of the existing RTAs having come into force after 1989. What is the relationship, if any, between AD and RTAs, and what is the role of AT? There is no simple answer to these questions because no clear pattern can be detected; though some general conclusions can be drawn.[14]

At one extreme of the spectrum, the European Union has substituted AD with AT for intra-union trade and applies common (supranational) AD procedures with regard to third countries.[15] There is no harmonization of national AT regimes but a supranational commission deals with union-wide competition matters. The European Union represents a very deep form of integration with free movement of goods, labor, capital, and coordinated monetary and fiscal policies. In this regard, the elimination of AD can be interpreted as a necessary step in order to reach a common market, more than the consequence of a supranational AT discipline.

The experience of the EU is particularly instructive when considering the various forms of agreements that the EU has signed with third countries. In fact, while AD has

been abolished within the European Economic Area (EEA)[16] without any reference to AT, the European Union retains the right to initiate AD investigations for imports from countries in the Agreements with Eastern European and Mediterranean Countries. Even if the EEA is a free-trade area (i.e., the simplest form of an RTA in which there are zero tariffs on bilateral trade flows but no common trade policy), the stated objective of creating a single market entails elimination of AD. By the same token, it does not come as a surprise that AD was not eliminated in the other cited agreements.

The goal of a common market seems to be the driving force when dealing with AD in RTAs, while AT does not appear to be a critical factor in the process. This interpretation of the European experience is confirmed by the Canada–Chile Agreement where AD was eliminated without any replacement with AT. The relationship between Australia and New Zealand (ANZCERTA) is less clear-cut since both countries changed their AT law to allow firms to be tried in both jurisdictions. The official document states that ". . . the maintenance of antidumping provisions . . . ceases to be appropriate as the Member States move towards the achievement of full trade in goods between them and a more integrated market."[17]

In the Americas, the firm support of the United States for AD made it impossible to accomplish much on the issue in the negotiations that led to NAFTA. The only concession has been the establishment of binational panels to ensure the correct application of the laws. Analogously, there has been no attempt at any harmonization of AT. In Mercosur, a double AD regime applies depending on the source of imports. National agencies are supposed to coordinate their actions for intra-region AD investigations. In contrast, a common legal framework, but without a supranational authority, applies for AD actions against third countries. Intra-region AD laws are to be completely eliminated and replaced by a common competition policy once the transition period is over and the as-yet unratified Protocol of Fortaleza is implemented. But this latter development seems far from becoming a reality since Paraguay and Uruguay still do not have AT laws.

In prospective terms, the Free Trade Area of the Americas (FTAA) and the Asia-Pacific Economic Cooperation (APEC) forum have working groups on AD but in neither case is the discussion made in conjunction with AT.

In conclusion, AD concerns in RTAs are not necessarily related to AT, even though some sort of coordination on competition policies is of help. The experiences surveyed also show that the particular trade arrangement (i.e., free-trade area versus custom union) is not the determining factor. The intention of creating a single market seems to be the critical element that triggers the phasing out of AD.

3 Justification for Competition Policies

We focus on the methods that countries can use to control the inappropriate exercise of market power by firms. Both AD and AT are instruments of competition policy and their use is justified on this basis. Specifically, they target price discrimination (i.e., charging different prices across markets and/or consumers), which is the result of segmented markets in an international setting. As such, AT and AD could be regarded as complementing each other where AT applies to domestic markets and AD pertains to the

domestic sphere of foreign firms. However, the ultimate objectives are quite different with AT aimed at protecting consumers' interests and AD designed to safeguard firms' businesses. Still, the two sets of policy *ideally* are meant to complement each other, as they act upon the same market distortion.[18]

The real-world experiences of the interface between AT and AD are far from showing strong complementarity, if any complementarity at all. Between the two, AD policies have moved farthest from this ideal scenario. Other economic and, particularly, political motives have shaped AD practices so that nowadays they are rather different from the idea of "fairness" that motivated them in the first place. In this process, any link between AT and AD has been lost and the two policies are now total strangers. In the following, a more detailed analysis of the justifications of AD and AT form the basis for the discussion of the possible solution(s) to the AD versus AT debate.

3.1 Antidumping Policy

Dumping is not, in itself, a cause for concern in international markets. It can, indeed, be a necessary component of pro-competitive trade. As trade barriers are lowered, firms have an increasing incentive to sell into foreign markets. This makes markets more competitive as domestic producers no longer have the market to themselves. Consequently, domestic prices are forced down to the benefit of consumers. It is very likely, however, that the foreign goods are being "dumped," in the sense that producers earn less on each good exported than if it were sold on their domestic market. For example, if there are still trade costs, these have to be absorbed by the exporting firm. Otherwise, exports would be more expensive than domestically produced goods that don't face the same trade costs. Thus, rather than being exceptional, anticompetitive behavior, dumping is a quite typical response to trade barriers and is generally pro-competitive, benefiting consumers. AD aids and abets domestic firms in exercising their market power, in that it either forces up the prices of imports or, in the extreme, drives the foreign firms out of the market entirely.

The only sound economic justification most economists would agree upon in defending the need of an AD law is the occurrence of *predatory* dumping. This is an extreme form of price discrimination with dominant firms lowering the price (i.e., "dumping" on foreign markets) to such a degree as to drive competitors from the market. The disappearance of domestic firms and the resulting monopolized market imply a net loss for the importing country's welfare, which makes AD provisions legitimate.[19] This justification is the motive that originally prompted the implementation of AD laws in the early years of the twentieth century. However, this concern seems to be vanishing. The literature shows that very few of the AD petitions filed in the 1980s were brought against instances of possible predatory dumping (Messerlin, 1996; Shy, 1998).[20]

Nowadays AD seems to have little concern with either predation or dumping. Some authors refer to AD as merely another form of protection entirely unconnected to any observed dumping behavior. Indeed, the criteria for proving dumping and consequent injury allow firms around the world to seek (and obtain) protection from the AD procedures. Industries at large have been granted protection through the numerous petitions they have filed in some countries.[21] In this way, AD duties have become

a substitute for tariff protection as trade liberalization has progressed (as shown in Figure 13.4) or have been pursued on pure protectionist grounds by some of the traditional users of AD (e.g., the European Union and the United States). AD protection benefits domestic producers and the direct impact on consumers is analogous to the effect of import taxes, in raising the prices for both imported goods and domestically produced import-competing products. The net effect on national welfare is negative.[22]

More worrying are the indirect effects of AD, as they are more difficult to quantify and may arise once an AD law is implemented even if it is not used. The long literature on this aspect as well as on the strategic behavior induced by the mere existence of the law (Blonigen and Prusa, 2003) points out that the adverse effects may be exacerbated or diminished depending on the exact market conditions. AD may facilitate collusion between home and foreign firms.[23] It could also be pro-competitive when collusion preexists the introduction of the law (Hartigan, 2000) or the law is simultaneously introduced in more than one country (Anderson et al., 1995). AD provisions may also induce foreign direct investment to avoid duties. Such "antidumping jumping" presents ambiguous welfare implications as domestic producers lose because of increased competition but consumers pay lower prices (Haaland and Wooton, 1998). As we have already mentioned, retaliation is another possible, and indeed likely, effect of AD.[24] However, it may be argued that retaliation will disappear in the long run if a "cold war" equilibrium is reached, where every country is equipped with an AD law but does not use it because of the credible threat of its trade partners responding in the same fashion. Judging from the data discussed in the previous section, it seems that we are still quite far from such an outcome, assuming this is indeed the eventual equilibrium.

If dumping is supposed to be a technical concept and the imposition of AD duties a bureaucratic process according to the WTO AD Agreement, nothing is further from reality. Political justifications as well as political influences in the administration of AD are crucial in explaining the resistance of some countries, and the United States in particular, to change the WTO AD Agreement.

The typical political justification is that AD is a safety valve for trade liberalization, which, it is argued, would not be otherwise accomplished because of the objections of the sectors disadvantaged in the move to freer trade. If this is the position of most industrialized countries, developing countries and Japan present an opposite view.

> Japan shares the view of many developing-country Members [to the WTO] which have expressed concerns about the abuse of anti-dumping measures and insisted on removing ambiguity and excessive discretion inherent in the Anti-Dumping Agreement. These characteristics have dangerously made it an instrument of protection rather than an instrument to counteract unfair dumping. Japan is concerned that this situation poses a threat to the multilateral trading system. Therefore, Japan supports clarifying and strengthening disciplines in the Anti-Dumping Agreement.[25]
>
> ***Communication of Japan to the WTO (1999)***

Clearly the political consensus is not general. The United States sees AD as necessary to sustain the multilateral trading system, while Japan thinks of it as a threat. This demonstrates how various interests can twist the same WTO Agreement and depart

from its principles in its implementation and use. A new round of multilateral trade negotiations is currently under way after its launch in Doha, Qatar. The United States has agreed to the inclusion of AD in the agenda with the objective of "clarifying and improving" its rules "while preserving the basic concepts, principles and effectiveness" (WTO, 2001). Still, the resistance of the United States to any substantial change is revealed in a resolution approved by the United States Congress just a few days before the WTO Ministerial Conference in Doha stating,

> . . . it is the sense of Congress that the President, at the WTO round of negotiations . . ., should (1) preserve the ability of the United States to enforce rigorously its trade laws, including the antidumping and countervailing duty laws, and avoid agreements which lessen the effectiveness of domestic and international disciplines on unfair trade, especially dumping and subsidies, in order to ensure that United States workers, agricultural producers, and firms can compete fully on fair terms and enjoy the benefits of reciprocal trade concessions; and (2) ensure that United States exports are not subject to the abusive use of trade laws, including antidumping and countervailing duty laws, by other countries.
>
> ***US House (November 7, 2001)***

Interestingly, this statement seems to show double standards with respect to the active use of AD and of being its target. In other words, the United States recognizes the possibility of a misuse of AD when it comes to American exports but it fails to admit any wrongdoing in its use of AD. In citing the constituency that AD protects (i.e., workers, agricultural producers, and firms), the resolution repeats the political justification of the law.

A reader looking for complementarity between AD and AT will notice that consumers' protection is not mentioned in the quote. Indeed, from the very beginning AD has always been concerned with "unfair" business practices and their relevance for domestic firms, leaving aside considerations for consumers and national welfare. Some countries do include a national welfare clause in their law (the European Union is a primary example). This does not seem to significantly change the way the law is administered. Such clauses are unlikely to resolve the conflict between AD and AT or reduce the misuse of AD, unless they are formulated in such a way as to become binding on the AD authorities.

The same list of interests in need of protection also explains the relevance of AD in the domestic political arena. Various contributions in the literature illustrate the channels of political influence over the application of AD (Finger et al., 1982; Moore, 1992; Hansen and Prusa, 1997). The mechanics of any channel basically rely on an "exchange" between interests groups (i.e., lobbies) and politicians, where favorable AD outcomes are the "price" paid for obtaining political support. This political economy aspect is probably the most important factor in causing the misuse of AD. At the same time, it is likely to be the most difficult to eliminate.

The interplay of economics and politics and the increasing role of the latter explain the evolution of AD from a tool originally conceived as an extension of AT for foreign firms' misbehavior to a standard protectionist instrument. AT is not necessarily administered free of bias, however, as we describe next, its underpinnings are more economic in nature and the clearly stated objective of consumers protection poses a limit to political influences.

3.2 Antitrust Policy

Competition policy (which we shall continue to refer to as "AT") shares a common target with AD, the prevention of predatory behavior. One firm undercutting the prices of a rival is not, of course, unequivocal evidence of predation. The price difference may simply be the result of differences in productive efficiency. In such circumstances, the competition is healthy and in the overall interests of the economy, despite the negative impact on the less efficient firm. The task of the AT authority is to distinguish between (what may be) intense price competition and true predation, as only the latter warrants intervention.

AT is more focused on encouraging competition than on preventing it. As trade barriers come down and markets become increasingly integrated, the costs of entry into foreign markets diminishes. These pro-competitive developments may not be in the interests of firms, who previously enjoyed greater market power in their domestic markets. As we have established, AD often works to reinstate some of this lost power. AT policies, instead, are designed to ensure that firms do not undercut the benefits of market integration by collusive behavior. The interests of consumers (and overall economic efficiency) are paramount with AT, while they are largely ignored by AD.

In determining whether AT measures are justified, a series of economic tests have been developed to evaluate when firms have dominant positions (depending on their market shares) and strict criteria of proof in each case. This is in sharp contrast to the practice of AD. There is a vacuum with regard to meaningful economic criteria and subjective decisions by commissioners (in the European Union even more than in the United States) are important, leaving the door open to interest groups to influence the outcome of the investigations.

Were governments intent on promoting efficiency and competition, AT policies would clearly be the superior instrument, as AD can do nothing to eliminate collusion between firms and actually goes a great length in facilitating such practices. The deficiency of AT is the inability of governments to apply it to firms other than domestic firms, or at least production taking place within domestic borders.

3.3 Summary

If AD and AT have a common target (i.e., anticompetitive behavior), the discussion so far makes clear that AD is serving other purposes as well. From a purely economic point of view, the distortions induced by AD decrease a country's welfare. However, redistribution of income and rent and their associated benefits make AD a very convenient tool for both interest groups and policymakers (and politicians).

AT appears to be the more effective instrument in dealing with anticompetitive behavior on the part of firms. Indeed, except in the instance of predatory behavior, there seems to be no economic justification for AD policies. Eliminating AD would consequently, and independent of any other actions, yield benefits to the world economy.[26] Thus it is argued by Niels and ten Kate (1997) that there is no need to replace AD with any alternative policy such as AT. In defense of AD, its elimination could make countries

the targets of predatory behavior. This might not be a big concern, however, as such a strategy seems very difficult to pursue at an international level.

A two-tier approach has been proposed by some authors. In such a framework, an AD case would be first judged using AT criteria and, only if it passes them, would it be allowed to proceed as an AD case. Thus AD actions would only be justified if there were a concern about abuse of market power, reducing the competitiveness of the market. Linking AD to AT in this way would allow AT to play a central role in shaping AD. This would help return AD to the perspective of the early nineteenth century when it was first introduced.

In the debate over competition policies and trade policies that has taken place in the WTO, the relationship between AD and AT is never tackled (and this remains true for the Doha Ministerial document). Academics (economists and lawyers) address the issue but this analysis does not seem to reach policymakers. However, some small positive signs are present. For example, Lithuania introduced its AD law in 1998 and it placed its authority in the same agency also responsible for consumer protection. Further, in a communication to the WTO in 1999, Mexico "give(s) some practical examples of the way in which certain basic concepts of competition policies could be used to improve the antidumping rules" (WT/WGTCP/W/136, July 15, 1999). It appears that some developing countries have learnt the lesson about the abuse of AD and they are willing to "fix" the problem instead of only retaliating.[27] It might be expected that any initiative to "solve" the problem of AD should come from one of the large global organizations (such as WTO, World Bank, OECD). However, there is concern that these bodies have been captured by US influence and are therefore silent on these issues.

4 Multidimensional Economic Integration

Both trade liberalization and competition policy can be pursued multilaterally or bilaterally. Table 13.3 attempts to provide a characterization of the policy combinations available to countries. In a world with high trade barriers, there is little trade as foreign firms are put at a serious disadvantage relative to local producers. Any competition problems will arise from domestic firms exercising their market power, and so domestic AT policies will provide the correct remedy.

As trade barriers begin to crumble as a result of multilateral trade negotiations, imports take an increasing share of domestic markets. As we have seen, and for the reasons elaborated, under the multilateral approach trade liberalization and

Table 13.3 Stages of bilateral integration

Trade	Competition policy
Self-sufficiency (autarky)	Autonomous competition (AT) policy
Unilateral tariffs	Unilateral AD actions
RTA (free bilateral trade)	Bilateral removal of AD
Customs union (common external tariff)	Common external AD policy
Common market	Common AT policy

competition policy (AD) appear to be substitutes. Figure 13.4 shows clearly that AD legislation/application increases as tariffs diminish.

The alternative path toward economic integration takes a more regional approach. The degree to which an RTA will also yield common competition policies will depend upon the degree of integration that the member countries wish to have. The lowest level of integration that is consistent with WTO obligations is the free-trade agreement. This simply obliges countries to set zero tariffs on their imports from other member nations. No aspect of national trade policy is placed in the hands of a regional or supranational agency and so there is no reduction in national sovereignty. Nations could (and we have given instances where some have) parallel the trade concession with an elimination of AD between members in the regional agreement. While this may do nothing to deal with disciplining the behavior of large firms, it would have the benefit of removing a protectionist policy, as Niels and ten Kate (1997) argue.

Steps toward deeper trade integration can also be matched on the competition front. Countries that establish a customs union must jointly agree on common external tariffs. Antidumping policy could be put on the same footing, with a regional policy replacing the individual policies of member countries with regard to imports from the rest of the world. It is not clear whether such regional coordination would improve welfare. On the one hand, firms from outside countries would have to face one consistent set of AD policies for sales in the region. On the other hand, the region may, through coordinating its actions, be able to wield AD more effectively as a protectionist device.

This problem remains even when countries pursue yet deeper integration as a common market. However, the completion of the regional market calls for the introduction of a region-wide AT regime. Within a common market, supranational institutions are already in place and so the issue of sovereignty has, to some degree, been resolved. As a result, AT can be introduced at the regional level, resulting in a more effective policing of firm behavior within the regional market. Then, to the extent that AD should be linked to AT, an RTA offers better prospects than multilateral negotiations carried out at the WTO level. Even though some advocate that the WTO should be involved with competition policies, the institution's trade focus may prevent it from being the ideal venue for discussion of the harmonization of AT regimes.[28]

5 Conclusions

We have considered the relationship between trade liberalization and competition policy. There is strong evidence of the increasing use of AD policies in the liberal trading regime that has been achieved in the global economy. But AD is not the best approach in dealing with corporate abuse of market power. Consequently, we have sought out alternatives to AD, particularly in the context of nations increasingly being members of some form of RTA.

It seems that, at least on political grounds, AD cannot be abandoned in the absence of an alternative policy, despite a compelling case for its elimination (except in the case of predation). Indeed a political economy approach seems to be the only way to explain the persistence of AD. The compromise is to find some form of alternative policy with which to replace AD. What might be considered the "best" policy on efficiency grounds,

a supranational application of competition policies, is not however considered to be politically feasible.[29] Thus, whatever policy is found, it will be "second best" from an economic viewpoint.

What feasible and efficient policies might be used? A national welfare clause, amending AD rules to move the focus of concern from domestic producers to overall national well-being, might be promoted. However, there is not much evidence of its success in countries that already have such a clause in place. If the WTO Agreement were to impose a national welfare clause with strict criteria, there is the possibility that it would be effective in limiting the existing abuse of AD laws. Then again, it is probably more realistic to acknowledge that special interest groups would find loopholes in any legislation (as they already do).

Probably RTAs can offer an "intermediate" solution as they improve regional AD legislation. They push national authorities into relinquishing their powers. The experience of doing this at a regional level may encourage countries in the long run to eliminate AD on a larger scale. Independently, RTAs that eliminate AD offer a micro model of how things should work and should be taken as exemplars of good practice. In this sense, the Treaty of Rome in 1957 was far more advanced than more recent forms of integration.

Acknowledgments

We wish to thank Michael Finger and Will Martin for comments and suggestions.

Notes

1 Many AT laws provide exceptions of some sort for export cartels. For example, in the United States, actions in violation of AT are permitted if firms are engaged exclusively in export activities.
2 Viner (1923) suggests that this mistaken belief may have been caused by "the wartime plague of mendacious propaganda."
3 In the United States, the first AT law (the Sherman Antitrust Act) was enacted in 1890. Australia refers to AD in the Federation papers of 1901, with provisions being enacted in 1906 to address predatory dumping.
4 The 15 countries of the European Union are counted individually. The total reported in Figure 13.1 is less than 94 because of missing data on implementation for some countries.
5 Intra-EU imports are not included since AD actions are not allowed among member countries.
6 See Blonigen and Prusa (2003) for a discussion.
7 All data presented in the following are derived from GATT/WTO documents. Known mistakes in these sources (Zanardi, 2004b) do not change the general picture of the phenomenon.
8 Imports and exports are measured in US dollars for all countries.
9 The World Bank offered technical and legal advice to many countries (Krishna, 1997) even if economists within the World Bank voiced their opposition (Finger, 1992). The IMF approved credit for Pakistan in 2000 subject to the enactment of a new AD law.
10 The very same article was present in the original law implemented in 1997. Russia's AD law contains a similar provision.
11 The European Community implemented its common AD law in 1968.

12 See Schott (1994) for a detailed summary.
13 For example, the United States never signed the Kennedy Round Code.
14 See Hoekman (1998) for an extensive survey on AD and AT in RTAs.
15 The Treaty of Rome (1958) constituting the European Community provided for the elimination of AD after a transition period.
16 The EEA is the European Union with Liechtenstein, Iceland, and Norway.
17 Preamble to the Protocol on Acceleration of Free Trade in Goods (1988).
18 Price discrimination is only one among many of the illegal practices covered by AT.
19 Both consumers and producers lose in the long run as higher prices will be charged by foreign producers to recuperate their short-run losses.
20 Recently, AD advocates have abandoned predatory dumping as their main justification and have invoked unfair arguments based on price discrimination and below-cost sales (Mastel, 1998). In turn, these claims have been proved unrelated to the actual AD practices (Lindsey, 1999).
21 Prominent examples include the steel industry, as well as the textile and chemical sectors.
22 Gallaway et al. (1999) estimate that the collective net cost of AD and countervailing duty orders (a less frequently used law that offsets export subsidies) was around $4 billion in 1993 for the United States. This amount is larger than the cost of any US trade restraint policy with the exception of the Multi-Fibre Arrangement.
23 See, among others, Staiger and Wolak (1992), Messerlin (1990), and Zanardi (2004a).
24 All these indirect effects would be present even if AD were aimed at predatory dumping only, though most likely in milder forms since firms would see that the law applies in a restricted number of cases.
25 This statement was received with vehement opposition from the United States and guarded support from the European Union.
26 The United States implementation of the Tokyo Round additionally makes it easier to claim cost dumping. Making domestic sales at prices below cost, for example, to unload inventories of seasonal items or apparel out of style or liquidation, are however not prohibited under AT law.
27 Or it may simply be an illusion.
28 See Guzman (2003) for the various positions in favor and against discussions of AT in the WTO.
29 In this regard, OECD (2000) notes that it “does not retain options that have been discarded in joint discussions as unrealistic, such as full harmonisation of competition laws, or an international antitrust authority with supranational powers.”

References

Anderson, Simon P., Nicolas Schmitt, and Jacques-François Thisse 1995: Who benefits from antidumping legislation? *Journal of International Economics*, 38, 321–37.

Blonigen, Bruce A. and Thomas J. Prusa 2003: Antidumping. In Choi, E. Kwan and James Harrigan (eds.), *Handbook of International Trade*, Oxford, UK and Cambridge, MA: Blackwell Publishers.

Finger, Michael J. 1992: Should developing countries introduce antidumping? Never. World Bank Outreach No. 1. Reprinted in Finger, Michael J. (ed.), *Institutions and Trade Policy*, Edward Elgar: Cheltenham, UK and Northampton, MA, USA. 2002.

Finger, Michael J. (ed.) 1993: *Antidumping: How It Works and Who Gets Hurt*. Ann Arbor: University of Michigan Press.

Finger, Michael J., Keith H. Hall, and Douglas R. Nelson 1982: The political economy of administered protection. *American Economic Review*, 72, 452–66.

Finger, Michael J., Francis Ng, and Sonam Wangchuk 2001: Antidumping as safeguard policy. World Bank Working Paper No. 2730.

Gallaway, Michael P., Bruce A. Blonigen, and Joseph E. Flynn 1999: Welfare costs of US antidumping and countervailing duty laws. *Journal of International Economics*, 49, 211–44.

General Agreement on Tariffs and Trade 1958: *Anti-Dumping and Countervailing Duties*. Geneva.

Guzman, Andrew T. 2003: International antitrust and the WTO: the lesson from intellectual property. *Virginia Journal of International Law*, 43, 933–57.

Haaland, Jan I. and Ian Wooton 1998: Anti-dumping jumping: reciprocal anti-dumping and industrial location. *Weltwirtschaftliches Archiv*, 134, 340–62.

Hansen, Wendy L. and Thomas J. Prusa 1997: The economics and politics of trade policy: an empirical analysis of ITC decision making. *Review of International Economics*, 5, 230–45.

Hartigan, James C. 2000: An antidumping law can be procompetitive. *Pacific Economic Review*, 5, 5–14.

Hoekman, Bernard 1998: Free trade and deep integration: antidumping and antitrust in regional agreements. World Bank Working Paper No. 1950.

Knetter, Michael M. and Thomas J. Prusa 2003: Macroeconomic factors and antidumping filings: evidence from four countries. *Journal of International Economics*, 61, 1–17.

Krishna, Ray 1997: Antidumping in law and practice. World Bank Policy Research Working Paper No. 1823.

Levinsohn, James 1996: Competition policy and international trade. In Bhagwati, Jagdish and Robert Hudec (eds.), *Fair Trade and Harmonization*, Cambridge: MIT Press.

Lindsey, Brink 1999: The U.S. antidumping law. Rhetoric versus reality. CATO Trade Policy Analysis Paper 7.

Lindsey, Brink and Daniel J. Ikenson 2001: Coming home to roost. Proliferating antidumping laws and the growing threat to U.S. exports. CATO Trade Policy Analysis Paper 14.

Mastel, Greg 1998: *Antidumping Laws and the U.S. Economy*. Armonk, N.Y.: M.E. Sharpe.

Messerlin, Patrick 1990: Antidumping regulations or pro-cartel laws? The EC chemical cases. *World Economy*, 13, 465–92.

Messerlin, Patrick 1996: Competition policy and antidumping reform: an exercise in transition. In Schott, Jeffrey J. (ed.), *The World Trading System: Challenges Ahead*. Institute for International Economics: Washington DC.

Moore, Michael O. 1992: Rules or politics? An empirical analysis of ITC anti-dumping decisions. *Economic Inquiry*, 30, 449–66.

Niels, Gunnar 2000: What is antidumping policy really about? *Journal of Economic Surveys*, 14, 467–92.

Niels, Gunnar and Adriaan ten Kate 1997: Trusting antitrust to dump antidumping, abolishing antidumping in free trade agreements without replacing it with competition law. *Journal of World Trade*, 31, 29–43.

OECD 2000: *International Options to Improve the Coherence between Trade and Competition Policies*. Trade Directorate, COM/TD/DAFFE/CLP(99)102/FINAL.

Prusa, Thomas J. and Susan Skeath 2002: The economic and strategic motives for antidumping filings. *Weltwirtschaftliches Archiv*, 138, 389–413.

Schott, Jeffrey S. 1994: *The Uruguay Round: An Assessment*. Washington, DC: Institute for International Economics.

Shy, Hyun J. 1998: Possible instances of predatory pricing in recent US antidumping cases. In Lawrence, Robert Z. (ed.), *Brookings Trade Forum: 1998*, Washington DC: Brookings Institution Press.

Staiger, Robert W. and Frank A. Wolak 1992: The effect of domestic antidumping law in the presence of foreign monopoly. *Journal of International Economics*, 32, 265–87.

US House, 107th Congress, 1st Session 2001: Concurrent resolution: whereas members . . .. *H. Con. Res.* 262.

Viner, Jacob 1923: *Dumping: A Problem in International Trade*. Chicago: University of Chicago Press.

World Trade Organization (WTO) 1998: *Communication from the US*, document WT/WGTCP/W/88.

World Trade Organization (WTO) 1999: *Communication from Japan*, document WT/GC/W/240.

World Trade Organization (WTO) 2001: *Doha WTO Ministerial Declaration*, document WT/MIN(01)/DEC/1.

Zanardi, Maurizio 2004a: Antidumping law as a collusive device. *Canadian Journal of Economics*, 37, 95–122.

Zanardi, Maurizio 2004b: Antidumping: what are the numbers to discuss at Doha? *The World Economy*, 27, 403–33.

14

Trade and the Globalization of Patent Rights

Rod Falvey, Feli Martinez, and Geoff Reed

Chapter Outline

This chapter examines the role of patent policy in the open economy. It begins by considering how the presence of patents affects trade in patentable products. A brief review of the general argument for patent protection is followed by consideration of the elements that comprise a patent system, and which determine "patent strength." Attention is then given to how the existence of the global market influences countries' choices of patent system. It is noted that the way patents are applied tends to push countries toward extreme choices, which may partly explain the pressure for some degree of international harmonization that led to the minimum standards specified in the TRIPS Agreement. But even with these minimum standards, countries retain discretion over important aspects of their patent systems.

1 Introduction

The argument underlying public policy intervention to protect intellectual property rights (IPRs – mainly patents, trademarks, and copyrights),[1] is that, without such protection, competitive market systems fail to provide private agents with sufficient incentives to undertake investment in new ideas and information that have "public good" attributes. Where "imitation" has lower costs than "innovation," imitating firms will gain an advantage over innovating firms, unless the latter can restrict the use of their innovation. We would expect limited investment in research and development (R&D) in such circumstances, and this provides the argument for strong IPR protection. On the

other hand, once such an innovation is made, its public good character suggests that welfare is maximized if its use is limited as far as possible, providing an argument for weak IPR protection. Choice of IPR policy then reflects a balancing of these arguments.

IPRs are territorial rights, conferred by national governments and valid only within the relevant jurisdiction. Indeed, until the early 1990s, the economic literature mainly addressed normative concerns on Intellectual Property (IP) systems from an industrial organization perspective with a focus on a closed economy. Similarly national IPR systems largely focused on what was perceived to be in the best interests of the country concerned. Thus (developed) countries with many potential innovators had relatively strong IPR systems, while (developing) countries with few potential innovators had relatively weak systems. Over time there has been a general strengthening and broadening of IPRs in developed countries.[2]

The globalization of the international economy has also had an impact on IPRs. First, the growing importance of international markets for patented goods has led to pressure from developed country innovators that the same or similar levels of protection be provided in foreign markets as are provided at home. Second, the success of trade liberalization has strengthened concerns that cross-country differences in the way IPRs are established and enforced could lead to non-tariff barriers to trade. Indeed the use of quantitative restrictions to protect IPRs is one of the limited exceptions allowed in the General Agreement on Tariffs and Trade (GATT) (Article XX(d)).

There have been international agreements on IPRs since the nineteenth century. Until recently the main instruments of international law regarding the substantive protection of IPRs were the Paris Convention for the Protection of Industrial Property (1883) and the Berne Convention for the Protection of Literary and Artistic Works (1886). More recently concerns expressed by the United States over international protection of IPRs led to their inclusion on the agenda of the Uruguay Round launched in September 1986. The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) of 1994, which provides *minimum* standards on IPRs for World Trade Organization (WTO) members, was the outcome.

As the diversity of IPRs inhibits a coherent discussion of all of them, in this chapter we restrict attention to patents, examining the role of patent policy in the open economy. Section 2 concentrates on how the presence of patents affects trade in patentable products. A brief review of the general argument for patent protection is followed by consideration of the elements that comprise a patent system, and which determine "patent strength" (i.e., the extent of imitation permitted by patent law and the way in which it is enforced). Firms producing and selling in a world of uncoordinated national patent systems must decide in which markets to produce, to which markets to export, and, correspondingly, in which countries to patent. These decisions will be influenced by the characteristics of the national patent systems, and in particular, the extent to which they provide a barrier to international arbitrage by other economic agents (parallel imports).

Section 3 switches focus, and considers how the existence of the global market influences countries' choices of patent system. It begins by noting that the way patents are applied tends to push countries toward extreme choices, which may partly explain why we observed developed countries having such strong systems and developing countries such weak ones. Hence the pressure for some degree of international harmonization

that led to the minimum standards specified in the TRIPS Agreement. But these standards still leave countries significant discretion over important aspects of their patent policy. In particular, they do not restrict a country's choice of principle of exhaustion, and hence allow choice over whether to allow parallel imports. We therefore review the economic arguments for and against parallel imports, from a global and from a national perspective. The section concludes with a brief discussion of issues for the future. The final section provides a summary of the chapter.

2 The Effects of Patents on Trade

A patent gives the patent holder a (limited) monopoly in the use of the patented innovation. This allows the patent holder to constrain the production and sale of products that use that innovation, but only where a patent has been obtained and to the extent specified by patent law. Elsewhere the innovation is in the public domain, available to any who are able to use it. Because patenting is costly, innovations are rarely patented in all potential markets. As a result patents present a barrier to trade in products containing a patentable innovation. This is obviously the case between markets where the innovation is and is not patented, but can also be the case, depending on how the patent laws are specified, between markets where parallel patents exist.

2.1 The Economic Role of Patents

IPRs protect the application of ideas and information giving rights to the owners, inventors, or authors. As such, IPRs are legal instruments to secure markets, and their exact content and the scope of the protection that they offer are legally determined and have evolved over time, and often differently, across jurisdictions. The main characteristic of all types of IPRs is that the rights granted are essentially negative. IPRs do not confer the right of exploitation of goods or services embodying IPRs, but rather give the IPR holder the right to prevent others from exploiting them without its license so as to guarantee the IPR holder a monopoly in the exploitation of its rights.

Even though the origin of patents can be traced back to the fifteenth and sixteenth centuries in some countries,[3] with the Crown granting a monopoly in its jurisdiction to traders and manufacturers for the exploitation of their inventions, it was only in the eighteenth and nineteenth centuries that the main features of modern patents took shape. Developments in these centuries included the requirement of sufficient disclosure in the description of the invention, provision of adequate patent length taking into account the interest of innovators, imitators, and consumers, extending patent protection to a wider range of innovation and improvements in the enforceability and administration of the patent system. These features nevertheless vary widely across countries, differences being more acute between developed and developing countries.[4]

To illustrate the market failure argument for granting patent protection, we consider the simplest case where R&D in a particular activity has resulted in the discovery of

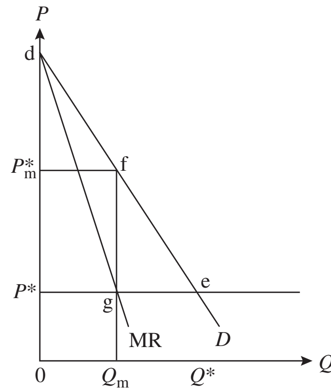

Figure 14.1 Patent protection under autarky

a new product.[5] In Figure 14.1, D is the demand curve for this product. Once the new product is available, social optimality requires that it be sold at a price equal to its marginal cost of production (MC) at point e. For simplicity we assume no fixed costs of production. From the equilibrium output Q^* that would result, society would derive maximum social surplus dP^*e. This solution would however provide each firm, including the innovator, with zero economic profits. Consequently, since R&D is costly, no firm will have an incentive to innovate (undertake R&D) as it could never recover its R&D outlays.

The usual solution to this problem is to grant the innovator a patent, thereby rewarding the development of new products by giving the innovator (or a licensee) the right to charge a monopoly price for the product for a limited period of time. Monopoly would result in a higher price P^*_m and a lower output Q_m. This outcome generates a deadweight loss relative to the competitive solution of area fge.[6] But compared to the situation of no innovation, society obtains a net gain of monopoly profit (equal to rectangle P^*_mfgP*), plus the remaining consumer surplus (triangle dfP^*_m), minus the associated R&D cost. Once the patent expires the competitive solution prevails. An optimal patent system will induce the development of all innovations for which gains exceed the R&D costs.[7] For those products that would be developed and sold anyway, the availability of patent protection simply reduces consumer surplus and raises monopoly profits, which nevertheless may encourage additional innovation.

Of course the direct rewarding of successful R&D activities is not the only role for patents in encouraging innovation. A patent may also help induce inventors to disclose their innovations, thus encouraging their wider use and placing the relevant knowledge (if not its use) in the public domain for the benefit of other inventors. Ownership of a patent may help inventors to obtain finance for the further investment necessary to develop and commercialize their innovation or allow others to undertake these activities, with their approval. Patents may also allow the "orderly development" of a new field in which there are likely to be many related innovations, by avoiding wasteful "patent races" over follow-on innovations (Mazzoleni and Nelson, 1998).[8]

2.2 Patent Strength

In this simple illustration, the only policy decision to be made is whether to grant a patent of a certain length or not. But in practice the design of a patent system involves the selection of a range of patent characteristics that have been referred to collectively as *patent strength*. These characteristics include: (1) *patent length* – the appropriate duration of the period for which the innovator is given the monopoly (Nordhaus, 1969; Scherer, 1972); (2) *patent breadth* – the range of applications reserved for the innovator. Many innovations have potential application in a variety of product markets, where existing firms (often including the innovator) are already producing. In general it will not be optimal to reserve all these potential applications for the innovator (Klemperer, 1990; Gilbert and Shapiro, 1990); and (3) *patent height* – the extent to which the patent for today's innovation, which may be the basis for subsequent innovations, also controls future improvements and developments (van Dijk, 1994).

The strength of a patent system depends not only on how the law is specified[9] but also on how it is administered. Patents are usually granted following a procedure involving a search and examination to confirm that the invention fulfils a set of patentability requirements. The invention must be, *inter alia*, *new* (i.e., not part of the existing state of the art), and involve an *inventive step* (i.e., not be obvious to a person skilled in the art). This granting procedure is open to interpretation, leaving countries some freedom to determine its scope and stringency, which may result in wide cross-country variations. Sometimes, national patent offices operate as mere registration offices, granting patents to applicants who submit applications fulfilling minimum formal requirements (e.g., filing a patent application and the payment of a filing fee) without any further substantial examination.[10] But a granted patent only entails a presumption of validity and the mere existence of a patent does not prevent its infringement, it being the patentee's decision whether or not to attempt to enforce a patent. Often patent infringements are not brought into court because litigation costs are too high, the administration of justice is inefficient, it seems likely that the patent will be found invalid, or the parties involved reach a private settlement.[11]

Countries therefore have to choose values for a range of characteristics in setting the strength of their patent systems. As we note below, this implies that countries acting in isolation would and could choose to have different patent strengths, depending, *inter alia*, on the optimal balance for them between the costs of innovation and the costs of temporary monopolization. Producers of patentable products must cope with an array of patent systems if their product is to be protected from imitation in the international market.

2.3 Patenting Strategies

Tradable products can be sold in many national markets. A firm selling a product containing a patentable innovation may wish to exercise the option of producing in some locations and exporting to others. Clearly the degree of competition this firm faces in each market will depend on whether a patent applies in that market, how strong

it is, and which firm holds it. In the absence of a universal patent, an innovator must decide in which markets to patent. This turns out to be a separate decision for each national market, a consequence of the two main principles governing the international trade in goods covered by IPRs as established in Article 4 (bis) of the Paris Convention and Article 5 of the Berne Convention. These are the principles of *territoriality* and *independence*.[12]

The principle of territoriality implies that the extent of protection of IPRs, their establishment, and transfer, as well as the means of redress afforded to the IPR holder, shall be governed exclusively by the laws of the country where protection is claimed. It also implies that the effect of an IPR is only valid within the boundaries of that country. Due to the principle of territoriality the strength of IPRs is nationally determined, and so differences in national IPR regulations can lead to cross-country variations in their strength. An innovator will enjoy protection only in those countries where a patent has been obtained. In other countries the invention remains in the public domain and is used freely.

The principle of independence implies that patents applied for in various countries shall be independent of patents obtained for the same invention in other countries in terms of establishment, transfer, and expiration. Therefore, the nullity, forfeiture, duration, and the like of a patent in country A do not impact on a dual patent in country B.

An innovator does have the advantage of the *principle of priority* (Article 4 of the Paris Convention), which applies to industrial property rights granted upon registration (e.g., patents and registered trademarks). This principle allows an applicant to file in signatory countries external patent or trademark applications on the basis of a first application in one country and demand recognition in those other countries over a *priority year*, during which the applicant decides its optimum strategy. Under this priority registration system, if a patent or trademark application is filed in one member country, nothing which occurs within a twelve-month period that runs from the first filing will affect the right to a patent or trademark in other member countries (the patentability requirements will be determined by the *priority date*, which is the date of the first filing). After the priority year and once the invention has been disclosed in the public domain, anyone is free to copy, produce, and market it, wherever patent protection has not been applied for. The patentee, however, can enforce any of its independent patents in each corresponding independent national court to prevent production in or importation to the protected markets.

Because of the principles of territoriality and independence, it is possible that an innovation could meet the conditions for patentability in one national market but not in another, and that even where patents are granted in different markets they will not be of the same strengths. We can consider the issues that arise when patents may or may not apply in two markets by returning to our earlier simple illustration. Figure 14.2 adds Country B to the analysis of the welfare economics of patent protection. The left panel shows the market in Country B for the new product developed in Country A from Figure 14.1. For convenience, demand in Country B (D_b) is shown as more elastic than demand in Country A (D_a) at any common price.

Suppose that the innovator holds a patent in Country A and hence sells the new product in that country at the monopolist equilibrium price of P_a^*, but that Country B provides no effective patent protection at all. Assuming that imitation is costless and

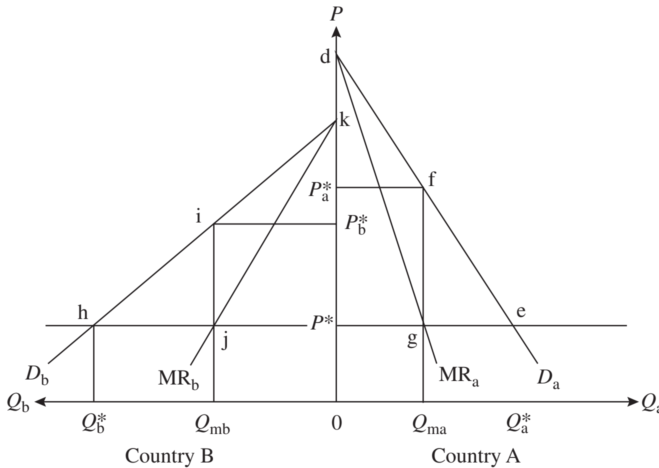

Figure 14.2 Welfare economics of product patents in a two-country model

immediate, and imitators can produce a perfect substitute using the same technology as the innovator, the innovator would have to sell the new product in Country B at its marginal cost of production at point h. From the equilibrium output Q_b^* that would result in Country B, society in B would derive maximum consumer surplus kP*h.

This solution provides zero economic profit for the innovator in Country B, but if the profits in Country A cover the R&D costs, this should not prevent the innovation. However, monopolist profits in Country B could provide additional incentives to develop new products that otherwise would not cover R&D costs.

If Country B provided patent protection, however, and if the innovator chose to take out a parallel or equivalent patent in Country B, it would in principle also enjoy a monopoly in Country B with equilibrium at point i. This outcome corresponds to the standard monopoly price-discrimination model, with the parallel patents generating the monopoly prices and segmenting the markets. The innovator would consequently charge a higher price (P_b^*) in Country B than would arise under the competitive solution. With identical marginal costs of production, the monopoly price is lower in Country B than in Country A because we have assumed a more elastic demand in Country B.

The innovating firm clearly earns greater gross economic profits from the sale of its product if it patents in all markets. But patenting in each single country of the world would be extremely costly, and very seldom, if ever, is patent protection enjoyed worldwide. Even where there is a patent in the innovator's home market, external applications involve high additional costs, mainly in terms of extra fees to the foreign patent offices, external patent agent fees, translation costs, and potential litigation costs to defend the external patents in the event of infringement. The usual strategy is to apply for patent protection in those countries where the applicant intends either to manufacture or market the invention and where the risk of imitation is high.[13]

In addition to procedural fees, annual renewal fees, and their changing rates over the length of the patent across countries influence the overall patenting strategy of the patent holder.[14] These schedules are often progressive, which, combined with value

of the patent decreasing over time as its rivals imitate, suggests that governments use patent renewal fees as a mean of shortening the effective length of patents.[15] If the patentee stops paying renewal fees in a given country, it loses its right in that country.[16] Thus firms tend to patent only in major markets, accepting the loss from the innovation being in the public domain in the others. Which markets they patent in will depend on the size of the market and the number of potential imitators based there.[17]

Another factor that decreases the value of patents in a given market is the ease of availability of *compulsory licenses*. As applied to patents, a compulsory license is an involuntary contract between a patent holder and a third party authorized by the national authorities that entitles the licensee to exploit the patent for a fixed period of time during the patent life, upon the payment of a reasonable remuneration to the right holder. They are usually granted on the grounds of public interest, dependency,[18] insufficient exploitation of the patent, or to remedy anticompetitive practices. Firms taking out a patent in a country that grants compulsory licenses will know that its attempts to restrict supply to exploit its monopoly power could be undermined in this way.

2.4 Parallel Imports

If dual patents[19] allow the patentee to price discriminate internationally then the profits of the innovating firm will be maximized by charging different prices in different markets. If these price differences exceed the costs of shipment between markets then profitable arbitrage opportunities arise for third parties if they ship legally sold products in the low-priced market for resale in the high-priced market. This phenomenon is known as *parallel importing*.

Parallel (or gray market) trade in goods covered by IPRs occurs when those goods that were intended for sale in one national market (say B) are exported or re-exported by a third party without the permission of the original authorized seller (normally the IPR owner, subsidiary, or licensee) from their original destination to another country (say A) with *parallel or corresponding IP rights* – for example, equivalent patents or industrial designs – exploiting price differentials. These products are identical to legitimate products sold in A except that they may be packaged differently, and may not carry the original manufacturer's warranty if the latter were restricted to apply only in B.[20]

The legality of parallel imports depends on which of the three approaches to the *principle of exhaustion* is adopted by the importing country. The principle of exhaustion (or "first-sale" doctrine) sets limits on the IPR holders in the exercise of their rights to control distribution of goods that have been placed on the market (first sale), either by themselves or by a third party with their consent. When a "first sale" takes place the IPR is considered to have served its purpose and therefore is exhausted. Hence, the effect of the IPR cannot be extended to acts such as use, assignment, or lease of the product covered by the patent in question, with the patent holder losing rights to any royalty after the initial sale.

The three approaches to the principle of exhaustion differ on the first-sale concept applicable. These approaches are as follows.

(a) Territorial, national, domestic, or narrow exhaustion. For this regime the relevant first sale has to be on the domestic market. Only if the goods covered by

an IPR are placed on the domestic market by the IPR owner or by a third party with his consent (a subsidiary, an affiliated company, or a licensee) can the IPR no longer be enforced within the domestic market. The IPR holder can block the importation of goods first sold abroad. Adhering to the territorial approach therefore has the consequence of partitioning international markets, which fully allows for international price discrimination by providing absolute import rights to the IPR holder. Any attempt by any trader to introduce commodities from a cheaper source covered by a parallel right in the importing country can be curbed by the IPR holder enforcing its right in the higher-price country.

(b) International or broad exhaustion. Under this approach a first sale anywhere on the relevant international market[21] (by the IPR holder or his authorized agent) will preclude any attempt by parallel IPR holders to enforce their parallel IPRs to prevent an arbitrageur from marketing those goods in the patentee's domestic market. International exhaustion can be regulated with caveats or safeguards, such as not considering IPRs as exhausted when the first sale has been subject to national price controls or to compulsory licenses forcing sales at a fixed price, or when the first sale takes place in a country without parallel rights, or when parallel IPR owners have no legal relationship (e.g., the case of simultaneous innovation or sale of parallel rights).

(c) Regional exhaustion. This approach can be chosen by members of a regional trade arrangement (RTA) such as the European Union (EU). In the absence of RTA-wide harmonization of IPRs, a balance must be struck between the protection of national IPRs and the free movement of goods. The outcome is a combination of the two approaches described above. International exhaustion applies within the integrated area, whereas territorial or RTA exhaustion applies with respect to countries outside. Under this regime, IPR holders in any country of the RTA would be legally barred from invoking their rights to prevent the parallel importation of goods that have been placed anywhere in the RTA by the IPR owner itself or a third party with its consent. However the first-sale concept is limited to commercial activities taking place within the RTA. Therefore, holders of IPRs in any country of the RTA can invoke their rights to prevent the parallel importation of goods coming from a country outside the RTA. Again, there can be various regimes of exhaustion, subject to different caveats, within the regional approach.

In the absence of an international rule governing the principle of exhaustion (Article 6 of TRIPS[22]), all WTO members are free to choose one among the three broad regimes of exhaustion pursuant to their own law.[23]

2.5 Summary

IPRs evolved initially as *private rights* under national law as part of a country's industrial policy, and therefore national differences in IP systems arose reflecting cross-country differences in economic development. Initially, when production was concentrated in one market, IPR protection in that market secured for the IPR holder control over the global market, as it could have full control over production and trade. However, the globalization of markets, the spread of multinationals with production capacities in different jurisdictions, and the growing recognition of the importance of technology and

international knowledge spillovers for economic growth, has brought the international dimensions of IPRs into prominence. In this section we have highlighted those features of patent systems important for trade in patentable goods – patent strength and parallel imports in particular.

Do strong IPRs encourage or discourage trade? In principle, the answer depends on the balance of two forces (Maskus and Penubarti, 1995): (a) the market power effect – strong IPRs give foreign patent-holders greater potential market power, which they can exploit by restricting supply (exports) to raise price; and (b) the market expansion effect – strong IPRs shift the demand curve facing the foreign patent-holders outward, since they reduce imitation, which tends to increase their exports. Smith (1999) argues that weak IPRs may discourage foreign patent-holders from exporting if they suspect that exporting raises the probability of imitation.[24]

3 The Effects of Trade on the Choice of Patent System

It was noted above that, prior to the TRIPS Agreement, countries could choose the characteristics of their patent systems independently, and that differences in patent strength across countries had implications for trade in patented goods. Here we switch perspective and examine the implications of trade in patented goods for a country's choice of patent system. We begin by investigating whether, given the characteristics of patent coverage, the possibility of trade in patented products would tend to generate divergence or convergence of nationally chosen patent systems, relative to those chosen in autarky. We find that divergence is the likely outcome.

3.1 The Choice of Patent Strength

In choosing the "strength" of its patent system a country is trading off the encouragement of innovation that comes from a strong patent system against the benefits of the wider dispersion of that innovation (through imitation) that comes from a weak patent system. Here we investigate how this trade off, and therefore the final policy choice, is affected by the fact that a country acts not in isolation but as part of the global trading system.[25]

Consider, first, a closed market where there are several firms each selling a differentiated product. Suppose one firm innovates (either improving its product quality or reducing its production costs) in a way that can potentially be imitated by its competitors. If the innovating firm patents its innovation, then the extent of (legal) imitation by the other firms will be determined by the strength of the relevant patent system. The stronger the patent system, the lower the level of imitation and the higher the profit gain to the innovating firm. Stronger patent systems therefore lead to higher levels of innovation. But the welfare gain will depend on the average level of product improvement (averaged over innovator and imitators) in this economy. While high rates of innovation raise this average level of product improvement, low rates of imitation reduce it. In general some intermediate level of patent strength will be optimal, depending on the costs of innovation and imitation and the number of potential imitators.

In a global market, each country's decision on its optimal patent strength is influenced by the decisions of others. All countries have an incentive to encourage innovation wherever it occurs, but they also gain from higher levels of imitation in the products that they purchase. Given different balances of interest, however, one would expect countries with different characteristics to choose patent systems of different strengths. Interestingly, this diversity is amplified by the way in which one country's choice of patent strength impacts on other national markets. Suppose there are two countries. Each country's decision on its patent strength affects not only the extent of permitted imitation on sales in its market by domestic producers and foreign exporters, but also the extent of imitation by domestic producers for export. This is because a patent allows the patent holder to prevent, for a limited period of time, the unauthorized *use, sale, offer for sale, importation, or manufacture* of goods embodying the patented subject matter within the territory for which protection is granted, whether these goods are imported or locally produced (Articles 27 and 28 of TRIPS). Thus a government's choice of patent strength affects not only domestic production for domestic sale, but also *imports and exports*. Traded goods are subject to the patent laws of both the importing and exporting countries. Indeed the level of imitation embodied in traded goods can only be the *minimum* permitted in the two trading partners.

To see the implications of this aspect of patents, we first suppose that domestic patent strength applies only to *sales* on the domestic market. Traded goods are then constrained only by the patent strength of the country of destination. Each country chooses its sales-patent strength taking that of the other country as given, with producers allowed different levels of imitation for domestic sales and exports, if the technology permits. All products for sale in the country with the weaker patent system would be likely to contain a higher level of imitation than those for sale in the other. Each country independently balances the (global) incentives for innovation against the average level of product improvement in its (national) market. Patent strengths will be weaker than if this decision were made cooperatively, because each country ignores the benefits of higher levels of innovation for its trading partners. Two identical countries will tend to choose identical sales-patent strengths. If countries differ in market size, the larger country will tend to have the stronger sales-patent, but the average global sales-patent strength may be little different.[26]

In practice, however, traded goods are subject to the patent laws of both countries. If the innovating firm takes out a patent in one country, then another firm producing in that country can, without violating the patent, legally produce and sell a product improved (by imitation) to the extent allowed by the patent law in that country. This firm can also export this product to all countries with weaker patent systems, where the innovating firm has taken out a patent, as the improved product clearly does not infringe the patent in these countries. But if the imitating firm attempted to export its improved product to a country that has a stronger patent system and where the innovator has taken out a patent, its product, though legal in the source country, would infringe the patent in the destination country.

Returning to our example of two identical countries, say A and B, we can distinguish four types of products — those produced in A for sale in A (AA), those produced in A for sale in B (AB), those produced in B for sale in A (BA) and those produced in B for sale in B (BB). Where countries are identical in terms of numbers of firms and

market size, the volume of sales of each type will be approximately equal. Suppose each country chooses its optimal patent strength taking the other's as given. If B's patent strength is relatively weak, A will tend to choose a stronger patent system, as otherwise the (global) incentive for innovation is too low. Then A's patent strength will completely determine the level of imitation for sales in its market (i.e., AA and BA), and half the sales in B's market (AB). As a result, A's choice of patent strength has the major impact (in fact three times as much) on the incentives for innovation, and A will choose a stronger patent system than it would, say, if B had the same patent strength and patents only applied to domestic sales. Alternatively, if B's patent strength is relatively strong, then A will tend to choose a weaker patent strength than B, and indeed a weaker patent strength than it would if patents applied only to domestic sales. The outcome is that *even two identical countries* would not find it optimal to choose identical patent strengths. The only (pure strategy) equilibria involve one country with a strong patent system and the other with a weak one. The country with the weaker patent system has the higher level of imitation, the higher level of average product improvement, and is consequently the better off of the two.

This tendency to choose patent systems of different strengths will continue where countries differ in market size and numbers of potential imitators.[27] Relatively small countries in terms of either criterion will tend to choose weaker patent systems. In these circumstances it is hardly surprising that we observe a global economy divided into two groups of countries. Developed economies, with large markets and where almost all of the innovation takes place, have relatively strong IPR systems in general. Developing countries have relatively weak systems, or none at all.[28] Almost all innovation is concentrated in developed countries, which already provide patent protection. If failure to provide such protection means that developing countries obtain those innovations of importance to them at competitive prices anyway, then they have no incentive to extend patent protection themselves. Any gain in consumer surplus from the few additional innovations that their providing protection would generate, would be more than offset by the losses from the monopoly pricing of those they otherwise would have obtained at competitive prices (Deardorff, 1990, 1992).[29] Conversely, developed countries would unambiguously gain from such an extension, so it is not surprising that the IPR holders in developed countries, particularly in the United States, have sought to extend developed country strengths of IPR protection globally.

3.2 International Agreements on Minimum Patent Strengths

Prior to the TRIPS Agreement, the World Intellectual Property Organization (WIPO), which administers, among others, the Paris Convention and the Berne Convention, provided the main international forum for the discussion of the legal and administrative aspects of IP protection and the harmonization of IP law. However, these conventions lacked enforcement mechanisms and were mainly ratified by developed countries. In the late 1970s, developed country patent-holders attempted to amend the Paris Convention to establish minimum levels of patent protection worldwide. But

the WIPO is a UN agency, and developing countries, which naturally opposed the changes, which they saw as not in their interests, had the votes to defeat them.

US firms with strong patent interests (e.g., pharmaceutical manufacturers) then combined with firms with strong copyright interests (e.g., publishing, motion pictures, recording, and software designing), all of whose products could be imitated or copied relatively cheaply, and successfully lobbied for US diplomatic support. This support took two main forms. One was bilateral action under section 301 of the Trade Act of 1974, as amended by the 1984 Trade and Tariff Act and the 1988 Omnibus Trade and Competitiveness Act. The 1984 amendment added failure to provide adequate IP protection to a list of "unreasonable practices," which could generate US Trade Representative (USTR) investigation and subsequently the imposition of trade sanctions. This was strengthened by the 1988 amendment, which added the Special 301 process, which required the USTR to name those foreign countries with the least adequate and effective protection of IPRs, and to announce an agenda for intervention (Stegeman, 2000). This bilateral action was accompanied by pressure for the inclusion of protection of IPRs on the agenda for the Uruguay Round of GATT negotiations, which were launched in September 1986. Other GATT members initially opposed this extension beyond traditional trade issues,[30] but conceded when it became clear that their inclusion was necessary for US participation.

The inclusion of IPRs in the Uruguay Round negotiations was important for the successful conclusion of TRIPS, because it allowed concessions between IPRs and other issues. Given the failure of the attempt to reform the Paris Convention, it was unlikely that if TRIPS had stood by itself it would have been accepted by the developing countries, although these countries were beginning to realize that the alternative would have been pressure (under section 301) to conclude bilateral agreements with the United States, to ensure continued access to the US market.[31] But the United States and the European Union never intended to make any serious concessions concerning their own IPR laws, so there was no question of reciprocity within the TRIPS. Instead the trade-off was better protection for IPRs, hence greater rents to developed country IPR holders in developing countries, in exchange for improved and more secure access for developing country products in developed country markets. There was no noticeable opposition in the United States or the European Union to an agreement that would force the rest of the world to strengthen its protection of IPRs. This lack of opposition is consistent with the view that consumers in developed countries had limited access to imitated products anyway, and that multinationals from developed countries felt they had more to gain from stronger IPR protection in developing countries than from taking advantage of their weaker systems to imitate in and for those markets in competition with developing country firms.

The TRIPS is a multilateral agreement establishing unconditional obligations on all WTO members, prescribing minimum standards for protection of IPRs, including enforcement and due process requirements. At the same time the Agreement allows enough flexibility in the design of IPR systems to encompass the different public policy objectives underlying WTO members' national systems, which include developmental and technological objectives (Preamble of the TRIPS Agreement, Recital 5).

These minimum standards apply to the availability, scope, and use of IPRs in copyrights and related rights, trademarks, geographical indications, industrial designs,

patents, layout-designs (topographies) of integrated circuits and protection of undisclosed information.[32] The TRIPS Agreement recognizes the applicability of the basic principles of GATT, including most-favored-nation (MFN) treatment[33] and national treatment,[34] and of the relevant international IP agreements and conventions. A significant aspect of the inclusion of TRIPS in the WTO package, rather than it being a stand-alone agreement, relates to dispute settlement. Under the WTO Understanding on Rules and Procedures Governing the Settlement of Disputes, members whose rights have been infringed under one agreement may be authorized to suspend obligations under a different agreement. This is likely to be important since offending countries are likely to have few patented innovations of national origin.

With respect to patents specifically, the TRIPS Agreement provides, "... patents shall be available for any inventions, whether *products* or *processes*, in *all fields of technology* ... patents shall be available and patent rights enjoyable *without discrimination* as to the *place of invention*, the *field of technology* and whether products are *imported or locally produced*" (Article 27.1). This increases the scope of patent protection available to any innovator in any WTO member, particularly developing countries (which have traditionally barred certain classes of products from patent protection).[35]

The fact that the TRIPS Agreement provides new international *minimum* standards for protection and enforcement of intellectual property law implies that no WTO member is obliged to provide for higher standards, although individual members are free to do so. It is recognized that members' choices will depend on their current level of development and technological objectives. Developing countries[36] still have discretion over key elements of "patent strength," particularly during the transition period (January 1, 2005, 2006 for the least developed countries). For example, before TRIPS developing countries traditionally provided for shorter periods of patent protection than the *minimum* of 20 years from the filing date specified in the Agreement. However expiration dates of parallel patents across countries may still differ due to the delayed implementation of the TRIPS obligation according to the transitional arrangements by developing countries (shorter periods of patent protection may last until 2005), and to individual states providing for longer periods of patent protection than 20 years. Even after the entry into force of the TRIPS Agreement countries still retain control over the cost of patenting, the stringency of the patentability requirements, the patentable subject matter (though discretion here has been reduced as noted above), patent enforcement in their judicial systems, and the exercise of patent rights through their competition policies.[37]

3.3 Parallel Imports

One area where TRIPS might have adopted common international standards but did not, is in limiting the choice over principle of exhaustion so as to constrain parallel imports. That TRIPS allows each country to choose its own regime with respect to parallel imports was important for its acceptance by many developing countries,[38] as few developing countries restrict parallel imports. As noted above, parallel imports arise where international price differences generate arbitrage opportunities for third parties. Such price differences can arise from several sources (Maskus, 2000). In some

cases the efficient distribution of a product may require the patent holder to retain some control over activities in the distribution chain (i.e., through exclusive dealerships), in order to control product quality and to ensure that wholesalers and retailers are rewarded for undertaking appropriate investments in marketing and service provision. Parallel imports are a channel through which unlicensed distributors can free-ride on some of these investments. In some countries it may be difficult to enforce a private contractual prohibition on sales outside the authorized distribution channel. Restrictions on parallel imports are then seen as a natural extension of the right to control distribution, although they may also permit collusion among exclusive dealers (Maskus and Chen, 1999).

Arbitrage opportunities may result from national price regulations established to achieve particular social objectives. These result in considerable price variation across countries, and allowing parallel imports would undermine the objectives of the regulation. Otherwise, arbitrage opportunities will arise if monopoly pricing in each market leads to international price differences of sufficient magnitude. We can infer the broad welfare effects of permitting parallel imports from Figure 14.2. The price will fall in Country A and rise in Country B. The firm will be worse off, since it would have charged a uniform price if that led to higher profits, as will consumers in Country B. Consumers in Country A will gain. Global welfare may rise or fall.[39] Further, in a world of many markets, the patent holder may prefer not to sell in small markets if parallel exports are a possibility (Malueg and Schwartz, 1994). A global regime of parallel imports may then reduce welfare in developing countries through higher prices and limited availability.

If this is the case, why do developing countries oppose restrictions on parallel trade? Some believe that their domestic prices are actually higher under price discrimination. Others see beneficial opportunities to parallel export. An option available to countries concerned that patented products will not be supplied to their market (possibly because the patent holder is afraid of parallel exports), is to require compulsory licensing. As noted earlier, compulsory licenses are usually granted on the grounds of public interest, dependency, insufficient exploitation of the patent or to remedy anticompetitive practices (TRIPS Article 31).[40] The TRIPS Agreement (Articles 7, 8, and Preamble) introduces new and more expansive concepts of the abuse and public interest exceptions as prerequisites of compulsory licenses.[41]

3.4 Current Issues

One area where price regulations and compulsory licenses are important issues is in patented medicines, where the trade-off is between providing an adequate return to the R&D investment of the patent holder and providing “essential” medicines at affordable prices to the populations of developing countries.[42] It was normal practice in some developing countries to exclude food, chemical, and pharmaceutical products from patentability. The TRIPS Agreement requires that product patent protection be extended to these areas by the end of the transition period, but not necessarily before then.[43]

Some developing countries view the local manufacturing of medicines as important for the functioning and development of their domestic health care systems, and

are therefore very reluctant to allow patent holders to supply their markets through exports. While voluntary licensing or other voluntary agreements (e.g., joint ventures) have the advantage of the cooperation of the patent holder, compulsory licensing can be employed on the grounds provided in TRIPS Article 31. This will not help those countries with limited manufacturing capacity in the pharmaceutical sector, however, and these markets could remain unsupplied if the patentee patents in all other countries with manufacturing capacities, as they would face difficulties in making effective use of compulsory licensing. Other countries resorting to compulsory licensing would have to do so primarily for their domestic market (TRIPS Article 31.f),[44] and so could not help their neighbors in this way. This problem was acknowledged in the Doha Declaration, and the Council of TRIPS was instructed to find a solution, to be reported by the end of 2002.

The ongoing debate on the desirability of allowing the patenting of life forms was another issue acknowledged in the Doha Declaration, which required the TRIPS Council, in pursuing its work program, to examine *inter alia* the relationship among the TRIPS Agreement (Article 27.3b), the Convention on Biological Diversity, and the protection of traditional knowledge, while taking into account the development dimension. In the light of the divergence of views on this issue, and the difficulties faced in the Uruguay Round in negotiating even a modest agreement on minimum standards for patent law, it seems unlikely that the issue will be resolved in the near future.

4 Conclusions

Patent rights are conferred by national governments and are valid only within the relevant national jurisdiction. In economic terms, a patent grants a successful innovator a (limited) monopoly in the use of that innovation. While this serves to encourage innovation, it is recognized that permitting some "imitation" is optimal, and the parameters of the patent system can be chosen so as to achieve a desired "patent strength." National patent systems therefore differ, reflecting different balances between the encouragement of innovation and the toleration of imitation, which combine to generate the desired level of average product improvement.

Innovators then need to decide in which countries to patent. In general, innovators do not patent in all countries. Their choices will be determined by market size, the location of potential imitators, and the costs of patenting. The ability of innovators to price discriminate internationally will depend on whether such pricing can be undermined by parallel imports. This in turn depends on the principle of exhaustion adopted in the relevant patent laws.

Where countries are free to choose their patent strengths, taking those of their trading partners as given, the outcomes will depend on country characteristics. One would expect that countries with small domestic markets and few potential innovators would prefer to have relatively weak patent systems (i.e., that permit significant imitation). But as we noted above, this tendency is amplified by the way in which patent laws are applied to traded goods. Such goods are subject to the laws of both the importing and the exporting countries and hence must satisfy the stronger of the two systems. The outcome is that countries choose extreme patent strengths, and, in light of this,

it comes as no surprise to observe a global economy where developed countries have strong patent systems and developing countries have weak ones.

Naturally, such a system does not serve the best interests of innovating firms in the developed countries, or their governments. IPRs were included among the issues negotiated in the Uruguay Round, and the TRIPS Agreement was the outcome. This Agreement provides for international *minimum* standards for protection and enforcement of intellectual property law. But even with these standards, countries retain discretion over the cost of patenting, the enforcement and functioning of their judicial systems, the stringency applied in the evaluation of patentability requirements, the extent of the patentable subject matter, and, through competition policy, the exercise of patent rights.

The right of countries to choose their own exhaustion regimes was confirmed by the TRIPS Agreement. This is important for developing countries, which typically allow parallel imports, even though this is not obviously in their interests. The prospects of strengthening and extending the TRIPS Agreement appear limited, at least until the transition period is successfully completed. Then its proponents will be able to observe whether it functions in the way they intended. Contentious issues remain, particularly the extent to which countries should be able to compel a patent holder to supply pharmaceutical products domestically, and whether life-forms should be patentable at all.

Acknowledgment

The authors acknowledge the support of the Leverhulme Trust under Programme Grant F114/BF.

Notes

1 See Cornish (1996) for a detailed description of IPRs.
2 See Mazzoleni and Nelson (1998) for a discussion of these changes.
3 Dates differ by country. The earliest known English patent for invention was granted by Henry VI to Flemish-born John of Utyman in 1449, giving him a 20-year monopoly for a method of making stained glass not previously known in England. In Spain, the first patent (*Real cedula de privilegio de invencion*) was granted in 1522. The web pages of many National Patent Offices provide a brief history of patents in their territory.
4 See Rapp and Rozek (1990) and Sherwood (1997) for a rating and description of patent characteristics in developing countries.
5 Similar arguments and analysis can be used for other IPRs.
6 This welfare loss could be avoided if the innovator is provided a large enough lump sum payment and the innovation is made freely available. A large literature exists on non-patent responses to this problem, such as state funding of R&D. For an introduction to this literature see Beath et al. (1995).
7 Ideally the monopoly would continue only for the time necessary for the expected discounted monopoly profits to cover the expected R&D costs. But the broad parameters of patent policy are not usually tailored to potential innovations in any particular sector, and it is likely that whatever the decision on patent length, the outcome will be to encourage too much R&D in some sectors and too little R&D in others.
8 In fact surveys of firms usually rate patents rather low among the factors affecting their competitive position and inducing innovation, except for some sectors, including

pharmaceuticals and chemicals, where firms undertake considerable R&D (Levin et al., 1987; Roland Berg Forschungs, 1995). In some industries, first mover advantages such as being the first to set up production facilities and sales and service networks, and thereby being furthest down the relevant learning curve, are more important. In others imitation itself is also time consuming and costly or the information necessary to allow imitation can effectively be kept secret.

9 Patents protect embodiments of inventive activities, and inventions "offensive to public morals" can be excluded from patentability. There are wide cross-country variations in public moral standards and in the definition of "embodiments of inventive activities," so that inventions patented in one location may not be patentable in another. The patentability of software and biotechnology are particularly sensitive issues.

10 For example, the Spanish patent system was a registration system until 1986. Even then, the Spanish Law of 1986 only introduced an additional requirement that a search of the state of the art was to be conducted with the report of the state of the art to be published allowing two months for third parties to express their views. From 2001 an examination procedure has been introduced by which a patent examiner determines whether the patent application fulfils the patentability requirements in the light of the report on the state of the art. Still, this examination procedure is optional for the applicant but once requested it is binding. This situation is particularly interesting given that patent laws in all European countries have largely been harmonized.

11 The usual defence of the alleged patent offender is to challenge the validity of the infringed patent in the first place. If the infringed patent is held invalid, further entry does not constitute infringement. Variations in judicial systems imply that a patent may be challenged and found totally or partially invalid in one jurisdiction and not in others. To illustrate the importance of court decisions on patent validity in infringement suits, "of the 294 patents contested in all US Federal Appellate Courts between 1966 and 1971, only 89 (about 30%) were found valid" (Kintner and Lahr, 1975, quoted in Choi, 1998, p. 1250).

12 At present the only exception to the principles of territoriality and independence is the (European) Community Trademark, which, although it coexists with national trademark systems, has once granted a unitary character for the whole European community and can be granted, transferred, revoked, and expire only in respect of the whole of the EC. There is no supranational or universal patent. When multinational protection is desired, a separate patent application and granting procedure has to be pursued in each different state leading to independent national patents. See Martinez (2001) on this issue and on current proposals for a Community patent.

13 Eaton and Kortum (1996) find that for OECD (Organization for Economic Cooperation and Development) countries, a high cost of patenting relative to market size has a negative effect on the decision to patent. EC countries are likely to be targeted for patent protection, but even here coverage is far from universal. For illustration, the European Patent Convention (or Munich Convention) provides a unified granting procedure for awarding patents in states party to the Convention. When filing a "European" patent application, the applicant must indicate the states in which it wishes to obtain protection, paying a fee for each designated state. This procedure does not provide a supranational patent however. Once the "European" patent is granted the applicant receives a bundle of national patents, one for each of the designated states and hence, protection and jurisdiction of each independent patent will be circumscribed to the national territory. If the "European" patent is granted the applicant has to validate the "European" patent in each of the designated states, filing the respective translation and paying the validation fee in accordance to the requirements of each country. However, the average number of designated states – countries where the invention is protected – per "European" patent *granted* in 1999 was 7.46 out of a possible 20. In that

year, “European” granted patents designated Germany in 98 percent cases, Great Britain in 92 percent, France in 91 percent, Italy in 72 percent, the Netherlands in 52 percent, Spain in 50 percent, Sweden in 42 percent, Belgium in 42 percent, Austria in 37 percent, Denmark in 29 percent, and Portugal, Ireland, Greece, Luxembourg, and Finland in less than 22 percent of the occasions (EPO 1999a, Table 7.4, pp. 70–1). In the nondesignated countries the invention remains in the public domain, adding complexity to EU intra-industry trade. Interestingly, Guellec and van Pottelsberghe de la Potterie (2000) find that the probability that a patent will be granted by the European Patent Office is maximized when six to seven countries are designated.

14 For theoretical analyses of patent renewal fees as a patent-related policy instrument see Cornelli and Schankerman (1999) and Scotchmer (1999).

15 For instance, in 1999 renewal fees in Germany increased from 51.13 Euros in the first year to 1687.26 Euros in the twentieth year. See EPO (1999b) for further examples.

16 Cross-country variations in renewal fees may explain how equivalent patents may have different cross-country duration even if the term of protection is harmonized. Most studies on the impact of the cost of patenting on the propensity to patent omit renewal fees from their analysis. The statistical study by Sánchez-Padrón et al. (1996) focuses on these fees but the results are inconclusive on the extent to which renewal fees influence the decision to maintain patents in force.

17 For a more detailed analysis of this decision in a multi-country setting see Falvey et al. (2001).

18 Dependent patents are patents (second patents) for the development and improvement of innovations, which cannot be exploited without infringing another patent (first patent).

19 Equivalent patents in different jurisdictions, or patents over the same invention.

20 It is important to distinguish parallel, or gray, imports, which are legitimate products authorized for original sale in the exporting country, from counterfeit or pirated goods, which have not been produced with the innovator’s approval.

21 This could be defined as the set of markets where the patentee enjoys parallel protection, or as the global market.

22 Article 6 of the TRIPS Agreement states: “For the purpose of dispute settlement under this Agreement, subject to the provisions of Articles 3 and 4 above, nothing in this Agreement shall be used to address the issue of the exhaustion of intellectual property rights.” Articles 3 and 4 deal with national and most-favored-nation treatment respectively.

23 Indeed, the current regulations differ in different trading blocks (Maskus, 2000).

24 Maskus and Penubarti (1995), Smith (1999), and Fink and Primo-Braga (1999) report empirical estimates that there is a positive relationship between strong IPRs and bilateral trade flows.

25 A more formal analysis of the following discussion can be found in Falvey et al. (2002).

26 In the model used in Falvey et al. (2002) we find no change in the equilibrium expenditure-weighted average patent strength where patents apply only to sales.

27 Zigic (1998) considers process innovation in a model of two firms, an innovator in the North and an imitator in the South. Under some circumstances the South may gain and the North may lose from enforcing stronger IPRs.

28 More generally, there has been work on the implications of strengthening IPRs, in particular patents, in developing countries on innovation – Diwan and Rodrik (1991), Helpman (1993), and Mansfield (1986); on foreign direct investment and the transfer and licensing of technology – Ferrantino (1993), Maskus and Penubarti (1995), Lee and Mansfield (1996), Maskus (1997), Fosfuri (2000), and Taylor (1993); and on growth – Taylor (1994) and Gould and Gruben (1996).

29 Where innovators are concentrated in developed countries, a further problem arises in providing incentives for innovations of particular interest to developing countries.

If the resulting innovation becomes a public good freely available to all developing-country producers, then which developing countries will provide the patent protection that compensates the innovator? (Yang, 1998).

30 IPRs were one of three new issues, the others being trade in services and trade-related investment measures. The role of IPRs has always been recognized in GATT, with the use of quantitative restrictions on trade to protect IPRs as one of the limited exceptions allowed under Articles XI(1) and XX(d).

31 With TRIPS, IPR-related actions under Section 301 can continue, but they are subject to multilaterally agreed limits and rules.

32 Article 39.3 of TRIPS provides that where the submission of undisclosed data is required as a condition of patent approval, the country shall protect that data against disclosure except where necessary to protect the public, or, where disclosure occurs, take steps to prevent unfair commercial use.

33 Traditional IP Conventions were silent on the MFN principle, and consequently allowed the possibility of establishing bilateral or special agreements in IPRs with trading partners.

34 Article 3(1) of TRIPS reads: "Each Member shall accord to the nationals of other Members treatment no less favourable that that it accords to its own nationals with regard to the protection of intellectual property." GATT and traditional IP Convention provisions fundamentally differ in the interpretation of the national treatment principle. National treatment in GATT applies to *goods*, requiring equal treatment of national and foreign goods, whereas in IP conventions it applies to *persons*, and requires no less favorable treatment between national and foreign applicants (Dhanjee and de Chazournes, 1990; Meesen, 1987).

35 "In 1988 it was established that at that time pharmaceutical products were not patentable in 49, food products in 35, chemical products in 22, pharmaceutical process in 10, process for the manufacture of food in 9 and micro-organism in 9 of a total of 92 Paris Union states" (Strauss, 1996, p. 174). Pharmaceutical products and processes only became patentable in Spain in October 1992, six years after its accession to the European Union in accordance with transitional arrangements (Transitional provision 1.1 of the Spanish Patent Law 11/1986). Nevertheless, many countries upgraded their patent systems before the entry into force of the TRIPS Agreement: ". . . as of January 1995 . . . fewer than 20 developing and least-developed country WTO Members excluded pharmaceutical products from the grant of patents and virtually all provided for process patent protection" (Watal, 2001, p. 8).

36 The signing of the WTO Agreements in 1994 has generated a number of articles on policy options open to developing countries in the implementation of the WTO Agreements. See, e.g., Correa (1998), Pacon (1996), Watal (1997), and Weissman (1996).

37 Although the existence of IPRs has to be respected, the exercise of those rights can be controlled by competition policies, and the TRIPS Agreement confers considerable leeway to WTO members to decide their competition policy for IPR-related matters.

38 This freedom was confirmed by the Ministerial Conference at Doha in November 2001.

39 The results on whether global welfare is higher with or without parallel trade are ambiguous. Such a judgment is also complicated because each regime results in different distributions of benefits within and across countries. After reviewing the theory and the empirical evidence, Maskus concludes, "it is impossible to support on economic grounds either a global policy to ban parallel imports or a mandate that there be a free global regime in parallel trade. The best advice is simply to permit the status quo ante to continue, with each country or region selecting its own policy" (Maskus, 2000, p. 1283). But see also Richardson (2002). In the context of examining the "European patent," Van Dijk and Van Cayseele (1993) show that if no international price discrimination is permitted, then it is optimal for patent strengths to be harmonized across countries; while if patent strengths differ, the prohibition of price discrimination leads to suboptimal outcomes.

40 See Julian-Arnold (1993) for a general discussion of international compulsory licensing.
41 For example, the protection of public health and nutrition, promotion of technological innovation, socioeconomic and technological development, and the international transfer of technology. As Watal (2000) shows, welfare gains to the country granting a compulsory licence depend crucially on the market share gained by the licensee, royalty payments, and the elasticity of demand. Such gains have to be assessed against the risk that excessive resort to compulsory licensing could become counterproductive if it unduly inhibits innovation.
42 While this is important in some instances, one should not exaggerate the extent to which patents are likely to restrict the availability of essential drugs to developing countries. Less than 5 percent of drugs on the World Health Organization Model List of Essential Drugs are under patent protection anywhere in the world (Watal, 2001, p. 6), and approximately 95 percent of essential pharmaceuticals that are off-patent are in the public domain, and can be produced freely by either local producers or the generic industry.
43 However, they have to accept the filing of product patent applications ("mail box" applications) for pharmaceutical and agricultural chemical products from the date of entry into force of the Agreement. Once TRIPS is implemented the granting procedure for these applications will be continued, with the patentability requirements assessed as at the filing date of the mail box application. If patent protection is granted, then, the patentee will enjoy protection for the remainder of the patent term, counted from the filing date (TRIPS Article 70.8).
44 Unless a compulsory licence is granted to remedy a practice determined after a judicial or administrative process to be anticompetitive (Article 31.k). As with most of its provisions, Article 31 of TRIPS is open to interpretation. A straightforward reading implies that patent holders can supply protected markets via exports (Article 27), and therefore the mere lack of domestic exploitation of the patent is not enough to grant compulsory licensing. But this interpretation is contested. The Ministerial Conference in Doha in November 2001 concluded with a Declaration on the TRIPS Agreement and Public Health, which acknowledged the flexibility of TRIPS and clarified certain provisions. Each member has the right to grant compulsory licenses and the freedom to determine the grounds upon which such licenses are granted.

References

Beath, John, Yanni Katsoulacos, and David Ulph 1995: Game-theoretic approaches to the modelling of technological change. In Paul Stoneman (ed.), *Handbook of the Economics of Innovation and Technological Change*, Oxford: Blackwell.

Choi, Jay Pil 1998: Patent litigation as an information-transmission mechanism. *American Economic Review*, 88, 1249–63.

Cornelli, Francesca and Mark A. Schankerman 1999: Patent renewals and R&D incentives. *RAND Journal of Economics*, 30, 197–213.

Cornish, William R. 1996: *Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights* 3rd edn. London: Sweet & Maxwell.

Correa, Carlos M. 1998: Implementing the TRIPS agreement in the patents field: options for developing countries. *The Journal of World Intellectual Property*, 1, 75–99.

Dhanjee, Rajan and Laurence Bosion de Chazournes 1990: Trade related aspects of intellectual property rights (TRIPS): objectives, approaches and basic principles of the GATT and of intellectual property conventions. *Journal of World Trade*, 24, 5–15.

Deardorff, Alan V. 1990: Should patent protection be extended to all developing countries? *The World Economy*, 13, 497–509.

Deardorff, Alan V. 1992: Welfare effects of global patent protection. *Economica*, 59, 35–51.

Diwan, Ishac and Dani Rodrik 1991: Patents, appropriate technology, and north–south trade. *Journal of International Economics*, 30, 27–47.

Eaton, Jonathan and Samuel Kortum 1996: Trade in ideas: patenting and productivity in the OECD. *Journal of International Economics*, 40, 251–78.

European Patent Office 1999a: *Annual Report*, Munich.

European Patent Office 1999b: *National Law Relating to the EPC*, Munich.

Falvey, Rod, Feli Martinez, and Geoff Reed 2001: Product innovation and international patent protection. Manuscript, University of Nottingham.

Falvey, Rod, Feli Martinez, and Geoff Reed 2002: Trade, innovation and patent policy. Manuscript, University of Nottingham.

Ferrantino, Michael J. 1993: The effect of intellectual property rights on international trade and investment. *Welwirtschaftliches Archiv*, 129, 300–31.

Fink, Carsten and Carlos A. Primo-Braga 1999: *How Stronger Protection of Intellectual Property Rights Affects International Trade Flows*, World Bank Report 2051.

Fosfuri, Andrea 2000: Patent protection, imitation and the mode of technology transfer. *International Journal of Industrial Organization*, 18, 1129–49.

Gilbert, Richard and Carl Shapiro 1990: Optimal patent length and breadth. *RAND Journal of Economics*, 21, 106–12.

Gould, David and William Gruben 1996: The role of intellectual property rights in economic growth. *Journal of Development Economics*, 48, 323–50.

Guellec, Dominique and Bruno van Pottelsberghe de la Potterie 2000: Applications, grants and the value of patent. *Economic Letters*, 69, 109–14.

Helpman, Elhanan 1993: Innovation, imitation, and intellectual property rights. *Econometrica*, 61, 1247–80.

Julian-Arnold, Gillian 1993: International compulsory licensing: the rationales and the reality. *IDEA*, 33, 349–400.

Klemperer, Paul 1990: How broad should the scope of patent protection be? *RAND Journal of Economics*, 21, 113–30.

Kintner, Earl W. and Jack L. Lahr 1975: *An Intellectual Property Law Primer*. New York: Macmillan.

Lee, Jeong-Yeon, and Edwin Mansfield 1996: Intellectual property protection and U.S. foreign direct investment. *Review of Economic and Statistics*, 78, 181–6.

Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney G. Winter 1987: Appropriating the returns from industrial research and development. *Brooking Papers on Economic Activity*, 3, 783–831.

Malueg, David A. and Marius Schwartz 1994: Parallel imports, demand dispersion and international price discrimination. *Journal of International Economics*, 37, 167–95.

Mansfield, Edwin 1986: Patents and innovation: an empirical study. *Management Science*, 32, 173–81.

Martinez, Feli 2001: Towards a patent reform in Europe. *The Journal of World Intellectual Property*, 4, 767–86.

Maskus, Keith E. 1997: The role of intellectual property rights in foreign direct investment and technology transfer. Prepared for the conference *Public–Private Initiatives After TRIPS: Designing a Global Agenda*.

Maskus, Keith E. 2000: Parallel imports. *The World Economy*, 23, 1269–84.

Maskus, Keith E. and Yongmin Chen 1999: Vertical price controls and parallel imports. Centre for International Economic Studies Discussion Paper No. 99/21.

Maskus, Keith E. and Mohan Penubarti 1995: How trade-related are intellectual property rights? *Journal of International Economics*, 39, 227–48.

Mazzoleni, Roberto and Richard R. Nelson 1998: Economic theories about the benefits and costs of patents. *Journal of Economic Issues*, 32, 1031–52.

Meessen, Karl M. 1987: Intellectual property rights in international trade. *Journal of World Trade Law*, 21, 67–74.

Nordhaus, William D. 1969: *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, Mass: MIT Press.

Pacon, Ana M. 1996: What will TRIPS do for developing countries? In Friedrick K. Beier and Gerrick Schricker (eds.), *From GATT to TRIPS – The Agreement on Trade-Related Aspects of Intellectual Property Rights*, IIC Studies, vol. 18. Max Planck Institute for Foreign and International Patent, Copyright and Competition Law, Munich, 329–56.

Rapp, Richard T. and Richard P. Rozek 1990: Benefits and costs of intellectual property protection in developing countries. *Journal of World Trade*, 24, 75–102.

Richardson, Martin 2002: An elementary proposition concerning parallel imports. *Journal of International Economics*, 56, 233–45.

Roland Berg Forschungs 1995: *Utilisation of Patent Protection in Europe*. Representative survey carried out on behalf of the European Patent Office, Munich.

Sánchez-Padrón, M., V. Cano-Fernandez, and E. Los Arcos 1996: An analysis of EPO renewals: a first step towards patent value estimation. Some conceptual considerations. Paper presented at the 54th International Conference on *Innovation Modelling: The Patent* organized by the Applied Econometric Association, Luxembourg.

Scherer, Frederic M. 1972: Nordhaus's theory of optimal patent life: a geometric reinterpretation. *American Economic Review*, 62, 422–7.

Scotchmer, Suzanne 1999: On the optimality of the patent renewal system. *RAND Journal of Economics*, 30, 181–96.

Sherwood, Robert M. 1997: Intellectual property systems and investment simulations: the rating of systems in eighteen developing countries. *The Journal of Law and Technology*, 37, 261–370.

Smith, Pamela J. 1999: Are weak patent rights a barrier to US exports? *Journal of International Economics*, 48, 151–77.

Stegemann, Klaus 2000: The integration of intellectual property rights into the WTO system. *The World Economy*, 23, 1237–67.

Strauss, Joseph 1996: Implications of the TRIPS agreement in the field of patent law. In Friedrick K. Beier and Gerrick Schricker (eds.), *From GATT to TRIPS – The Agreement on Trade-Related Aspects of Intellectual Property Rights*. IIC Studies, vol. 18. Max Planck Institute for Foreign and International Patent, Copyright and Competition Law, Munich, 160–215.

Taylor, M. Scott 1993: TRIPS, trade and technology transfer. *Canadian Journal of Economics*, 26, 625–38.

Taylor, M. Scott 1994: TRIPS, trade and growth. *International Economic Review*, 35, 361–82.

Van Dijk, Theon W. P. 1994: *The Limits to Patent Protection*. Maastricht: Universtaire Pers Maastricht.

Van Dijk, Theon W. P. and Patrick Van Cayseele 1993: Economic implications of converging patent breadth in Europe. In J. Hagedoorn (ed.), *Technical Change and the World Economy*, chapter 6. Edward Elgar: Cheltenham.

Watal, Jayashree 1997: Implementing the TRIPS agreement: policy options open to India. *Economic and Political Weekly*, 32, 2461–8.

Watal, Jayashree 2000: Pharmaceutical patents, prices and welfare losses: policy options for India under the WTO TRIPS agreement. *The World Economy*, 23, 733–52.

Watal, Jayashree 2001: *Background note* prepared for the WHO–WTO secretariat workshop on differential pricing and financing of essential drugs. This document can be downloaded at http://www.wto.org/english/tratop_e/trips_e/tn_hosbjor_e.htm.

Weissman, Robert 1996: A long, strange TRIPS: the pharmaceutical industry drive to harmonize global intellectual property rules, and the remaining WTO legal alternatives available to third world countries. *University of Pennsylvania Journal of International Economic Law*, 17, 1069–125.

Yang, Yong 1998: Why do southern countries have little incentive to protect northern intellectual property rights? *Canadian Journal of Economics*, 31, 800–16.

Zigic, Kresimir 1998: Intellectual property rights violations and spillovers in north–south trade. *European Economic Review*, 42, 1779–99.

15

Mixed Markets with Counterfeit Producers

E. Kwan Choi

Chapter Outline

This chapter investigates a mixed market in which one or more producers of genuine brands compete with firms that produce counterfeit goods. The counterfeit product is assumed to be a good with positive marginal utility. Consumers recognize a counterfeit but knowingly purchase it because of its low price. A Stackelberg-type leader is shown to reap a higher profit than a Nash-type monopolist. When duopolists of genuine brands compete for a larger market share, counterfeiters may actually help the brand they copy. Developing countries do not directly benefit from recognizing the intellectual property rights of developed economies and law enforcement is costly. Prosecution and elimination of counterfeiters may not be optimal. Instead, tolerance of registered counterfeiters may be socially desirable, especially when the trademark holder operates in a country where law enforcement is lax and by tradition intellectual property rights are not well defined.

1 Introduction

Four types of markets have been thoroughly analyzed in economics literature: monopoly, oligopoly, perfect competition, and monopolistic competition.[1] These types of markets have been developed in Western European and North American nations. As these countries trade with others of different cultures, their firms may find themselves operating in a different economic environment. Thus, international trade between

different cultures may force a change in the market structure. For instance, international trade may force domestic monopolies to behave as competitive firms. The interaction of firms also may engender a mixed market, an entirely new type of market where different types of firms compete against one another across national borders.

In a mixed market, different types of firms may sell identical products. There are at least two kinds of mixed markets. First, state-owned public firms may compete with private firms in the same market. Merrill and Schneider (1966) considered regulation of a government firm in an oligopolistic market. Ishibashi and Matsumura (2001) considered a mixed market in which public firms compete with private firms.

Second, when intellectual property rights are not perfectly enforced, legitimate firms may be forced to compete against producers selling counterfeits or low-quality imitations. Both the legitimate and counterfeit producers may sell close substitutes, so similar that they cannot be viewed as operating in different markets.

This chapter considers the second type of mixed markets in which producers of genuine brands compete with producers of counterfeit goods. The presence of counterfeit goods' producers and imperfect law enforcement make it possible for the two different types of firms to coexist side-by-side in the same market.

According to the US International Trade Commission (1984, p. vii), counterfeiting is "the unauthorized use of a registered trademark on a product that is identical or similar to the product for which the trademark is registered and used."

Counterfeit products are not popular in the United States due to heavy legal penalties, but are widespread in Asian countries. Increasing US trade deficits with Asian countries are likely to make the copyright and trademark infringements a contentious trade issue between the two regions.[2] However, counterfeiting is not a trade issue *per se*. Grossman and Shapiro (1988a) noted, "[i]n practice, however, most counterfeits originate in certain countries in which the laws governing the protection of trademarks are not so strict and enforcement is lax."

Producers of a genuine brand can stay in business as long as production is profitable, but any counterfeit producer vanishes when caught and prosecuted. If there are only a few counterfeiters, the monopolist may be able to expose and eliminate them by legal means. However, counterfeiters in Asian countries may be numerous, anonymous, and even invisible to the brand-name producers for all practical purposes. Prosecution of counterfeiters may be counterproductive, because law enforcement is costly and if one is caught another soon fills the void as long as expected profits are positive. Confiscation and fines may not be sufficient to dissuade counterfeiters from reentering the market. Periodic crackdowns have not deterred counterfeiters in Asian countries.

In a seminal paper, Grossman and Shapiro (1988a) first considered trade in counterfeit products. In their model, foreign firms export counterfeits of domestic brand-name products, and the counterfeit products are of inferior quality. Examples of common counterfeit goods are watches and drugs.

In other industries, counterfeiters are producing imitations that are identical to the original products because they are manufactured in the same production plant. For instance, a brand-name producer of textiles may design a product and outsource its production to a firm in China. The actual manufacturer then supplies the product to the trademark holder in the United States, but also may market extra or unauthorized output in China or other Asian countries.[3] Since these counterfeits are made by the

same manufacturer, they may be functionally and qualitatively identical to the originals. Similarly, digitized copies often are functionally identical to the originals. For instance, music CDs and movie DVDs produced in the United States can be copied and sold easily in other countries. Likewise, copies of software CDs are identical to the originals.

The plan of this chapter is as follows. Section 2 derives consumer demand for counterfeits. Section 3 considers a short-run Nash equilibrium in a mixed market with a monopolist and many counterfeiters, while Section 4 shows the advantages of a Stackelberg-type monopolist. Section 5 introduces free entry and examines long-run Nash and Stackelberg equilibria. Section 6 investigates a mixed market with duopolists of genuine brands and counterfeiters who imitate only one brand. Section 7 considers a monopolist subsidizing law enforcement in a foreign country. Section 8 analyzes the benefits of an alternative policy of tolerance rather than prosecution, while Section 9 shows that tolerance of counterfeits may be socially optimal. Section 10 contains concluding remarks.

2 Demand for Counterfeit Products

Often counterfeits are not as good as the originals. Sometimes, counterfeits are useless.[4] Nonfunctional counterfeit cameras have been produced, and bogus pills are marketed at low prices. In other cases, counterfeits are somewhat inferior to the genuine brands they imitate. For instance, counterfeit watches resemble the genuine brands but are lacking in quality or function. In still other cases such as music CDs or DVDs, information is digitized and the copies are functionally identical to the originals. Similarly, copies of software are functionally identical to the original. Crampes and Jacques-Laffont (2002) analyze software copying, which appears to be widespread even in the United States. Often an infringer buys one legal copy and makes illegal copies for use within the firm.

Grossman and Shapiro (1988a, p. 60) suggest that counterfeiting is intrinsically harmful to society:

> Counterfeiting undermines the functioning of the property rights system. Not only do consumers suffer the direct harm associated with the purchase of low-quality copies purporting to be the originals, but the infringement on the legitimate firms' rights alters the incentives to invest in their reputation.

Higgins and Rubin (1986) analyzed production decisions of a monopolist when consumers are aware that they are purchasing counterfeit goods.[5] Counterfeits are generally sold in the black market, whereas the genuine brands are sold in authorized stores. Consequently, consumers usually can discern counterfeits from the genuine brands. Few consumers unknowingly buy counterfeits at the price of the genuine brands. Moreover, there is a large difference in prices. Counterfeit watches often are sold at a very small fraction of the price of the genuine products. The large price difference indicates that these counterfeits are not purported to be the originals and are sold as counterfeits. In this situation, there may be little intent on the part of the counterfeiters to mislead consumers.

In this chapter we only consider counterfeit *goods*, that is, counterfeits that have positive marginal utilities without dangerous side effects. Consumers knowingly purchase counterfeits because the price is lower and counterfeits fulfill some functions of the originals. For the purpose of illustration, we consider a two-country model for DVD products such as movies and computer games; however, the theory also applies to markets where counterfeits are inferior imitations.

How money is spent on other consumption goods or the numéraire in the presence of counterfeits is not an issue in this chapter. Accordingly, we simplify the derivation of demand for counterfeits, and consider a representative consumer spending fixed income I on two products, a genuine product x and its counterfeit X.[6] The counterfeit product is not a bad, but a good, and marginal utility of the counterfeit good is positive. The quality of the genuine product is unity. The counterfeit serves some functions of the genuine brand. That is, the counterfeit is inferior to the genuine brand product. Thus, one unit of the counterfeit is assumed to be qualitatively equivalent to h units of the genuine product and $h \leqslant 1$. Counterfeit movies or software may be identical to the original, in which case $h = 1$. The consumer does not observe the quality of the counterfeit beforehand and hence h is a random variable, $0 < h \leqslant 1$. Let $z = x + hX$ be the weighted total quantity.

The consumer's utility is represented by a von Neumann–Morgenstern utility function $u(x+hX)$. The consumer's problem is to maximize expected utility $Eu(x+hX)$ subject to a budget constraint, $px+PX = I$, where p and P are the prices of the genuine product and the counterfeit, respectively. Expected utility is written as:

$$Eu\left(hX + \frac{I - PX}{p}\right). \tag{1}$$

Differentiating (1) with respect to X gives

$$Eu'(h - P/p) \equiv Eu'g(Eh - P/p) + \mathrm{Cov}(u', h - P/p) = 0, \tag{2}$$

for $X > 0$, where $\mathrm{Cov}(g)$ is the covariance operator. Note that $\mathrm{d}u'/\mathrm{d}h = u''X < 0$, and hence $\mathrm{Cov}(u', h - P/p) < 0$. If $P \geqslant pEh$, the left side of the equality in (2) is negative, and the consumer will not purchase any counterfeits. Thus, the first-order condition in (2) implies that

$$pEh > P \tag{3}$$

must hold for the consumer to purchase some amount of the counterfeit. That is, the price of the counterfeit must be less than the expected quality times the price of the genuine product. The question is whether the consumer *will* purchase some counterfeit product when $P < pEh$.

We now evaluate the left side of the equality in (2) at $X = 0$.

$$Eu'[I/p](h - P/p) = u'[I/p](Eh - P/p) > 0. \tag{4}$$

That is, if the relative price of the counterfeit good is less than its expected quality – the quality of the genuine brand is unity – then the expected marginal utility of the counterfeit is positive when evaluated at $X = 0$. Thus, a risk-averse consumer will purchase some amount of the counterfeit product.

Proposition 1. If the relative price of the counterfeit good is less than its expected quality ($P/p < Eh$), then a risk-averse consumer purchases some amount of the counterfeit product. If $P/p \geqslant Eh$, a risk-averse consumer does not purchase the counterfeit product.

2.1 Properties of Counterfeit Demand Function

We now examine the properties of the demand function for the counterfeit product. Differentiating (2) with respect to I yields

$$\frac{\partial X}{\partial I} = -\frac{Eu_{XI}}{Eu_{XX}},$$

where $Eu_{XX} = Eu''(h - P/p)^2 < 0$ for a risk-averse consumer, and

$$Eu_{XI} = \frac{1}{q} Eu''(h - P/p).$$

The counterfeit product is a normal good ($\partial X/\partial I > 0$) if, and only if, $Eu''(h - P/p) > 0$. We assume that the counterfeit product is a normal good. It is shown in the Appendix that diminishing absolute risk-aversion (DARA) is sufficient for the counterfeit to be a normal good.

Next, differentiating (2) with respect to P and p gives

$$\frac{\partial X}{\partial P} = -\frac{Eu_{XP}}{Eu_{XX}} \quad \text{and} \quad \frac{\partial X}{\partial p} = -\frac{Eu_{Xp}}{Eu_{XX}},$$

where

$$Eu_{XP} = -\frac{X}{p} Eu''\,(h - P/p) - \frac{1}{q} Eu' < 0, \tag{5}$$

$$Eu_{Xp} = -(x/p) Eu''(h - P/p) + (P/p^2) Eu'. \tag{6}$$

Since the two terms on the right side of (6) have opposite signs, the sign of $\partial X/\partial p$ is generally ambiguous. These results are summarized below:

Proposition 2. Assume that the consumer exhibits DARA. Then the counterfeit product is a normal good and its demand curve is negatively sloped, that is, $(\partial X/\partial I) > 0$ and $(\partial X/\partial P) < 0$.

Properties of demand function for the genuine product can be obtained from the budget constraint. Differentiating $x = (I - PX)/p$ with respect to P and p gives

$$\frac{\partial x}{\partial P} = -\frac{P}{p} \frac{\partial X}{\partial P} > 0, \tag{7}$$

$$\frac{\partial x}{\partial p} = -\frac{x}{p} - \frac{P}{p} \frac{\partial X}{\partial p} = -\frac{x}{p} + \frac{\partial x}{\partial P}. \tag{8}$$

Thus, given DARA, the demand for the genuine product also is *positively* related to the price of the counterfeit.

3 Monopoly and Counterfeit Producers

We consider a mixed market with counterfeit producers in a two-country framework. There are two countries, the home country (America) and the foreign country (Asia). There is a single firm with the genuine brand, which is sold to the world market, but many foreign firms produce counterfeits. The monopolist does not license its technology to foreign firms.[7] Due to strict law enforcement, counterfeiters do not exist in the home market. They exist and operate only in the foreign market where law enforcement is lax. Their outputs are sold in the foreign market and no firms are allowed to import counterfeits into the home country. However, consumers visit the foreign country as tourists and may bring counterfeits to the home market.[8]

The foreign government does not rigorously enforce copyrights and trademarks within its territory. Law enforcement is costly, and the foreign government has no strong incentive to crack down on the counterfeiters contributing to the national economy. Accordingly, only some, but not all, infringers are caught and prosecuted. Let e denote the number of counterfeiters caught. In general, the number of arrests is a function of how many law enforcement agents are hired in the foreign country. The quality of counterfeit products may vary in practice, depending on the technology used by counterfeiters. For simplicity, counterfeiters are assumed to be identical firms.

Once an infringer is caught, the counterfeit goods are confiscated and destroyed, and they do not reach the market. Let n denote the number of counterfeiters, and let X now denote the output of a representative counterfeiter. Recall that due to inadequate law enforcement, not all counterfeiters are caught ($e < n$). The number of surviving counterfeiters is $(n - e)$. The total output of the surviving counterfeiters is

$$X^* = (n - e)X. \tag{9}$$

In the previous section, some properties of demand function for the counterfeit good were derived. We now focus on the *market* demands for the genuine brand and counterfeit product. For simplicity, market demands for the counterfeit and the genuine products are assumed to be linear in prices. Specifically, the inverse demand for the genuine product is given by

$$p = a - bx - (n - e)gX, \tag{10}$$

where x is now the output of the monopolist. All parameters, a, b, n, e, and g are positive. The presence of the counterfeit product decreases the demand for the genuine product since g is positive. If the genuine product is not well known, the presence of imitations may actually increase the demand for the genuine product.

The inverse demand function for the counterfeit is given by

$$P = A - (n - e)BX - Gx, \tag{11}$$

where P is the price of the counterfeit and A is a positive constant.

Since the genuine product and the counterfeit are close substitutes, increased production of one good depresses the demand for the other. Accordingly, the price of a product is inversely related, not only to its own quantity, but also to the quantity of the other product. Since they are not perfect substitutes, the slopes of cross-demand curves are less than those of own-demand curves, that is, $g < b$ and $G < B$. Solving the system of two equations in (10) and (11), we obtain:

$$x = \frac{B(a - p) - g(A - P)}{\Delta},$$

$$X = \frac{b(A - P) - G(a - p)}{(n - e)\Delta},$$

where $\Delta \equiv bB - gG > 0$ by assumption. Note that $g > 0$ and $G > 0$ imply that cross-demand curves are positively sloped, that is, an increase in the price of one good increases the demand for the other product.

3.1 Monopoly

We first consider the production decision of the monopolist in the presence of counterfeiters. Monopoly profit is

$$\pi = (a - bx - (n - e)gX)x - c(x). \tag{12}$$

The first-order condition is

$$\pi_x \equiv a - 2bx - (n - e)gX - c' = 0, \tag{13}$$

where $m_x = (a - 2bx - (n - e)gX)$ is marginal revenue. The monopolist's marginal revenue, as well as the demand curve, are adversely affected by the counterfeit sales. Equation (13) defines the monopolist's output as a function of counterfeit production, that is, it defines a reaction function $x = \eta(X)$ of the monopolist, along which $\pi_x = 0$. For the purpose of graphical illustration, however, its inverse function, $X = r(x) \equiv \eta^{-1}(x)$ is used hereafter. The slope of the monopolist's reaction curve is

$$r'(x) = -\frac{\pi_{xx}}{\pi_{xX}} = -\frac{2b + c''}{g} < 0. \tag{14}$$

Thus, the monopolist's reaction curve is negatively sloped as shown in Figure 15.1.

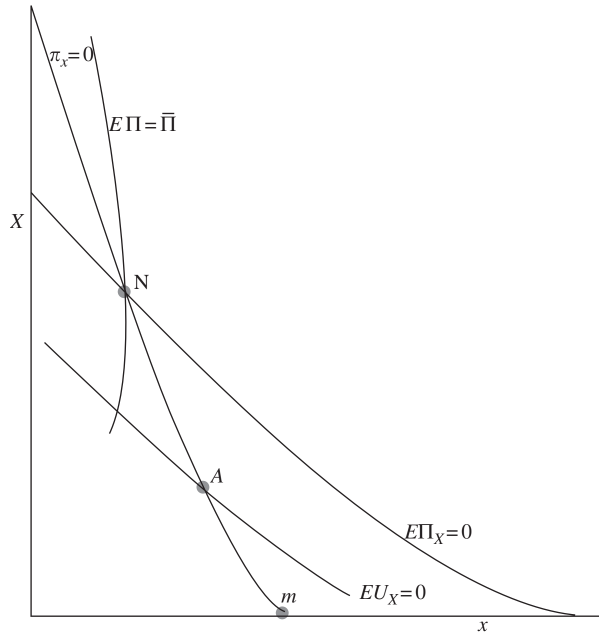

Figure 15.1 Risk aversion and Nash equilibrium

3.2 Producers of Counterfeit Goods

Entry into the mixed market is free except for the risk of confiscation of illegally produced output and the fine a counterfeiter pays when caught. All counterfeit producers face the same risk of confiscation. These counterfeiters do not collude and are assumed not to exhibit any strategic behavior. Each counterfeiter is a price-taking firm, and assumes that the prices of the counterfeit and the genuine product are given. Accordingly, we consider a representative firm selling the counterfeit product.

If a counterfeiter is not caught by the authorities, his profit is

$$\Pi^o = PX - C(X),$$

where $C(X)$ is the production cost, which is incurred before the counterfeit is sold in the black market.

If the counterfeiter is caught, the total output is confiscated and the counterfeiter receives no revenue. Moreover, the counterfeiter is fined. His profit is then

$$\Pi^T = -C(X) - T,$$

where T is the fine the counterfeiter must pay. Imprisonment is either ruled out, or the opportunity cost of imprisonment is included in the penalty T.[9]

In practice, a counterfeiter may be caught at any stage of production and sales. One may be caught after selling some or all outputs. However, for simplicity, we assume

that once detected, both the revenue from the outputs sold and the remaining outputs are confiscated. This amounts to confiscation of the total output before selling any products, as was assumed earlier.

All counterfeiters face the same risk of detection and confiscation. However, for the entire industry producing counterfeits, the number of arrests e is deterministic. The probability that any individual counterfeiter will be caught is $\sigma = e/n$.

3.2.1 RISK AVERSION

Consider the production decision of a risk-averse counterfeiter. The expected utility of a representative counterfeiter is

$$EU = (1-\sigma)U(PX - C(X)) + \sigma U(-C(X) - T), \tag{15}$$

where $U(g)$ is a von Neumann–Morgenstern utility function. Let $V = U(-C(X) - T)$ denote the utility when the counterfeiter is arrested and his output is confiscated in order to distinguish it from the utility $U = U(PX - C(X))$ when the counterfeiter is not caught. The first-order condition for optimal production is:

$$EU_X \equiv (1-\sigma)U'(P - C') + \sigma V'(-C') = 0,$$

where $C' = C'(X)$ is marginal production cost. Thus,

$$(1-\sigma)(P - C') = \sigma C' \frac{U'}{V'}. \tag{16}$$

Note that $V' = U'(-C(X) - T)$ is the marginal utility of income when the counterfeiter is fined and his product is confiscated, and hence $V' > U'$. Thus, $(1-\sigma)(P - C') > \sigma C'$, or

$$E\Pi_X \equiv (1-\sigma)P - C'(X) > 0. \tag{17}$$

If the counterfeiter is risk neutral, the first-order condition reduces to $E\Pi_X = 0$. Let X^o and X^a denote the outputs of risk-neutral and risk-averse counterfeit producers, respectively. Then the above results can be summarized as follows.

Proposition 3. Assume that law enforcement is imperfect and the probability of arrest of a trademark infringer is positive but less than unity ($0 < \sigma < 1$). Then a risk-averse counterfeiter produces less than a risk-neutral producer, that is, $X^a < X^o$.

Proposition 3 shows that regardless of the monopolist's output, risk aversion reduces the output of the counterfeiter. Figure 15.1 illustrates a Nash equilibrium with risk-averse counterfeiters. The locus of points (x, X) that satisfy $\pi_x(x, X) = 0$ defines the monopolist's reaction curve. On the other hand, the locus of points that satisfies $EU_X(x, X) = 0$ defines a risk-averse counterfeiter's reaction curve.[10] Along this curve, $E\Pi_X(x, X) > 0$. The intersection A of these reaction curves determines a Nash equilibrium with risk-averse counterfeiters.

3.2.2 Risk neutrality

Proposition 1 also suggests how a Nash equilibrium may be affected by the risk attitude of counterfeiters. Consider the production decision of a risk-neutral counterfeiter. The expected profit of the counterfeiter is

$$E\Pi = (1 - \sigma)PX - C(X) - \sigma T. \tag{18}$$

Note that each counterfeiter takes market price $P = A - (n - e)BX - Gx$ as given. The first-order condition for optimal production is

$$\frac{\partial E\Pi}{\partial X} \equiv P(1 - \sigma) - C' = 0. \tag{19}$$

Recall that σ is the probability that the counterfeiter is caught and receives no revenue. Intuitively, (19) states that the counterfeiter takes the confiscation risk into account and discounts market price, equating expected price $P(1 - \sigma)$to marginal cost. The second-order condition requires that $E\Pi_{XX} = -C''(X) < 0$.

Equation (19) describes implicitly the reaction function $X = R(x)$ of the risk-neutral counterfeiter, that is, it shows the relationship between x and X along which $E\Pi_X = 0$. Differentiating (19) with respect to x gives,

$$R'(x) = -\frac{E\Pi_{Xx}}{E\Pi_{XX}} = -\frac{G(1 - \sigma)}{C''} < 0, \tag{20}$$

which implies that the counterfeiter's reaction curve is negatively sloped if $G > 0$. In deriving the second-order condition, $E\Pi_{XX} < 0$, market price P is assumed to be given. However, a change in the monopolist's output affects the price of the counterfeit when deriving Π_{Xx}.

At this juncture, it is helpful to compare the slopes of the two reaction curves. From (14) and (20), we obtain

$$r'(x) - R'(x) = -\frac{2b + c''}{g} + \frac{G(1 - \sigma)}{C''}.$$

It follows that the reaction curve of the monopolist may be steeper or flatter than that of the counterfeiter. For instance, consider the case where the genuine product and the counterfeit are functionally identical, as in the case of DVDs. If marginal cost of duplication is constant, ($c''(x) = C''(X)$ is zero), the reaction curve of the counterfeiter is vertical. On the other hand, if $c'' \geqslant g$ and the two demand curves are symmetric ($c'' = C''$, and $g = G$), then the reaction curve of the counterfeiter is steeper than that of the monopolist.

Figure 15.1 also illustrates a Nash equilibrium with risk-neutral counterfeiters. Along the reaction curve $R(x)$ of the representative counterfeiter, expected marginal profit is zero, that is, $E\Pi_X(x, X) = 0$. Thus, a Nash equilibrium with risk-neutral counterfeiters is established at point N. For a risk-averse firm, (17) indicates that $E\Pi_X(x, X) > 0$. Thus, a risk-averse counterfeiter's reaction curve lies below that of the risk-neutral firm.

3.3 Law Enforcement and Short-Run Nash Equilibrium of the Mixed Market

In the absence of counterfeiters, the monopolist would enjoy his monopoly position at point m in Figure 15.1. Once risk-neutral counterfeiters enter the market, a Nash equilibrium obtains at point N. In the short run, counterfeiters may earn positive or negative profits at the Nash equilibrium.

Recall that expected price is equal to marginal cost along the risk-neutral counterfeiter's reaction curve. If a counterfeiter is risk averse, expected price is greater than marginal cost. Thus, the reaction curve of a risk-neutral counterfeiter lies below that of the risk-neutral one, thereby reducing the output of the counterfeiter and raising that of the monopolist.

It is important to note in (19) that the presence of law enforcement agents directly affects the output of a counterfeiter by increasing the probability of arrest. On the other hand, a fine has no effect on the output of the counterfeiter.[11] However, it affects the level of profit in the short run and hence entry in the long run.

Next, consider the effect of law enforcement on Nash equilibrium. From (13), we note that a fine on the counterfeiters has no direct effect on the monopolist, but the presence of the police force reduces the number of counterfeiters and hence increases the monopolist's demand and marginal revenue. Specifically,

$$\frac{d\pi_x}{dT} = 0, \quad \frac{d\pi_x}{de} = gX > 0.$$

Thus, an increase in e shifts the monopolist's reaction curve to the right.

Differentiating the left side of (19) holding P constant gives

$$E\Pi_{Xe} = -\frac{P}{n} < 0.$$

Thus, an increase in e shifts the counterfeiter's reaction curve to the left.

Figure 15.2 illustrates the effects of law enforcement when the monopolist's reaction curve is steeper than that of the counterfeiter. Assume that in the beginning a Nash equilibrium occurs at point N. Raising the fine has no effect on the behavior of the counterfeiter. However, an increase in the police force shifts the reaction curve of the monopolist to the right, while shifting that of the counterfeiter to the left. Thus, the Nash equilibrium moves to point N′. In this case, law enforcement unambiguously curtails counterfeit production while increasing the monopolist's output.

Figure 15.3 illustrates the adverse effect of law enforcement on the output of the counterfeiters. The monopolist's reaction curve is flatter than that of the counterfeiter. An increase in police force shifts the monopolist's reaction curve to r' and the counterfeiter's reaction curve to R'.

Accordingly, a new Nash equilibrium prevails at N'. Consequently, increasing the police force paradoxically raises the counterfeiter's output.

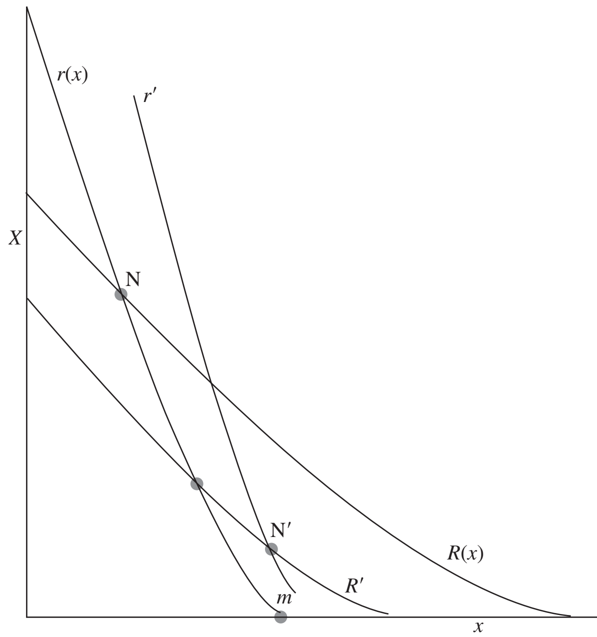

Figure 15.2 Law enforcement and Nash equilibrium

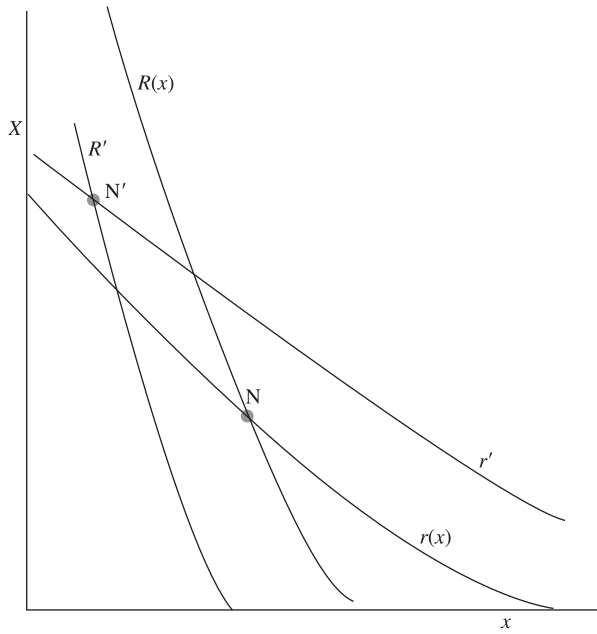

Figure 15.3 Adverse effect of law enforcement

Proposition 4. An increase in police force decreases (increases) the production of the counterfeiter and increases (decreases) the output of the monopolist if

$$-\frac{2b + c''}{g} + \frac{G(1-\sigma)}{C''}$$

is positive (negative).

Proposition 4 shows that the presence of the police force generally has an ambiguous effect on counterfeit production.

How does law enforcement affect the number of counterfeiters? Differentiating (18) with respect to T gives

$$\frac{\partial E\Pi}{\partial T} = -\sigma < 0. \tag{21}$$

Thus, a fine is a deterrent to counterfeiters and reduces the number of counterfeiters. Differentiating (18) with respect to e yields

$$\frac{\partial E\Pi}{\partial e} = -\frac{PX + T}{n} < 0. \tag{22}$$

Thus, the police force also is a deterrent to counterfeiting.

Proposition 5. Assume that the probability of arrest is positive. Then both the fine and police force are deterrents to the entry of counterfeiters into the mixed market.

4 Stackelberg Leader Equilibrium

In countries where law enforcement is lax, the monopolists may not be able to persuade the government to exterminate or drive out the counterfeiters. However, they can utilize the predictable responses of counterfeiters and learn to live with them. The monopolist recognizes that counterfeiters exhibit competitive behavior, that is, they take the prices of the genuine product and the counterfeit as given and adjust their outputs. In this situation, the monopolist can take into account such predictable competitive behavior and choose its output along the representative counterfeiter's reaction curve. In other words, the monopolist can act like a Stackelberg leader, except for the fact that he is operating in a mixed market with many competitive counterfeiters.

A Stackelberg leader's problem is to maximize (12) subject to the counterfeiter's reaction function. Substituting $X = R(x)$ into (12), we have

$$\pi = (a - bx - (n-e)gR(x))x - c(x). \tag{23}$$

Recall that in a Nash equilibrium, the first-order condition is

$$a - 2bx - g(n-e)X - c_x = 0. \tag{13}$$

For the Stackelberg leader, the first-order condition is

$$a - 2bx - g(n - e)(R'x + X) - c_x = 0. \tag{24}$$

Comparing these two conditions, we see that the left side of (24) has an extra term,

$$-g(n - e)R'x > 0,$$

which does not appear in (13). That is, when evaluated at a Nash equilibrium, the left side of (24) is positive, which means that the Stackelberg-type monopolist's profit is still rising as it increases its output along the counterfeiter's reaction curve. Note that in this problem the Stackelberg leader moves along the counterfeiter's reaction curve, whether the latter is risk neutral or averse.[12] The Stackelberg leader's output is to the right of the Nash equilibrium, regardless of the risk attitude of the counterfeiter. Thus, the Stackelberg leader's output is greater than a Nash-type monopolist, as shown in Figure 15.4.[13]

Proposition 6. In the presence of competitive counterfeiters, a Stackelberg-type monopolist produces more and earns a higher profit than a Nash-type monopolist.

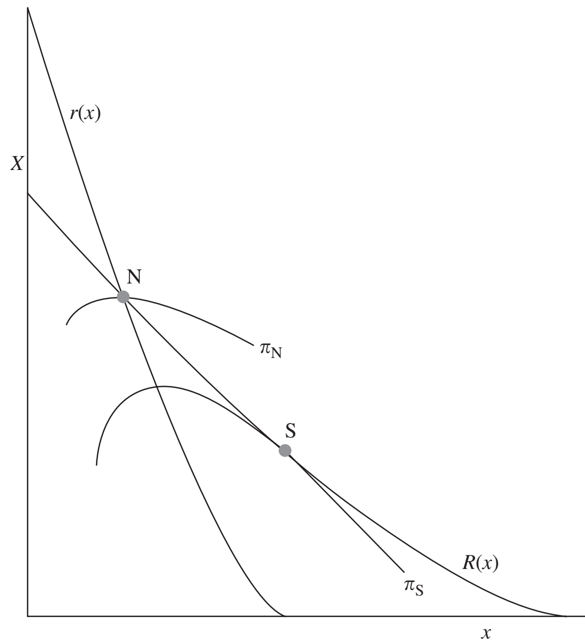

Figure 15.4 Stackelberg equilibrium

5 Long-Run Equilibria

Law enforcement is a barrier to entry in a mixed market with counterfeit producers. However, entry is free to all potential counterfeit producers who are willing to assume the risk of confiscation or imprisonment. Thus, counterfeit producers resemble competitive firms, except for the risk of getting caught. They are illegal competitive firms, as opposed to legal firms operating in a perfectly competitive market.

While a Stackelberg leader enjoys a higher profit than a Nash-type monopolist, the Stackelberg equilibrium may not be sustainable in the long run if the counterfeiters are earning positive expected profits. Similarly, a short-run Nash equilibrium may not be sustainable if counterfeit producers are earning positive expected profits. The entry of counterfeiters will continue until expected profits are zero.

Recall that the expected profit of the counterfeiter is

$$(A - (n - e)BX - Gx)X(1 - \sigma) - C(X) - \sigma T.$$

How does entry affect the reaction curve of the representative counterfeiter? In order to examine the effect of entry, the risk-neutral counterfeiter's reaction function in (19) is rewritten as

$$(A - (n - e)BX - Gx)(1 - e/n) - C_X = 0. \tag{25}$$

This condition shows that marginal cost is equal to expected price, which is market price times the probability of survival. Entry affects both price and the probability of survival. Thus, entry of new counterfeiters affects the incumbent's reaction curve through the probability of arrest and market price P. Specifically, as entry occurs, the probability of arrest decreases but so does the price of the counterfeit goods. Differentiating (25) with respect to n, we get

$$\frac{\partial E\Pi_X}{\partial n} = \frac{\partial P}{\partial n}\left(1 - \frac{e}{n}\right) + \frac{Pe}{n^2}, \tag{26}$$

where

$$\frac{\partial P}{\partial n} = -BX < 0. \tag{27}$$

Thus, the effect of entry on the counterfeiter's reaction is generally ambiguous.

Note that the probability of arrest, e/n, remains constant when the numbers of counterfeiters and arrests increase proportionately. When $n = e$, all counterfeiters are caught. To simplify, let us normalize the number of arrests so that $e = 1$. When there is one counterfeiter, he is caught, and the (expected) price is

$$P(1) = A - Gx.$$

With two counterfeiters, the expected price is

$$P(2)\tfrac{1}{2} = (A - BX - Gx)\tfrac{1}{2} = \tfrac{1}{2}(P(1) - BX) < P(1).$$

Thus, the increment in the expected price when the number of counterfeiters increases from one to two is:

$$\Delta_1 \equiv P(2)\tfrac{1}{2} - P(1) = -\tfrac{1}{2}(P(1) + BX) < 0.$$

With n counterfeiters, the expected price is:

$$P(n)\frac{n-1}{n} = (A - (n-1)BX - Gx)\frac{n-1}{n}. \tag{28}$$

When there are $(n+1)$ counterfeiters, the expected price is:

$$P(n+1)\frac{n}{n+1} = (A - nBX - Gx)\frac{n}{n+1} = (P(n) - BX)\frac{n}{n+1}. \tag{29}$$

Subtracting (28) from (29), we get

$$\Delta_n \equiv P(n+1)\frac{n}{n+1} - P(n)\frac{n-1}{n} = P(n)\left(\frac{n}{n+1} - \frac{n-1}{n}\right) - Bx\frac{n}{n+1}.$$

After rearranging, the increment in expected price when n increases by one is

$$\Delta_n = P(n)\frac{1}{n(n+1)} - BX\frac{n}{n+1}. \tag{30}$$

The first term on the right side of (30) is positive but the second term is negative. For small values of n, we observe

$$\Delta_2 = P(2)\tfrac{1}{6} - BX\tfrac{2}{3},$$

$$\Delta_3 = P(3)\tfrac{1}{12} - BX\tfrac{3}{4},$$

$$\Delta_4 = P(4)\tfrac{1}{20} - BX\tfrac{4}{5},$$

and so on. It is clear that for some small values of n, it is possible for Δ_n to be positive. However, as the number of counterfeiters increases, the first term approaches zero while the second term approaches BX. Thus, the second term in (30) begins to dominate the first beyond some n^o. When there are a few counterfeiters ($n < n^o$), the first term on the right side of (26) may more than offset the second term. In this case, in (26) $E\Pi_{Xn}$ is positive and the counterfeiter's reaction curve shifts upward. However, as the number of counterfeiters increases further, $E\Pi_{Xn}$ becomes negative, and eventually entry of new firms will move the reaction curve downward.

Figure 15.5 illustrates the case where the counterfeiter's reaction curve shifts downward. The initial Nash equilibrium is the intersection of reaction curves $r(x)$ and $R(x, n)$. As more firms enter the market, the reaction curve shifts to $R(x, n')$. Accordingly, the Nash and the Stackelberg equilibria move to N′ and S′, respectively.

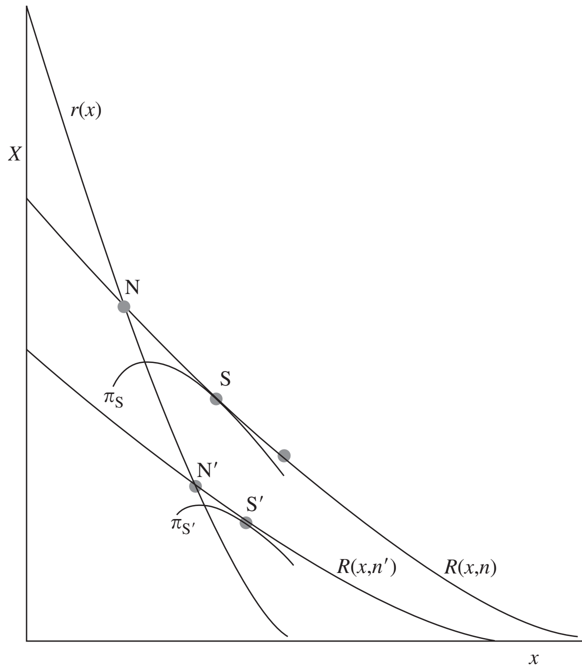

Figure 15.5 Entry of new counterfeiters

It is important to note that these equilibrium points refer to per capita output of counterfeiters. Let $X^* = nX$ denote the total output of counterfeiters. Then

$$\frac{\mathrm{d}X^*}{\mathrm{d}n} = X + n\frac{\mathrm{d}X}{\mathrm{d}n}.$$

Thus, even if the representative counterfeiter's output declines as n increases, the aggregate output of counterfeiters may still increase.

5.1 Long-Run Nash Equilibria: Coexistence or Extinction of the Monopolist?

There are two possible long-run Nash equilibria in the mixed market. First, the counterfeiters and the monopolist may coexist in a long-run Nash equilibrium. Second, the monopolist may be forced to exit the market due to the continued entry of counterfeiters.

We now consider the first type of long-run Nash equilibrium, shown in Figure 15.6. The entry of new counterfeiters continues until the representative counterfeiter earns zero expected profit. Thus, in the long-run equilibrium of a mixed market, total profits of all counterfeits are zero,

$$nE\Pi = (A - (n - e)BX - Gx)(n - e)X - nC(X) - eT = 0. \tag{31}$$

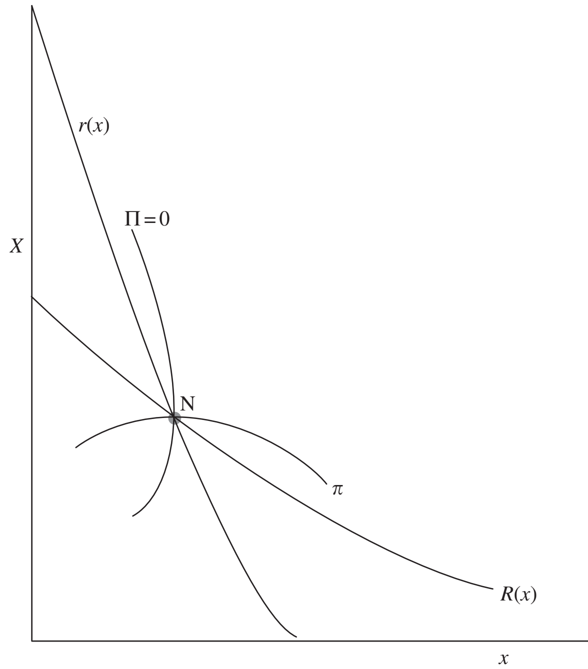

Figure 15.6 Long-run Nash equilibrium

Dividing (31) by X^* yields

$$P(1-\sigma) = \theta(X) + \sigma\frac{T}{X}, \tag{32}$$

where $\theta(X) = C(X)/X$ is unit cost.

In Figure 15.6, the Nash equilibrium is the intersection of two reaction curves. An iso-profit curve is a locus of (x, X) along which expected profit of a representative counterfeiter is constant, that is, $E\Pi(x, X) = \overline{\Pi}$. To be in a long-run Nash equilibrium, expected profit of the counterfeiter is zero ($E\Pi = 0$). In contrast, at the intersection of the two reaction curves, the monopolist must make a non-negative profit.

A second type of Nash equilibrium occurs when the monopolist makes a negative profit at the intersection of the two reaction curves. In this case, the monopolist vanishes from the market and the Nash equilibrium degenerates into the long-run equilibrium of a competitive market in which only the counterfeiters survive. A further output adjustment is made to fill the void left by the vanishing monopolist. However, without the genuine product supplied by an active monopolist, the demand for counterfeits may gradually decline.

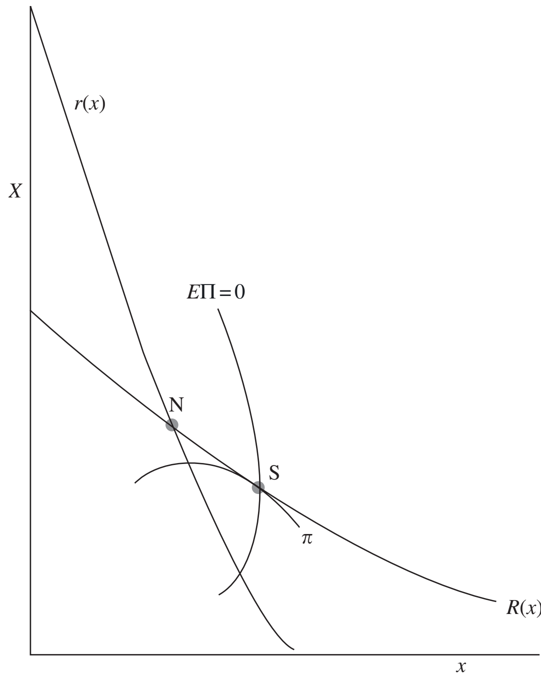

Figure 15.7 Long-run survival of a Stackelberg monopolist

5.2 Long-Run Stackelberg Equilibrium

When a Nash equilibrium does not guarantee a positive profit, the monopolist has a better chance to survive as a Stackelberg leader. Even if the monopolist can survive at the Nash equilibrium, a still higher level of profit can be guaranteed by a Stackelberg-type monopolist. Stackelberg leadership might be a superior market strategy when the monopolist cannot ensure a positive profit in long-run Nash equilibrium.

In Figure 15.7, a Nash monopolist may suffer a loss and must soon exit the market. Alternatively, the monopolist can choose a point S along the counterfeiter's reaction curve. If $\pi > 0$, then point S is a possible long-run equilibrium for the Stackelberg-type monopolist. If the profit is negative even at point S, the monopolist will exit the market, in the absence of more stringent law enforcement.

6 Duopoly and Counterfeiters

In this section we consider a mixed market with duopolists producing genuine brands. In the real world, counterfeiters are likely to make copies of both the genuine brands. However, we consider a simple case where the counterfeiting firms produce copies of only one brand. For instance, one duopolist may rigorously prosecute counterfeiters,

little realizing that they might help him grab a larger market share. We then investigate whether counterfeiters hurt or benefit the brand that they copy.

Let x be the output of one duopolist and X the output of the imitating firm that makes counterfeits of x. Let y be the output of the other duopolist, whose brand is not being copied. Inverse demand functions of duopolists are given by

$$p = (a - bx - fy - (n - e)gX), \tag{33}$$

$$q = \alpha - \beta y - \gamma x, \tag{34}$$

where p and q are the prices of duopolists one and two, and α, β, and γ are positive constants which are counterparts of a, b, and f.

Profits of duopolists are given by

$$\pi(x, y, X) = (a - bx - fy - (n - e)gX)x - c(x), \tag{35}$$

$$\psi(y, x) = (\alpha - \beta y - \gamma x)y - c(y). \tag{36}$$

The expected profit of the counterfeiter is

$$E\Pi(X, x, y) = P(1 - \sigma)X - C(X) - \sigma T, \tag{37}$$

where

$$P = A - (n - e)BX - Fy - Gx. \tag{38}$$

Note that in (37) the output of the second duopolist does not directly affect the counterfeiter's profit. Also, the counterfeiter does not produce any counterfeit products to compete with the second duopolist. However, the inverse demand function in (38) indicates that as long as $F \neq 0$, the counterfeiter indirectly competes with the second duopolist because the latter's output may affect the demand for the counterfeit that competes with the first duopolist's genuine brand.

The first-order condition for optimal counterfeit production is

$$E\Pi_X \equiv P(1 - \sigma) - C_X = 0. \tag{39}$$

Differentiating (39) with respect to y gives

$$\frac{\partial E\Pi_X}{\partial y} = (1 - \sigma)\frac{\partial P}{\partial y} = -(1 - \sigma)F < 0. \tag{40}$$

That is, the presence of the second duopolist reduces the output of the counterfeiter if $F > 0$. Thus, indirectly the counterfeiter is competing with the second duopolist even though he does not make counterfeits of the latter's output.

The first-order conditions for duopolists are

$$\pi_x \equiv a - 2bx - fy - (n - e)gX - c_x = 0, \tag{41}$$

$$\psi_y \equiv \alpha - 2\beta y - \gamma x - c_y = 0. \tag{42}$$

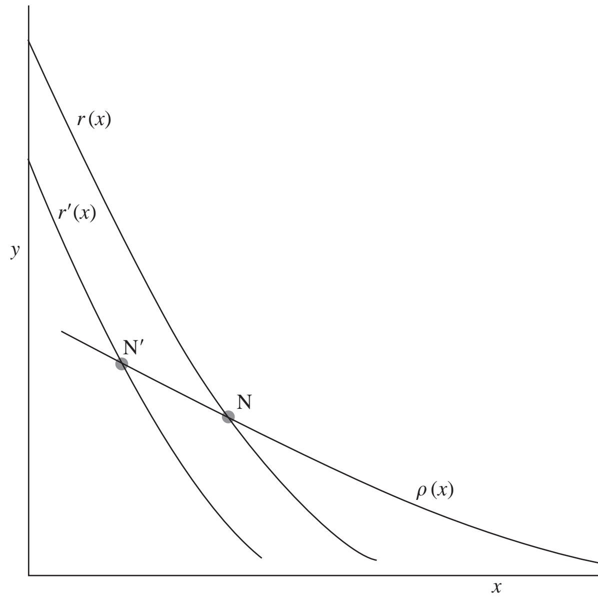

Figure 15.8 Duopoly equilibrium with counterfeiters

Since there are three producers, a Nash equilibrium cannot be drawn satisfactorily in a two-dimensional diagram. Figure 15.8 shows only the reaction curves of duopolists in the (x, y) space, while the Nash equilibrium output of the counterfeiter is not drawn.

The slope of the first duopolist's reaction curve is obtained by differentiating (41) with respect to x:

$$r'(x) = \frac{dy}{dx} = -\frac{\pi_{xx}}{\pi_{xy}} = -\frac{2b}{f} < 0. \tag{43}$$

Similarly, the slope of the second duopolist's reaction curve is

$$\rho'(x) = -\frac{\psi_{yx}}{\psi_{yy}} = -\frac{\gamma}{2\beta} < 0. \tag{44}$$

Then

$$r'(x) - \rho'(x) = -\frac{2b}{f} + \frac{\gamma}{2\beta}.$$

If the duopolists have symmetric demand curves ($b = \beta$ and $f = \gamma$), then

$$r'(x) - \rho'(x) = \frac{f^2 - b^2}{bf} < 0,$$

because the two genuine brands are close, but not perfect, substitutes ($b > f$). Thus, the reaction curve of the first duopolist is steeper than that of the second, as shown in Figure 15.8.

6.1 Counterfeiters: Friends or Foes?

How does the presence of counterfeiters affect the Nash equilibrium of genuine brand producers? It is important to distinguish the role of counterfeiters that copy the monopolist's product from that of counterfeiters who copy one brand but not the other. When there is a single producer of the genuine product, the counterfeiters undoubtedly reduce the monopolist's demand for the genuine product.[14] However, when there are two or more genuine brand producers, they compete for brand-name recognition. Thus, the presence of counterfeiters may enhance brand recognition and increase the demand for the genuine brand they imitate.[15] In other words, counterfeiters may actually help the brand they copy to dominate the market for the genuine brands. In the last analysis, whether counterfeiters increase or decrease demand for the genuine brand is an empirical question.

Differentiating the left side of (41) with respect to X yields

$$\frac{d\pi_x}{dX} = -(n - e)g > 0. \tag{45}$$

This implies that the presence of counterfeiters shifts the reaction curve of the first duopolist to the left from $r(x)$ to $r'(x)$, as shown in Figure 15.8.

Proposition 7. Assume that counterfeiters sell unauthorized copies of only one duopolist's product. If $g > 0$, then the presence of counterfeiters reduces the output of the brand they copy and increases that of the other brand. On the other hand, if $g < 0$, the counterfeiters increase the output of the brand they imitate and decrease the output of the other brand they do not copy.

This proposition suggests that the counterfeiter is an enemy of the trademark holder when $g > 0$, but is a friend if $g < 0$. In the latter case, the presence of counterfeiters helps the trademark owner grab a larger market share by making consumers more familiar with the brand. In this case, counterfeiters help the trademark owner by increasing the demand for the brand they copy.

How does law enforcement affect the Nash equilibrium? Differentiating (41) with respect to e gives

$$\frac{\partial\pi_x}{\partial e} = gX. \tag{46}$$

Thus, an increased police force shifts the first duopolist's reaction curve to the right if $g > 0$. Consequently, the Nash equilibrium output of the first duopolist increases and that of the second declines. On the other hand, if $g < 0$, then an increased police force shifts the first duopolist's reaction curve to the left from $r(x)$ to $r'(x)$, as shown in Figure 15.8.

Proposition 8. Assume that counterfeiters sell only illegal copies of one duopolist's product. If $g > 0$, then an increased police force increases the output of the duopolist

whose brand they imitate and decreases that of the other duopolist. If $g < 0$, then an increased police force decreases the output of the former and increases the output of the latter.

6.2 Long-Run Effects of Fines and Arrests

From the first-order conditions of the duopolists and the counterfeiter, it is clear that a fine does not directly affect the outputs of counterfeiters or genuine brand producers in the short run. However, from (21), a fine does reduce expected profit and discourages entry. Thus, eventually there will be fewer firms. Similarly, from (22), more rigorous law enforcement decreases expected profit and discourages entry. However, the results in Proposition 5 only compare the number of counterfeiters in the short and long run. They do *not* show how the *long-run equilibrium number of counterfeiters* responds to changes in fines or arrests.

We now consider the effects of fines and law enforcement agents only on the long-run equilibrium number of counterfeiters. A long-run Nash equilibrium satisfies the zero profit condition for counterfeiters:

$$E\Pi(X, x, y) = (A - (n - e)BX - Fy - Gx)(1 - e/n)X - C(X) - eT = 0. \tag{47}$$

Differentiating (47) partially yields

$$\frac{\partial n}{\partial T} = -\frac{E\Pi_T}{E\Pi_n} \quad \text{and} \quad \frac{\partial n}{\partial e} = -\frac{E\Pi_e}{E\Pi_n},$$

where $E\Pi_T < 0$, $E\Pi_e < 0$, and

$$E\Pi_n = \frac{X}{n}(-nBX(1 - e/n) + Pe/n). \tag{48}$$

Thus, the effects of fines and law enforcement on the number of counterfeiters are generally ambiguous.

Proposition 9. An increase in fines or law enforcement agents increases or decreases the long-run equilibrium number of counterfeiters according to whether

$$Pe - nBX(n - e)$$

is positive or negative.

7 Subsidized Law Enforcement

China signed the Madrid Pact for trademark protection in 1989, and acceded to the Bern Convention and Universal Copyright Convention in 1992. However, enforcement of intellectual property rights since then has been a problem. The concept of property ownership, much less intellectual property rights, has not been well defined in China and other Asian countries. It may take a while for them to respect intellectual property rights as fully as do high-income countries in Europe and North America.

Free trade often requires trading countries to respect the intellectual property rights of other countries. Developed countries have a mutual interest in protecting their patent rights. However, governments of developing countries do not directly benefit from granting operating rights to foreign firms who want to extend their monopoly rights to these countries. They may acknowledge them only under pressure to gain access to foreign markets. Protection of intellectual property rights of multinational firms headquartered in Western countries shuts out domestic counterfeiters that contribute to the national economies of low-income countries. Moreover, foreign governments may argue that counterfeiting is their domestic problem, as long as they do not export counterfeit products to the Western countries. Accordingly, governments of low-income countries may lack a strong incentive to enforce copyrights or trademarks when their firms infringe on the intellectual property rights of Western firms. It is noteworthy that enforcement of copyrights and trademarks is not a problem in Japan – a high-income Asian nation – but is a problem in low-income Asian countries.

Governments of low-income countries might be persuaded to enforce trademarks more rigorously if home firms were more willing to bear the cost. Even if the identities of infringers are exposed, foreign governments are in no hurry to use their resources to enforce US trademarks. In this situation, it may be wise for trademark holders to assume some of the enforcement costs necessary to reduce the number of counterfeiters. Alternatively, trademark owners may hire private agents to expose and prosecute counterfeiters. For example, Kaikati and LaGarce (1980) noted that firms are hiring detective services to expose and prosecute counterfeiters, even though prosecution does not guarantee that the infringers will cease counterfeit production. Higgins and Rubin (1986) further showed that private law enforcement results in too much enforcement and lower social welfare than are socially optimal.

Suppose the monopolist pays the cost the foreign government incurs to hire policing agents or hires his own detective agents to prosecute counterfeiters. Let $e(L)$ denote the number of counterfeiters arrested. All arrested counterfeiters are assumed to disappear and do not return to the market in the short run. The arrest function is an increasing and concave function of policing agents hired, L. That is, as more agents are hired, the number of counterfeiters caught increases but at a decreasing rate.

Recall that the monopolist is selling the product in a foreign country and cannot actively pursue counterfeiters. It can only assist the foreign government in catching the counterfeiters by incurring the cost wL. The monopolist's profit is then

$$\pi = (a - bx - (n - e(L))gX)x - c(x) - wL. \tag{49}$$

The first-order conditions are:

$$\pi_x = a - 2bx - (n - e)gX - c_X = 0, \tag{50}$$

$$\pi_L = e'(L)gXx - w = 0, \tag{51}$$

where $e'(L)gXx$ is the marginal benefit of hiring a law enforcement agent. As long as the marginal benefit of employing an agent is greater than the wage, it pays for the monopolist to hire policing agents to expose and eliminate counterfeiters. It is possible for the monopolist to neutralize all counterfeiters so that $n = e(L)$ but, at that level, profits are lower due to the increasing cost of apprehending those who infringe on the trademark.

Proposition 10. Assume that the monopolist bears the cost of law enforcement and assists the foreign government in hiring policing agents to prosecute counterfeiters. Then (i) optimal policing occurs where $e'(L)gXx = w$, and (ii) exterminating counterfeiters is not optimal.

Proposition 10 suggests that when the trademark owners bear the cost of law enforcement, it is not optimal to eliminate all counterfeiters. This is consistent with empirical observation. Trademark owners do need to prosecute some infringers, because inaction may cause them to lose the trademark. However, trademark owners need not prosecute all infringers.

8 Tolerating Counterfeits

Crampes and Langinier (2002) consider three options. The patent holder may do nothing (renunciation solution). It may settle with the infringer for a sum, or prosecute the infringer. Even when the trademark holders pay the full cost of litigation and law enforcement, the foreign government may not be willing to shut down counterfeiters that contribute to its national economy. Thus, prosecution does not necessarily guarantee that the infringers will not reenter the market later. This may be especially true when counterfeiters are numerous. Also, it may not be politically expedient for the home government to retaliate or threaten economic sanctions against the infringing nation that houses the counterfeiters.

In this situation, trademark holders may learn to live with counterfeiters. Specifically, trademark owners may settle for a fee and tolerate the infringers instead of litigating and prosecuting them. A settlement is arranged usually after an infringer is caught. In this section, we consider tolerance, an ex ante settlement, before the infringer is caught. Tolerance also should be distinguished from licensing to transfer technology in return for a royalty payment. The settlement fee may be paid for the right not to be prosecuted. It does not mean the counterfeiters are authorized to use the trademark.

Suppose the monopolist issues a permit not to prosecute a counterfeiter in return for a sum W. Tolerance does not authorize a counterfeiting firm to produce counterfeits. In a sense, tolerance is like the indulgences that the Roman Catholic Church issued

in the fifteenth century to raise funds to build St. Peter's Basilica. Indulgences did not authorize believers to commit sins, but supposedly forgave sinners after the fact. In contrast, counterfeiters may register for a settlement before counterfeits are produced. It is only a guarantee not to persecute the registered counterfeiters. This policy may help trademark owners to monitor counterfeiters. Whether this is legally feasible is an entirely different question.

Consider a counterfeiter contemplating registration ex ante. Registration is a form of insurance to the counterfeiters. It is costly but it eliminates the risk of confiscation ex post. If the counterfeiter is risk averse, the maximum amount it is willing to pay to avoid prosecution is implicitly defined by

$$(1-\sigma)U(PX-C)+\sigma U(-C-T)=U(PX-C-W). \tag{52}$$

Thus, the maximum amount a risk-averse counterfeiter is willing to pay is the sure profit without the risk of prosecution,

$$W^{\max}=PX-C(X). \tag{53}$$

The maximum amount a risk-neutral counterfeiter is willing to pay is defined implicitly by

$$(A-(n-e)BX-Gx)(1-\sigma)X-C(X)-\sigma T=(A-(n-e)BX-Gx)X-C(X)-W.$$

Thus, the tolerance fee he is willing to pay is the expected loss he would incur without registration, that is, the loss of revenue and fine weighted by the probability of arrest,

$$W=\sigma(PX+T). \tag{54}$$

This amount is less than the amount a risk-averse counterfeiter is willing to pay. Thus, the trademark owner can ask any amount W in the following range:

$$PX-C(X)\geqslant W\geqslant\sigma(PX+T). \tag{55}$$

Once the tolerance fee is announced, the trademark owner cannot discriminate among counterfeiters. Registration of non-prosecuted counterfeiters may thus be voluntary. Let N and n denote the numbers of protected and unprotected counterfeiters, and let Y and X denote the outputs of protected and unprotected counterfeiters, respectively. Both types of counterfeiters are assumed to produce identical products and their prices are the same. Total profits of N tolerated counterfeiters are:

$$N\widetilde{\Pi}=N(PY-C(Y))-NW, \tag{56}$$

where W is the fee the protected counterfeiters pay. A protected counterfeiter is willing to pay a fee,

$$W\leqslant PY-C(Y).$$

The monopolist's revenue is then

$$\pi=(a-bx-gNY+(n-e)X)x-c(x)+NW. \tag{57}$$

The first-order condition is

$$\frac{\partial \pi}{\partial x} = -bx + p - c'(x) = 0. \tag{58}$$

Recall that the trademark owner does not choose the number of registrants. Differentiating (57) with respect to N gives

$$\frac{\partial \pi}{\partial N} = -gYx + W, \tag{59}$$

where gYx is marginal revenue loss from issuing a certificate of tolerance. Thus, as long as the registration fee exceeds marginal loss, it pays for the monopolist to issue tolerance certificates. However, the right-hand side of (59) does not depend on N. The actual number of registered counterfeiters is determined by the counterfeiters.

Proposition 11. Issuing tolerance certificates not to prosecute registered counterfeiters is an optimal policy for the monopolist if $PY - C(Y) \geqslant W > gYx$.

Recall that the amount a risk-neutral counterfeiter is willing to pay for the tolerance permit is $\sigma(PX + T)$. It should be noted that as the number of registered counterfeiters increases, price declines and the representative counterfeiter's output shrinks as well.

9 Socially Optimal Counterfeiting

We now consider a seemingly outrageous idea. The ultimate question is whether counterfeiting is harmful to a society and should be eliminated at all costs, or whether counterfeit products yield some benefits to society and hence counterfeit production should be controlled. Recall that we are not dealing with firms that produce intrinsically harmful products and that the marginal utility of the counterfeit product is positive.

Consider the joint profits of the monopolist and counterfeiters. Adding profits in (56), (57), and (31), we get

$$\begin{aligned}\pi + (N\widetilde{\Pi} + n\Pi) = {} & (a - bx - NgY - (n-e)gX)x - c(x) + NW + NPY - NC(Y) \\ & - NW + (A - NBY - (n-e)BX - Gx)(n-e)X - nC(X) - eT.\end{aligned}$$

Since the side payment NW is canceled out, the aggregate profit reduces to

$$\Theta \equiv px - c(x) + N(PY - C(Y)) + P(n-e)X - nC(X) - eT. \tag{60}$$

Differentiating (60) with respect to x and X, and retaining the monopolistic behavior of the genuine producer and competitive behavior of counterfeiters, we obtain

$$\Theta_x = -bx + p - c'(x) = 0,$$

$$\Theta_X = P(n-e) - nC'(X) = 0,$$

and

$$\Theta_Y = N(P - C') = 0.$$

Note that in long-run equilibrium, the protected firms earn zero profits after the tolerance fee, and hence

$$W = PY - C(Y). \tag{61}$$

Differentiating (60) with respect to N and utilizing these first-order conditions, we obtain in long-run equilibrium,

$$\frac{\partial \Theta}{\partial N} = \frac{\partial p}{\partial N} x + (PY - C) + (NY + (n - e)X) \frac{\partial P}{\partial N} = W - (gx + NBY + (n - e)BX).$$

When evaluated at $N = 0$, this reduces to

$$\frac{\partial \Theta}{\partial N} = (W - gx) - (n - e)BX. \tag{62}$$

The first term on the left side of (62) is positive but the second term is negative. Thus, if (62) is positive, there is a socially optimal number of counterfeiters, and extermination of counterfeiters is not optimal.

Crampes and Langinier (2002) delineated a condition under which renunciation or abandonment would be optimal. We now consider the case of free registration or laissez-faire. Note that when $W = 0$, all counterfeiters are registered and none are arrested ($n = e = 0$). Evaluate (62) at $W = 0$, that is, when no fees are collected:

$$\frac{\partial \Theta}{\partial N} = -gx < 0. \tag{63}$$

Thus, if there is no tolerance fee, all counterfeiters are registered and full registration is not socially optimal. That is, renunciation is not optimal.

Proposition 12. Issuing certificates to tolerate counterfeits is socially optimal. There is a socially optimal number of counterfeiters if

$$W > gx + (n - e)BX.$$

Moreover, free registration is not socially optimal.

While the tolerance fee transfers income from counterfeiters to the trademark holder, it also may improve social welfare and efficiency of resource allocation. However, some fee should be assessed, equal to the sure profits of counterfeiters, for socially optimal resource allocation.

10 Concluding Remarks

Grossman and Shapiro (1988a, p. 60) noted that "Counterfeiting undermines the functioning of the property rights system" for three reasons. First, consumers suffer harm

by purchasing low-quality copies purported to be originals. Second, consumers face a risk of obtaining fakes. Third, consumers may not recognize a bogus good and may attribute its poor quality to the trademark owner.

In this chapter we have examined consumer demand for counterfeits when they purchase these counterfeits knowingly. Consumers face no risk because they are aware that their purchases are fakes, and they do not attribute poor performance, if any, to the trademark owner. Often these counterfeits are functionally, if not aesthetically, identical to the original.

There may be only one self-portrait of Vincent Van Gogh at the National Gallery of Art in Washington, DC. If it is copied and prints are made commercially without the permission of the owner, consumers may purchase them knowing full well that they are copies. It also is a well-known fact that many paintings of Van Gogh began to appear after his death, and even experts have difficulty discerning the fakes.

Dissemination of copies of original paintings also may enhance their value as consumers become more familiar with them. This may be especially true when many paintings are vying for consumer attention. Copies may not necessarily damage the value of the original paintings. Similarly, counterfeit copies of Rolex watches may not reduce the demand for genuine Rolex watches.

When counterfeiting firms sell unauthorized copies of European paintings in countries where law enforcement is lax, detection and prosecution of counterfeiters may be costly. Owners may well benefit more from a tolerance fee and counterfeiters may be willing to pay the fee to avoid prosecution. Since copies are not intrinsically harmful but have positive marginal utilities, extermination of counterfeits may not be socially optimal.

In a sense, counterfeit products are like goods that generate some pollution and impose spillover costs on other producers. Control, rather than extermination, of counterfeits may be socially optimal, just as regulation, rather than elimination, of pollution is socially optimal. Pollution control is socially accepted in the United States and Europe, but regulation of counterfeits is not. Without this understanding of the similarity between pollution generating production and counterfeiting, rigorous enforcement of intellectual property rights on foreign soil could cause friction with our trading partners, especially in Asian and Middle Eastern countries where property rights historically have not been well defined. The majority of intellectual property rights are owned by high-income countries, and hence they have mutual interest in protecting their copyrights and trademarks. However, low-income countries that possess few trademarks have little interest in protecting or rigorously enforcing the intellectual property rights of Western firms. They do not benefit directly from extending and protecting the patent rights of developed economies.

Acknowledgments

I am indebted to James Hartigan and Henry Thompson for their valuable comments. Much of the work was inspired and completed while I was visiting City University of Hong Kong in 2002. The usual caveats apply.

Appendix

Let $A(z) = -U''(z)/U'(z)$ denote the Arrow–Pratt absolute risk aversion, where $z = x + hX$. Assume that the consumer exhibits DARA, i.e., $A'(z) < 0$. Let z^o denote the weighted quantity when $h = P/p$. Then since $A(z)$ is decreasing,

$$A(z) \leqslant A(z^o) \quad \text{for } h \geqslant P/p.$$

Thus,

$$-\frac{U''(z)}{U'(z)} \geqslant A(z^o) \quad \text{for } h \geqslant P/p. \tag{64}$$

Also, observe that

$$-U'(z)(h - P/p) \leqslant 0 \quad \text{for } h \geqslant P/p. \tag{65}$$

Now multiply (64) by the left side of (65) to get

$$U''(z)(h - P/p) \geqslant -A(z^o)gU''(z^o)(h - P/p).$$

This condition holds for $h \geqslant P/p$, and also for $h < P/p$. In the latter case, the inequalities in (64) and (65) are reversed. Taking expected values, we get

$$EU''(z)(h - P/p) \geqslant A(z^o)EU'(z)(h - P/p) = 0. \tag{66}$$

The latter equality in (66) follows from the first-order condition.

Notes

1 Contestable market as popularized by Baumol (1982) also may be added to this list.
2 According to the US Census Bureau (2000), the US trade deficit with China has hovered around \$80 billion since 2000.
3 This activity must be distinguished from parallel exports or imports in that marketing of extra output in China or other countries is unauthorized. See, for instance, Maskus and Chen (2002) for parallel imports.
4 Counterfeit drugs may be harmful to the human body and the use of counterfeit airplane parts may be hazardous to air travelers.
5 Higgins and Rubin (1986) also consider the case where consumers are deceived and think that they are buying genuine goods when they are not.
6 In this chapter, we consider a single consumer who purchases both the genuine product and the counterfeit good. See Grossman and Shapiro (1988b) for a model with different tastes for quality.
7 See Yang and Maskus (2001a, 2001b) for models with licensing.
8 This scenario practically creates one world market, and whether counterfeiters exist and operate only in the foreign market is immaterial.
9 For instance, in the United States an infringer is subject to four years of imprisonment and/or \$250,000.
10 Recall that (17) shows that along this curve, $E\Pi_X > 0$. Thus, the reaction curve of a risk-averse counterfeiter is below that of the risk-neutral counterfeiter.
11 This is because the fine itself is a lump sum. If the fine increases with the volume of counterfeiting, it will affect output.

12 It is immaterial whether its reaction curve as a Nash player is steeper than the counterfeiter's reaction curve, since it is not used.
13 It should be noted that in a sequential game in which the monopolist moves first, this Stackelberg equilibrium also is a Nash equilibrium.
14 This is because positive slopes of cross-demand curves guarantee that $g > 0$ and $G > 0$ when there are only two products, x and X. This amounts to forcing the signs of the elements of the inverse of a 2×2 matrix. However, when there are more products, knowing the signs of a 3×3 or 4×4 matrix does not enable us to predict the signs of its inverse. Let

$$\mathbf{x} = \mathbf{A}\mathbf{p} + \mathbf{b},$$

where $\mathbf{x} = (x_1, x_2, x_3)'$ and $\mathbf{p} = (p_1, p_2, p_3)'$ are output and price vectors, $\mathbf{b} = (b_1, b_2, b_3)$ is a constant vector, and $\mathbf{A}$ is an $n \times n$ positive matrix with diagonal elements $a_{ii} = \partial x_i / \partial p_i < 0$ and off-diagonal elements $a_{ij} = \partial x_i / \partial p_j > 0$, $i \neq j$. The price vector is written as

$$\mathbf{p} = \mathbf{B}(\mathbf{x} - \mathbf{b}),$$

where $\mathbf{B} = \mathbf{A}^{-1}$. Then, for instance,

$$p_1 = b_{11}(x_1 - b_1) + b_{12}(x_2 - b_2) + b_{13}(x_3 - b_3).$$

The assumption that $\mathbf{A}$ is a matrix with negative diagonal and positive off-diagonal elements does not guarantee any definite signs of b_{12} or b_{13}. Thus, the sign of g in (35) is generally ambiguous.
15 Producing more output is not helpful due to declining marginal revenue. Counterfeiters "advertise" the genuine product and hence may produce positive spillover benefits to the brand they imitate.

References

Baumol, William 1982: Contestable markets: an uprising in the theory of industry structure. *American Economic Review*, 72, 1–15.
Crampes, C. and C. Langinier 2002: Litigation and settlement in patent infringement cases. *RAND Journal of Economics*, 33, 258–74.
Crampes, Claude and Jean Jacques-Laffont 2002: Copying and software pricing. Working paper, University of Toulouse.
Grossman, Gene and Carl Shapiro 1988a: Counterfeit-product trade. *American Economic Review*, 78, 59–75.
Grossman, Gene and Carl Shapiro 1988b: Foreign counterfeiting of status goods. *Quarterly Journal of Economics*, 103, 79–100.
Higgins, Richard and Paul Rubin 1986: Counterfeit goods. *Journal of Law and Economics* 29, 211–30.
Ishibashi, Ikuo and Toshihiro Matsumura 2001: Patent rate in a mixed market. Working paper, University of Tokyo.
Kaikati, Jack and Raymond LaGarce 1980: Beware of international piracy. *Harvard Business Review*, 58, 52–60.
Maskus, Keith E. and Yongmin Chen 2002: Parallel imports in a model of vertical distribution: theory, evidence, and policy. *Pacific Economic Review*, 7, 319–34.
Merrill, William and Norman Schneider 1966: Government firms in oligopoly industries: a short run analysis. *Quarterly Journal of Economics*, 80, 400–12.

U.S. Census Bureau 2002: *US Trade Balance with China*. Accessed December 29, http://www.census.gov/foreign-trade/balance/c5700.html

U.S. International Trade Commission 1984: *The Effects of Foreign Production Counterfeiting on U.S. Industry*. USITC Publication No. 1479.

Yang, Guifang and Keith Maskus 2001a: Intellectual property rights, licensing, and innovation in an endogenous product-cycle model. *Journal of International Economics*, 53, 169–87.

Yang, Guifang and Keith Maskus 2001b: Intellectual property rights and licensing: an econometric investigation. *Weltwirtschaftliches Archiv*, 137, 58–79.

16

Endogenous Injury

James C. Hartigan

Chapter Outline

An economic justification for including an injury requirement in antidumping (AD) law is provided through a finite horizon two-stage model with home and foreign duopolists. Cost reducing technology arrives at each period. Firms make a decision to install or hold an option to install the technology in stage one of each period. They are Nash–Cournot competitors in stage two. Although the firms initially are technically symmetrical, a low (high) national market demand realization induces the home (foreign) firm to not invest (invest) in the new technology arriving in period one. As the cost of installing the new technology is higher for the lagging firm, cyclical reciprocal dumping can generate endogenous injury through inducing a permanent technology disparity. The recommendation is that injury investigations focus upon the rates of industry demand cyclicality and innovation, and the vintage of firms' capital stock. While providing justification for injury-based AD measures, this chapter contends that conditions meriting such measures are uncommon.

1 Introduction

In the US antitrust case against Microsoft, there has been a departure from the traditional emphasis on pricing. In what is potentially a landmark case, the focus is upon the monopoly deterrence of innovation. This can be manifest in at least three ways: (1) denial of access to a network, (2) collusive withholding of innovations, and (3) reduction in research and development expenditure concomitant to a merger or acquisition. See Mandel, France, and Carney (2000).

It is the contention of this chapter that innovation can, and perhaps should, be considered in trade policy decisions. The context that I suggest is in the material and serious injury verdicts of investigations under administered protection law. The legal requirements for the finding of injury are rather vague, and focus upon a wide array of variables. This includes profits, sales, outputs, inventories, R&D expenditures, capacity utilization, and employment of production and nonproduction labor. I aim to provide more discipline to the injury decision, and to disclose that injury can arise in certain contexts as a concomitant to international trade. I devise a discrete time finite horizon (three periods) duopoly model in which each period consists of two stages. In the first stage, a firm is faced with an investment decision. This decision takes place because an innovation arrives in every period. The extent of the innovation is, however, uncertain. The source of the innovation is external to the firm. That is, I do not address a research and development race. The second stage is the Nash–Cournot output decision. The equilibrium concept is subgame perfection. There are home and foreign markets, in each of which a duopolist is domiciled. There is an international shipping cost. Hence each firm has a larger share of its national market than does its rival. This is consistent with the Brander and Krugman (1983) and Murray and Kurdaliev (1999) reciprocal dumping models. Demand realizations in each period can be high or low, and are not correlated between markets. This is consistent with the cyclical dumping models of Ethier (1982), Bernhardt (1984), Das (1992), and Staiger and Wolak (1992a).

When a firm must make, and chooses to postpone, an investment decision in the presence of a low demand realization in its national market, it may induce a technological lag in the case of innovation. This requires the rival firm to invest in the technology as a result of a high demand realization in its primary market. I suppose that the cost of adopting the new technology is greater for a lagging firm than for one which was up to date at the time of innovation. This disparity can induce a permanent technology gap and ultimately exit by the lagging firm. This is more apt to occur in an industry in which innovation is frequent.

Consideration of exit and trade policy in the context of declining industries has been given by Cassing and Hillman (1986), Van Long and Vousden (1991), Anis et al. (2002), Brainard (1994), and Brainard and Verdier (1994, 1997). In this chapter, however, the industry is not declining.[1] Hence, this is not a last to exit problem. The industry is viable as a duopoly. My demonstration thus provides an argument for the use of administered protection laws when the domestic firm incurs competition from a foreign rival. However, the requirements for the manifestation of endogenous injury are rather unlikely to hold simultaneously. Thus, while justifying the use of these laws, I also contend that their utilization is abusive. A firm that is not mismanaged and is in a long-term viable industry nonetheless is in a weakened position due to a confluence of events outside of its control.[2] This technology gap and the possibility of exit would not occur in the absence of international trade. The relative inefficiency of a firm arises endogenously, and may be deemed a manifestation of injury. Policy options available include an investment subsidy, or import duties. Either of these can raise the sum of discounted profits that the firm expects to accrue.

The production and investment stages of the model are developed in Section 2 and policy implications are considered in Section 3. Subsequently, a conclusion is given.

2 A Simple Model

2.1 Production

Suppose that there are two countries, each of which has a resident firm. The countries and firm are designated as home (h) and foreign (f). Each firm consists of a unit of capital. It is installed at a cost of $r^i_{\Im} \in \Re^+$, where $\Im \in Z^+$ (the set of positive integers), and $r^i_{\Im'} < r^i_{\Im}$ for $\Im' < \Im$, and $i = h,f$.[3] The subscript $\Im$ refers to the currency of the firm's technology. That is, a firm using the most up to date technology when it replaces its unit of capital does so at a cost of r_1. (Technology is embodied in the unit of capital.) If it was not utilizing the most recent technology, then $r^i_{\Im} > r^i_1$. Thus if there was one innovation since the firm replaced its unit of capital, the cost of doing so is r^i_2. (It was using the second-best technology.) For simplicity, I will assume that the differential in the costs of installing units of capital embodying different techniques is a function only of the number of techniques advanced. That is, $r_{\Im''} - r_{\Im'} = r_{\Im'} - r_{\Im}$ for $\Im'', \Im', \Im \in Z^+$ and $\Im'' = \Im' - 1 = \Im - 2$. Denoting the output of firm i in any period t as q^i_t, its cost function is given by

$$C^i_t = v^i_{\Im} q^i_t + r^i_{\Im} \tag{1}$$

if its unit of capital is replaced, and

$$C^i_t = v^i_{\Im} q^i_t \tag{2}$$

if capital is not replaced. Variable costs of production are $v^i_{\Im}$, where $\Im$ denotes the level of technology. If $\Im'$ is a more current technology than $\Im$, then $v^i_{\Im'} < v^i_{\Im}$.

The availability of a more productive unit of capital is the only reason for a firm to install a new unit.[4] Capital never wears out. However, unless the innovation were of sufficient magnitude to reduce cost to a level that permitted monopolization of the market by the rival if a firm did not invest, a firm can continue to produce without undertaking investment. It will however, lose market share if its rival invests and it does not. If its rival purchases a new unit, a firm will be at a competitive disadvantage (because of higher variable costs) if it does not replace its capital. Innovations are generated externally to the firm. This is for modeling simplicity. However, it is not unreasonable for many industries.[5]

I assume that there are two realizations of demand in each market: high and low. These realizations are not correlated over time and across markets (countries). I further suppose that each firm's unit of capital provides capacity sufficient to satisfy consumer demand in both markets when each demand realization is high. A firm must incur a shipping cost of s per unit to export its product to the other country. Let demand in the home market be given by

$$P^h_t = a_t - b(q^h_t + q^f_t), \qquad a_t \in \{\underline{a}, \bar{a}\}, \quad \bar{a} > \underline{a} \text{ in } \Re^+ \tag{3}$$

in each period t. The pertinent variables and parameters in the foreign market are denoted by an asterisk (*). The firms play a Nash–Cournot game in each period in each

market, so long as neither of them has chosen to exit rather than replace its unit of capital. As can be inferred from (3), the firms produce a homogenous good, and the high (low) demand realization is given by $\bar{a}(\underline{a})$. Equation (3) can be combined with the cost functions of (1) or (2) to form a profit function. For the home market, the gross profit functions are

$$\pi_t^h = (a_t - b(q_t^h + q_t^f))q_t^h - v_{\Im}^h q_t^h \text{ and } \pi_t^f = (a_t - b(q_t^h + q_t^f))q_t^f - (v_{\Im}^f + s)q_t^f. \quad (4)$$

These profits are calculated without taking into account the costs of capital replacement. Profit functions for the foreign market are specified analogously. Differentiating with respect to own output, and setting the results to equality with zero permits derivation of the best response functions of the stage game for each firm in each market. These can be solved simultaneously for the Nash–Cournot stage game equilibrium. For the home market, this is

$$\underline{q_t^h = (a_t - 2v_{\Im}^h + v_{\Im}^f + s)}_{3b} \quad \text{and} \quad \underline{q_t^f = (a_t - 2v_{\Im}^f + v_{\Im}^h - 2s)}_{3b}. \quad (5)$$

The equilibrium in the foreign market is derived analogously.

Two observations are apparent immediately. The first is that the international shipping costs s induce an asymmetric equilibrium, in which each firm has a larger share of its national market, when the firms incur the same variable costs. The second is that the firm's relative share of a market increases as it obtains a lower relative variable cost.

Given the equilibrium outputs in each market, market prices can be obtained from (3) and its foreign market analogue. This permits the calculation of profits for the home market (4). Its unstated analogue allows the computation of foreign market profits. Starting from an assumption of identical variable and international shipping costs, as well as identical high and low inverse demand function intercepts and inverse demand function slopes in each market, I can infer that the profits of the duopolists are identical when the same (high and low) demand realizations occur in each market. However, when demand is high in one market and low in the other, the firm domiciled in the market with high demand will earn greater profits than the firm whose national market realizes low demand in that period. This has interesting implications for the investment decision of the firms.

2.2 Investment

Recall the circumstance under which a firm would face an investment[6] decision. Each firm is assumed to possess the best available technology as play begins in period one. That is, the firms are identical initially. They are assumed to be able to produce with unit costs $v_{\Im}^i$. An innovation occurs in stage one of period one, permitting production with unit costs of $v_{\Im'}^i$, where $v_{\Im'}^i < v_{\Im}^i$, for $i = h,f$. Each firm considers an investment (the purchase of a new unit of capital) by taking into account current period profits, as well as the continuation payoff of expected Nash–Cournot duopoly profits in the subsequent periods. Current profits are given by (4) after substitution from (3) and the solution to (5) for the home market. Recall that the equations

for the foreign market are suppressed. To determine the expectation, denoted by E, of future profits, the firms must know the density functions of the demand realizations for the two markets. Term these $f(a_t)$ and $f(a_t^*)$, and let $f(\underline{a}) = f(\underline{a}^*)$. The firms know that the current technology will be superseded by a superior manifestation with near certainty in the ensuing period. What is uncertain is the extent of the cost advantage that it provides. Since the best technology for producing the good in period 1 permits unit costs of $v_{\Im'}$, define $G(v_{\Im''}/v_{\Im'})$ as the conditional probability distribution of unit variable costs under the period 2 innovation. Note that because the source of the innovation is external to the firm, $G(\cdot)$ is conditional on the best available technology, irrespective of whether or not it has been adopted by either firm. This distribution has support given by $[0, v_{\Im'}]$, and density $g(\cdot)$. Let $g(v_{\Im''}/v_{\Im'}) = 1/v_{\Im'}$. Hence G is the uniform distribution. Under this distribution, the probability of a significant innovation is equal to that of a minor one, as all innovations are equally likely. Although this may be somewhat objectionable, the uniform is the easiest distribution with which to work. The expectation of G, held in period 1, is $v_{\Im'}/2$.

From the density functions $f(a_t)$ and $f(a_t^*)$, and from (3), (4), and (5), the expected duopoly profits can be calculated for each period and each country. These can be calculated for any pair of variable costs for the firms.

To solve the game, let us denote current (period t) Nash–Cournot equilibrium (net of the costs of capital replacement) by

$$\pi_t^i = \pi_t^i(a_t, a_t^*, v_{\Im'}^{\mathrm{h}}, v_{\Im'}^{\mathrm{f}}, r_{\Im'}^i), \qquad i = h,f \tag{6}$$

when both firms invest in the technology, permitting unit costs of $v_{\Im'}$, that becomes available in period 1. When either or both firms do not invest, $v_{\Im}^i$, $i = h,f$, replaces $v_{\Im'}^h$ and/or $v_{\Im'}^f$ in (6) accordingly. In order to keep the problem manageable, and to demonstrate the point that I wish to make about endogenous injury in the context of dumping, I will suppose that a firm finds it unprofitable to invest in the newly available technology when its national market demand realization is low and both firms produce with symmetric (identical) technologies. That is, when $a_t = \underline{a}$ and $v_{\Im}^h = v_{\Im}^f$, the home firm will not invest. When its national market demand realization is low, it holds an option to invest in the subsequent period.[7]

The expectation, held in period t, of Nash–Cournot profits in period $t+1$ as a result of a period 1 investment is

$$\begin{aligned}
\mathrm{E}_t^i(\pi_{t+1}^i) = \delta\big[& f(\bar{a})f(\bar{a}^*)\pi_{t+1}^i\big(\bar{a}, \bar{a}^*, v_{\Im'}^i/2, v_{\Im'}^j/2, r_{\Im''}^i\big) \\
&+ f(\bar{a})f(\underline{a}^*)\pi_{t+1}^i\big(\bar{a}, \underline{a}^*, v_{\Im'}^i/2, v_{\Im'}^j, r_{\Im''}^i\big) \\
&+ f(\underline{a})f(\bar{a}^*)\pi_{t+1}^i\big(\underline{a}, \bar{a}^*, v_{\Im'}^i, v_{\Im'}^i/2, r_{\Im''}^i\big) \\
&+ f(\underline{a})f(\underline{a}^*)\pi_{t+1}^i\big(\underline{a}, \underline{a}^*, v_{\Im'}^i, v_{\Im'}^i, r_{\Im''}^i\big)\big],
\end{aligned} \tag{7}$$

where δ denotes the common discount factor. Let us suppose that for the purpose of evaluating an investment in period t, terms raised to the second power are vanishingly small. Hence, $\mathrm{E}_t^i(\pi_{t+2}^i)$ is not relevant to the calculation of the value of an investment

in period t. Hence the value of an investment in a unit of capital is the sum of the appropriate analogues of (6) and (7). As was noted before, arguments pertaining to each firm's unit costs depend upon whether or not investment takes place.[8]

In order to determine the equilibrium investment decisions of the firms, I invoke subgame perfection. In the third (final) period of the finite horizon game, the profits accruing to an investment in that period are given by an analogue of (6) for both firms, as period $t + 1$ does not exist. The appropriate levels of $v^h_{\Im}$ and $v^f_{\Im}$, as well as a_3 and a^*_3, must be substituted into each firm's profit function. Because of the absence of a continuation payoff for a period 3 investment, I will assume that neither firm invests in the newly available technology in this period.

Now that period 3 equilibrium has been determined, subgame perfection dictates that we proceed to stage 2 of period 2. Solving for the Nash–Cournot outputs, price, and profits require that we know the unit costs (technologies) of the two firms. Recall that the firms are initially endowed with technology $\Im$ at the start of play. Technology $\Im'$ becomes available in period 1, and technology $\Im''$ becomes available in period 2. The firms learn the technology and the demand realization prior to having to make their investment decisions. Recall also the simplifying assumption that firms invest in the new technology only when their national market demand realization is high. This implies the following pairs of unit costs for the duopolists: $(v^h_{\Im}, v^f_{\Im})$, $(v^h_{\Im'}, v^f_{\Im})$, $(v^h_{\Im''}, v^f_{\Im})$, $(v^h_{\Im}, v^f_{\Im'})$, $(v^h_{\Im'}, v^f_{\Im'})$, $(v^h_{\Im''}, v^f_{\Im'})$, $(v^h_{\Im}, v^f_{\Im''})$, $(v^h_{\Im'}, v^f_{\Im''})$, and $(v^h_{\Im''}, v^f_{\Im''})$. Because of the focus in this chapter on endogenous injury arising as a result of demand fluctuations in the context of reciprocal dumping, the only circumstance that is of interest is the one in which the home firm enters period 2 at a technological disadvantage. That is, when the foreign firm has unit costs of $v^f_{\Im'}$ and the home firm has unit costs of $v^h_{\Im}$.

In stage 2 of period 2, the foreign firm can have unit costs of $v^f_{\Im'}$ or $v^f_{\Im''}$, depending on whether or not it invests in the new technology $\Im''$. The home firm will have technology $\Im$ if it does not acquire a new unit of capital. If it does invest, it has the choice of technologies $\Im'$ and $\Im''$. For any level of technology that the foreign firm has selected, the profits gross of the installation costs of the new unit of capital are highest with technology $\Im''$ for the home firm. This is because unit costs are lowest with this technique. However, $r^i_{\Im'} < r^i_{\Im''}$. Further, if the home firm continues to use technology $\Im$, it does not incur installation costs. For any technology that the home firm possesses, the gross profits of the foreign firm are higher with technology $\Im''$ than with technology $\Im'$. It does, however, incur installation costs with $\Im''$. In stage 2 of period 2, each firm knows the demand realization (high or low) in each market, as well as the newly available technology. Further, they know the stage 1 optimal technology choice of their rival. Given these stage 1 technology choices, the optimal outputs for each firm in each market are determined by solution of (5) and its unstated analogue for the foreign market. Prices in each market are given by (3) and its unstated foreign market analogue. Profits are generated by (4) and its unstated analogue. The sum of profits from each market yield total gross profits for each firm.

Now that the equilibrium for stage 2 of period 2 has been ascertained, subgame perfection permits the determination of the investment decision for each firm in stage 1. Recall that under the backward induction of subgame perfection, each firm decides upon its optimal action in a stage based upon its calculation of what its rival will rationally play in the subsequent stage of the game. Each firm makes its period 2 investment decision

knowing that it will generate its profits in both countries in periods 2 and 3. It knows what these profits will be for any technology choice that its rival may make. Expected demand in period 3 in the home market is given by $f(\underline{a})\underline{a} + (1 - f(\underline{a}))\bar{a}$. Expected demand in the foreign market is determined analogously. The period 2 investment decision is given by the sum of (6) and (7) with the appropriate substitutions for unit production costs and the unit of capital installation costs.

There are four pairs of demand-level realizations possible in period 2. These are $(\underline{a}, \underline{a}^*)$, $(\bar{a}, \bar{a}^*)$, $(\bar{a}, \underline{a}^*)$, and $(\underline{a}, \bar{a}^*)$. Recall that to make this chapter relevant to the material injury decision in AD investigations, I will suppose that the home firm utilized technique $\Im$ in period 1, and the foreign firm utilized $\Im'$. I assumed, with symmetric costs, that each firm invests in the new technology when its national market demand realization was high, and held the option to invest when its national demand realization was low. This implies that the foreign firm will install a new unit of capital embodying $\Im''$ if $a_2 = \bar{a}^*$. Expected profits would be greater under $(\underline{a}, \bar{a}^*)$ and $(\bar{a}, \bar{a}^*)$ than they would be under initial symmetry. This is because the foreign firm could, at worst, advance to stage 2 of period 2 at technical symmetry, and its unit costs would be less than in period 1. If play began with the foreign firm having a one-level technology advantage over its domestic rival, it could advance to a two-level lead if the home firm did not invest. Since (3), (4), and (5) disclose that stage game output and profits are inversely related to the rival duopolist's unit production costs, a firm will install the new technology under a high national market demand realization when it has a technical lead if it would do so under technical symmetry. Similarly, the home firm will not invest under a one-level technology disadvantage if it would not do so under technical symmetry when the national market demand realization is low. That is, not investing is a dominant strategy.

If $(\underline{a}, \bar{a}^*)$ is realized in period 2, the foreign firm's technology advantage expands to two levels. This is due to my assumption that a firm will not invest when its national demand realization is low under symmetric technology.

If $(\bar{a}, \bar{a}^*)$ is realized, the home firm has three choices. It can remain at technology level $\Im$. It can adopt the new technology level $\Im''$. Finally it can advance one level of technique, by installing $\Im'$. It is more expensive to install $\Im''$ than $\Im'$, as $r^i_{\Im''} > r^i_{\Im'}$. From technical symmetry and high national market demand, I have assumed that a firm will always invest. However, investing when at a technical disadvantage results in lower stage game profits because of the higher investment cost. Installing $\Im'$ is less costly. However, stage game profits gross of $r_{\Im'}$ will be lower than under technical symmetry, because the foreign firm will install $\Im''$. The home firm can avoid the capital installation charge if it continues to utilize $\Im$. However, its stage game profits gross of investment costs are the lowest with this strategy. The decision of the home firm will be based upon the sum of (6) and (7) with the appropriate substitutions. That is, $\bar{a}$ and $\bar{a}^*$ will be substituted into (6). In addition, $v^f_{\Im''}$ will be substituted into (6) and (7). The current period and continuation profits of the home firm will be determined by alternatively substituting $v^h_{\Im''}$ and $r^h_{\Im''}$, $v^f_{\Im'}$ and $r^f_{\Im'}$, and $v^f_{\Im}$ and zero (the capital is already installed) into (6). The three unit production costs are also substituted into (7). Recall that investment will not take place in period 3, but its expected cost in period 2 is $r^i_{\Im'''}$. The home firm will make the investment decision that yields the highest level of profits.

The solution by backward induction for subgame perfection now proceeds to period 1. I only consider the first period equilibrium in which $\underline{a}$ and $\bar{a}^*$ are the demand realizations, and only the foreign firm invests in the newly available technology $\Im'$. This is to make the game pertinent to consideration of the material injury requirement of AD law. In so doing, I have disclosed that simultaneous realization of $\underline{a}$ and $\bar{a}^*$ can induce material injury through the creation of a permanent technological disadvantage for the home firm in a model of cyclical and reciprocal dumping.

Now let us consider the realization $(\underline{a}, \underline{a}^*)$ in period 2. From initial technical symmetry, neither firm would invest in the newly available technology. The home firm will certainly not invest when it is at a technical disadvantage, if it would not do so from symmetry. As a result, not investing may not be a dominant strategy for the foreign firm. If it invests, it can establish a two-level technique advantage over the home firm. This will enable it to earn higher stage game profits in period 2, relative to the symmetric case. Once again, profits are compared by substituting $\underline{a}$, $\underline{a}^*$, and $v^h_{\Im}$ into (6). The latter is also substituted into (7). Now we can alternatively substitute $v^f_{\Im''}$ and $r^f_{\Im''}$, and $v^f_{\Im'}$ and zero into (6). The appropriate substitutions into (7) are $v^f_{\Im''}$ and $r^f_{\Im''}$, and $v^f_{\Im'}$ and $r^f_{\Im''}$. (Recall, however, that there is no investment in equilibrium in period 3.) The foreign firm will implement the strategy yielding the highest stage game and continuation profits. Once again, material injury to the home firm can arise, as it will begin period 3 trailing the foreign firm by either one or two levels of technology.

The realization of $\bar{a}$ and $\underline{a}^*$ offers the only opportunity for the home firm to restore symmetry when it begins play in period 2 at a technical disadvantage. Investment by the home firm can be in either $\Im'$ or $\Im''$. Investment in $\Im''$ would reestablish technical symmetry if the foreign firm also invested, but I have assumed that not investing is a dominant strategy for a firm at technical symmetry in a low national market demand realization. Thus investment by both firms in $\Im''$ is not an equilibrium. If the home firm invested in $\Im'$, the foreign firm would still have a one-level technology lead if it invested in $\Im''$. If the home firm invested in $\Im'$ and the foreign firm did not invest, technical symmetry would be restored, albeit with neither firm utilizing the best technology. Investment by the home firm in $\Im''$ would give it a technological advantage over the foreign firm. The realization of the pair $(\bar{a}, \underline{a}^*)$ is more difficult to analyze at this level of generality than the pair $(\bar{a}, \bar{a}^*)$ in period 2. For the latter, the foreign firm will invest in $\Im''$. This is not necessarily the case with $(\bar{a}, \underline{a}^*)$. The profits for each firm are lower with $(\bar{a}, \underline{a}^*)$ than with $(\bar{a}, \bar{a}^*)$. However, the foreign firm is affected more adversely, because firms have a higher share of their national market. However the home firm has to incur a higher capital installation cost than the foreign firm to advance to $\Im''$. It is more likely that the home firm will advance at least one level of technology with $(\bar{a}, \underline{a}^*)$ than with $(\bar{a}, \bar{a}^*)$, because the foreign firm is less likely to invest with the former pair of realizations than with the latter. Without imposing additional structure, more cannot be said. Because the purpose of this chapter is to provide a credible argument as to how material injury can arise endogenously in a framework with cyclical reciprocal dumping, additional structure is not required.

Now that the period 2 production and investment equilibria have been considered, subgame perfection requires that period 1 be addressed. Because of my objective in establishing that endogenous injury can occur, the only period 1 equilibrium of interest

is that in which the demand realizations are $(\underline{a}, \bar{a}^*)$, the foreign firm invests in the newly available technology $\Im'$, and the home firm does not.[9]

3 Policy Implications

Now that it has been established that a technological disadvantage can arise through uncorrelated cyclical demand fluctuations in a framework that corresponds to the legal definition of dumping (exporting at a price below that charged in the national market coupled with material injury in the form of depressed profits), it is natural to inquire as to the role of AD law in alleviating the injury to the home firm. Since the home firm's profits are depressed because of its technical disadvantage, resulting in higher production costs than its foreign rival, it should be apparent that the home firm cannot restore a competitive balance (technical symmetry) without investing in the technology that the foreign firm possesses. Further, it has been established that once the home firm becomes a technological laggard, it may continue to fall further behind its rival. This is because the current period and continuation payoffs of the home firm are lower than that of the foreign firm, as the former has a higher cost of investment in the newest technology than the latter. Thus a firm can incur material injury even if it is not mismanaged or is not in a declining industry. The inefficiency that arises in this context is endogenous.

It should be apparent that AD policy will not be effective in restoring the so-called "level playing field" that supporters of AD law espouse, unless it induces the home firm to achieve technical symmetry with its rival.[10] That is, it must raise the sum of (6) and (7) for the home firm and/or reduce this sum for the foreign firm. The conventional application of AD duties to the foreign firm's sales of its product in the home market can do this. To see this, suppose an AD duty d^f is introduced to the foreign firm's gross profits function (4) for production for the home market. That is, rewrite (4) as

$$\tilde{\pi}_t^f = (a_t - b(q_t^h + q_t^f))q_t^f - (v_{\Im}^f + s + d^f)q_t^f. \tag{8}$$

Solutions of (4) for the home firm and (8) for the foreign firm, using (3) and a rederivation of (5) will disclose that the stage game (per period) equilibrium profits for the home firm will rise and those of the foreign firm will decline. That is, AD policy induces a shift in profits from the foreign to the home firm. Inspection of (5) discloses that d^f in (8) has the same incremental effect on q_t^h and q_t^f as does v_t^f and s. Hence $\partial q_t^h/\partial d^f = \frac{1}{3}b$ and $\partial q_t^f/\partial d^f = -\frac{2}{3}b$. Thus the negative effect of d^f on q_t^f is larger than its positive effect on q_t^h. Hence per period output in the home market will decline as a result of the AD duty. Therefore, by (3), the equilibrium price will rise, resulting in a decline in home consumer surplus. Since the markets are segmented, there is no impact on the foreign market equilibrium. The effect on the home market equilibrium will raise the current period and continuation payoffs for the home firm and lower them for the foreign firm. Thus the AD duty does provide an inducement to a technologically upgrading investment by the home firm, and makes the restoration of technical symmetry more likely. The AD duty does not, however, ensure that the investment by the home firm takes place. For example, as was discussed previously, demand realizations

such as $(\underline{a}, \underline{a}^*)$ and $(\underline{a}, \bar{a}^*)$ generate not investing as a dominant strategy for the home firm in the absence of AD duties. The presence of AD duties, particularly in the case of $(\underline{a}, \bar{a}^*)$, do not guarantee that the technology gap will be closed. This is because for the latter pair, it is likely that the home firm will have to advance two levels of technology and incur a higher capital installation cost than the foreign firm. This is because investing is a dominant strategy for the foreign firm in the absence of AD duties when its national market demand realization is high. The presence of home AD duties does guarantee that the home consumers will lose surplus as a result of a higher equilibrium price.

Because the source of material injury to the home firm is its technical disadvantage, the Principle of Targeting suggests that a more directly applied remedy, such as a conditional investment subsidy, would attain technical symmetry (a level playing field) at a lower social cost than that imposed by AD duties.[11] This is because a conditional subsidy would be disbursed only if the home firm invests in a new unit of capital. In contrast, consumers of this product absorb a loss in surplus irrespective of the home firm's investment decision under AD duties. Even if these consumers were taxed to finance the subsidy in the period in which it took place, this would be preferable to a loss in surplus due to the AD duty in every period until the investment took place. (That is, in a longer horizon game, or, in the present three-period model, if the investment did not take place.) Although this is not a permitted policy response under the AD Code of the World Trade Organization (WTO), it does suggest a sensible reform.

The above discussion also suggests a possible reform in the material injury decision in AD law. As was noted in the introduction, there are many characteristics of firm performance that an AD investigation is required to take into account. In the United States, at least, none of these variables is to be dispositive. My contention is that the investigation should be more focused. In particular, emphasis should be given to the currency of the home firm's technology, as well as that of its foreign competitors. Further, the above discussion and simple model suggests that investigators give attention to the rate of innovation in a petitioning industry, as well as to the degree to which demand is cyclical. Although the model does not address it formally, a natural extension would be to consider the rate of depreciation in an industry, as well.[12–14] Both the rate of innovation and the rate of depreciation suggest that the vintage of a firm's capital stock is important to an assessment of the viability of a firm that is confronting dumping. Sharpening the focus of the injury criteria in material injury investigations, as well as the serious injury investigations of the escape clause, can contribute to decisions that better reflect the ability of petitioners to compete. Although the failure to adopt new technology may reflect mismanagement, a clearer understanding of the manifestation of injury should make mismanagement more difficult to conceal.

4 Conclusion

I have demonstrated that cyclical demand in a finite horizon dynamic duopoly model in which firms may face an investment decision in stage 1 of each period and are Nash–Cournot competitors in stage 2 of each period can induce a permanent technological disadvantage. In addition to international shipping costs, this requires the coincidence

of low demand in a firm's primary market, high demand in its rival's primary market, and a cost reducing innovation. Because low demand results in low current period profits, the discounted flow of profits accruing to an investment in the cost reducing technology may fall below the acquisition costs, thereby inducing a firm to hold an option to invest. The option is held until demand uncertainty is resolved through the realization of demand in the subsequent period. If a firm's acquisition costs of a new technology depend upon the level of technology that it presently uses, holding the option to invest results in higher costs in the acquisition of the newest technology when further innovation takes place. If the rival firm has invested rather than held the option, a permanent cost advantage can arise on the part of the investing firm. This requires that the innovation occurring while the option is held raise the acquisition costs of the holding firm by an amount sufficient to induce it to not invest under plausible future demand realizations.

A policy implication is that investigative authorities in administered protection verdicts should take the age of a firm's and its rival's capital stock and technology, the degree to which demand is cyclical in its constituent firm's national and export markets, and the rate of innovation in the industry into account when making material and serious injury decisions. If one takes the view that the objective of administered protection law is to prevent injury in circumstances that may be deemed fair (safeguards) and unfair (dumping and subsidies), rather than to serve efficiency, then the issues raised in this chapter may be important. This is particularly the case when the inefficiency that arises here is a result of the unfortunate confluence of events, rather than mismanagement. This occurs in the context of cyclical demand and reciprocal dumping generated by an oligopolistic interaction, both of which are plausibly present in a majority of administered protection investigations. However, the requirements for endogenous injury in our framework are rather unlikely to manifest simultaneously. Thus, I both justify the use of an injury requirement, as well as suggest that investigative authorities are too lenient in their determination.[15]

Acknowledgments

I wish to thank the Australian National University, where this chapter evolved, for its supportive environment. Thanks are due to Rod Falvey, Jiandong Ju, Geoff Reed, seminar participants at Nottingham University, and participants at the Midwest International Economics Meetings for their invaluable comments.

Notes

1 The model, with some modification, is amenable to declining or growing industries.

2 The issue of mismanagement or self-infliction is nonetheless important in the provision of a public policy that is contingent upon an injury determination. The difficulty in making the distinction between injury and mismanagement in a model of an injury investigation is discussed in Hartigan (2002).

3 An additional complexity would be to make $r^i_{\Im'}$ an explicit function of the extent of the innovation. It is not necessary to do so to establish the primary points of this chapter.

4 Eaton and Lipsey (1981) invoked an investment model of entry deterrence in which an incumbent firm entailed a unit of capital.

5 Such industries are often termed cyclicals. Metal and chemical products typically do not have high R&D budgets relative to firm size. In any case, we wish to avoid the formidable additional complexity of introducing research and development to the model.

6 Nash–Cournot games can be interpreted as choices of capacity rather than output. However, games with capacity as the strategic variable may generate mixed strategy equilibria. See Staiger and Wolak (1992b). So as to avoid this complexity as much as possible, I have restricted investment in capacity to a discrete choice.

7 For further discussion of the option theory of investment, see Dixit and Pindyck (1994).

8 Grossman and Levinsohn (1989) devised a model of the effect of imports on the returns to capital in which the value of the firm is a function of current and past investment. Capital stock is indexed by the date at which the investment took place. They assess the impact of unanticipated changes in import prices on stock market returns. In this chapter, the concern is with the effects of (predictable) demand fluctuations on the decision to invest, and the implications of that decision upon the future viability of the firm.

9 As Guiso and Parigi (1999) discuss, the relationship between investment and uncertainty cannot be determined in a purely theoretical context. That is, models can be constructed in which uncertainty increases or decreases investment. However, using Italian manufacturing data, they disclose a negative effect of uncertainty on investment. In the present model, there is not only uncertainty, but also a low price realization in the period in which the investment must be made. Furthermore, there is the possibility of innovation embodied in the investment and the attendant possibility of technology differentials between the firms.

10 Lindsey and Ikenson (2002) discuss the framing of the primary political justification in the United States for AD law and policy in terms of creating a "level playing field" for US producers.

11 See Bhagwati et al. (1998).

12 Grossman (1986), Murray and Rousslang (1989), and Pindyck and Rotemberg (1987) have devised methods for disentangling the domestic causes of injury from that induced by imports. Kelly and Morkre (1998) discuss the use of a computable general equilibrium model to compare a petitioning industry's equilibrium with a simulated equilibrium in the absence of dumping. Hartigan (2000) provides an argument for weak injury criteria when small numbers of firms behave collusively in the absence of an AD law. In contrast, this chapter discusses the appropriate criteria for ascertaining injury.

13 Although exit may occur as a result of the endogenous injury depicted in this chapter, it is to be distinguished from predatory dumping. For examples of the latter, see Hartigan (1994, 1996).

14 Klepper and Simons (2000) demonstrate that technological change had a significant influence on firm survival in the evolution of the US tire industry.

15 In the United States, material injury determinations are in the affirmative in 70 percent of the AD petitions adjudicated by the US International Trade Commission.

References

Anis, Aslam H., Michael Benarroch, and Quan Wen 2002: Persistent protection in an international exit game. *Journal of International Economics*, 56, 465–87.

Bhagwati, Jagdish, Arvind Panagariya, and T.N. Srinivasan 1998: *Lectures on International Trade*, 2nd edn. Cambridge, MA: MIT Press.

Bernhardt, Dan 1984: Dumping, adjustment costs and uncertainty. *Journal of Economic Dynamics and Control*, 8, 349–70.

Brainard, S. Lael 1994: Last one out wins: trade policy in an international exit game. *International Economic Review*, 35, 151–72.

Brainard, S. Lael and Thierry Verdier 1994: Lobbying and adjustment in declining industries. *European Economic Review*, 38, 586–95.

Brainard, S. Lael and Thierry Verdier 1997: The political economy of declining industries: senescent industry collapse revisited. *Journal of International Economics*, 42, 221–37.

Brander, James and Paul Krugman 1983: A ‘reciprocal dumping’ model of international trade. *Journal of International Economics*, 15, 313–21.

Cassing, James H. and Arye L. Hillman 1986: Shifting comparative advantage and senescent industry collapse. *American Economic Review*, 76, 516–23.

Das, Satya 1992: Market uncertainties and cyclical dumping. *European Economic Review*, 36, 71–82.

Dixit, Avinash K. and Robert S. Pindyck 1994: *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.

Eaton, B. C. and R. G. Lipsey 1981: Exit barriers are entry barriers: the durability of capital as a barrier to entry. *Bell Journal of Economics*, 12, 721–9.

Ethier, Wilfred J. 1982: Dumping. *Journal of Political Economy*, 90, 487–506.

Grossman, Gene M. 1986: Imports as a cause of injury: the case of the U.S. steel industry. *Journal of International Economics*, 20, 201–23.

Grossman, Gene M. and James A. Levinsohn 1989: Import competition and the stock market return to capital. *American Economic Review*, 79, 1065–87.

Guiso, Luigi and Giuseppe Parigi 1999: Investment and demand uncertainty. *Quarterly Journal of Economics*, 114, 185–227.

Hartigan, James C. 1994: Dumping and signaling. *Journal of Economic Behavior and Organization*, 23, 69–81.

Hartigan, James C. 1996: Predatory dumping. *Canadian Journal of Economics*, 29, 228–39.

Hartigan, James C. 2000: An antidumping law can be procompetitive. *Pacific Economic Review*, 5, 5–14.

Hartigan, James C. 2002: Endogenous obfuscation in injury investigations. *Economica*, 69, 505–13.

Kelly, Kenneth H. and Morris E. Morkre 1998: Do unfairly traded imports injure domestic industries? *Review of International Economics*, 6, 321–32.

Klepper, Steven and Kenneth L. Simons 2000: The making of an oligopoly: firm survival and technological change in the evolution of the U.S. tire industry. *Journal of Political Economy*, 108, 728–60.

Lindsey, Brink and Dan Ikenson 2002: Reforming the antidumping agreement: a road map for WTO negotiations. Cato Institute, Trade Policy Analysis Papers, No. 21.

Long, Ngo Van and Neil Vousdan 1991: Protectionist responses and declining industries. *Journal of International Economics*, 30, 87–103.

Mandel, Michael J., Mike France, and Dan Carney 2000: The great antitrust debate. *Business Week* (June 26) 40–2.

Murray, Tracy and Donald J. Rousslang 1989: A method for estimating injury caused by unfair trade practices. *International Review of Law and Economics*, 9, 149–64.

Murray, Tracy and Nurlan Kurdaliev 1999: Universal dumping of homogenous products. *Review of International Economics*, 7, 580–9.

Pindyck, Robert S. and Julio J. Rotemberg 1987: Are imports to blame? Attributes of injury under the 1974 Trade Act. *Journal of Law and Economics*, 30, 101–22.

Staiger, Robert W. and Frank S. Wolak 1992a: The effect of domestic antidumping law in presence of foreign monopoly. *Journal of International Economics*, 32, 265–87.

Staiger, Robert W. and Frank S. Wolak 1992b: Collusive pricing with capacity constraints in the presence of demand uncertainty. *RAND Journal of Economics*, 23, 203–20.

17

International Trade in Services: More Than Meets the Eye

Lawrence J. White

Chapter Outline

This chapter presents an overview of international trade in services. Though often intangible and "invisible," trade in services accounted for almost a fifth of total world trade in 2000 and has been growing in both absolute and relative terms. The General Agreement on Trade in Services (GATS) – came into existence in 1995, as part of the World Trade Organization (WTO). GATS has had only limited success in achieving reductions in barriers to trade in services, though a negotiating round among member countries is currently under way. This slow progress is largely due to the extensive national regulation that frequently surrounds services and the entry (and trade) barriers that are often part of that regulation. The structure of the GATS document itself, as well as the WTO's post-Seattle defensiveness, reflects the political sensitivity of these national regulatory issues; but this sensitivity, unfortunately, has impeded progress in reducing trade barriers.

1 Introduction

The common image of international trade is that of goods – sacks of grain, rolls of steel, pallets of machinery – traveling across borders by boat, rail, or truck. Though this image does represent the dominant form of trade, it nevertheless leaves out another significant, and growing, form of trade: international trade in services.

Traditionally, services have generally been considered to be inherently non-tradable – or at least difficult to trade. There were acknowledged to be exceptions,

such as tourism or the transportation and other services that attached to the trade in goods. Still, trade in services seemed almost to be an oxymoron.

This is no longer the case. With the expansion of services as a proportion of national output for many economies, and especially for developed economies,[1] there has also come an expansion of international trade in services and an expanded recognition of and interest in that trade. Among the indications of that expanded interest and concern are the establishment within the World Trade Organization (WTO) of a General Agreement on Trade in Services (GATS) that roughly parallels the General Agreement on Tariffs and Trade (GATT), with active negotiations among the 144 WTO members and a Council Trade in Services to administer it; an annual report by the US International Trade Commission (USITC), *Recent Trends in US Services Trade*; and a growing pile of statistics, academic studies, and monographs on the topic.[2]

This chapter is about the growing importance of international trade in services. After highlighting the dissimilarities, as well as the similarities, between trade in services and trade in goods, we present recent data that highlight the absolute and relative growth of international trade in services and then provide an overview of the multilateral trade negotiations that have occurred with respect to services.

2 The Dissimilarities, and Similarities, between Trade in Goods and Trade in Services

2.1 The Dissimilarities

International trade in goods is easy to visualize. The goods are tangible. They physically move from one geographical location to another across a border – by ship, rail, or truck. So long as the geographic exit and entry points of the goods are well specified, the reporting, documentation, classification, taxation (if any), and data gathering that pertain to the shipment of the goods are relatively straightforward.[3]

Further, even the barriers to trade in goods are relatively easy to visualize and understand. A tariff is a tax on the import of a good. A quota is a quantitative limitation on the import of a good. Even the subtler regulatory restrictions on goods imports, such as customs delays at borders or "safety"-motivated restrictions, are easy to understand.

Services are different in at least two important respects. First, they are usually intangible; they cannot be seen, held, smelled, or touched; they also cannot be stored. They usually do not cross international borders the way goods do. Second, they are often regulated by governmental entities. Both of these differences make trade in services different from trade in goods and help explain why trade in services was considered relatively unimportant and why the liberalization of trade in services has proceeded much more slowly than has been true for trade in goods.

2.1.1 Intangibility

The intangibility of services means that their trade cannot occur in the same, physical-shipment way that trade in goods occurs. In an important sense, they are invisible.

Indeed, international services remittances are often described as part of the "invisibles" in the discussion of a country's balance of payments.

International trade in services can occur in one of four ways or "modes":[4]

(1) Cross-border: Some services do actually originate with a provider in one country and "move" across a border to a user-recipient in another country. Some examples include:

- the sale of electricity from one country to another;
- telephone communications across borders;
- catalog- or Internet-based retail transactions across borders;
- the transportation services that attach to the movement of goods in international trade;
- the provision of a loan, deposit, insurance, or other financial services product across a border;
- distance learning: a student located in one country taking a course offered electronically by an educational institution located in another country.

Even in these instances, the "movement" is invisible and instantaneous; the service does not stop and wait at the border while a customs official inspects, classifies, and records it and determines the appropriate tariff.

(2) Consumption abroad: Some services are consumed through the travel of the customer from one country to the provider in another country. Examples include

- tourism;
- study (education) abroad;
- medical treatment abroad.

Though a country's nationals who have traveled abroad may be stopped at the border upon their return and asked what *goods* they have purchased (and the appropriate tariffs levied), they are rarely (if ever) asked what services they have bought and consumed.

(3) Commercial presence: For many services, the most effective delivery mechanism is often the establishment of a physical presence (a branch office or subsidiary) in the country in which the service is to be sold and consumed. Underlying the desirability of the physical presence abroad are at least two phenomena: First, the individual units of the service are not sufficiently valuable so as to warrant the customers' traveling abroad for them (many people may be willing to travel 5 miles for a restaurant meal, but few are willing to travel 5,000 miles solely for a restaurant meal); and second, the transportation and scale economies of provision argue for bringing the physical establishment and its service into proximity with the customers (rather than having them all travel abroad to obtain the service). Further, a local presence will be more desirable when local knowledge of the customers is especially important for the seller (e.g., banks want to have a considerable amount of information about the creditworthiness of their borrowers); when the customers want firsthand contact with or convenience related to the sellers (e.g., depositors want to be able to have ready access to their

funds and/or to interact with tellers; retail merchandise customers want to be able to see the items that they are buying, perhaps sample them or try them on); and/or when continuing physical contact between provider and customer is necessary (e.g., for some professional services, short visits supplemented by telephonic contact may not be sufficient).

Consequently, commercial presence abroad is a common phenomenon for (among others):

- banks and other financial services providers;
- restaurant chains;
- hotel chains;
- retail merchandise chains;
- accountancy branches;
- law offices branches.

Commercial presence is, of course, also important for many categories of goods, where branch/affiliate factories are alternatives to direct shipments. But for goods, direct shipments from the home country are often a fairly close substitute for production abroad. By contrast, direct "shipment" is likely to be an exceedingly poor substitute for the commercial presence mode of delivering many services.[5]

(4) Temporary presence (presence of natural persons): A service may be delivered through the temporary presence abroad of an individual or group of individuals who deliver the service while abroad. Examples include:

- visiting entertainers (e.g., orchestras, theater companies, rock stars);
- temporary consultancies;
- short-term construction projects.

Some services may be capable of being delivered through more than one mode.

Since the commercial presence method is a common form of delivery, it is worth additional attention. Commercial presence abroad will mean the necessity for investment (foreign direct investment, or FDI) in the host country from the provider in the home country. Also, personnel from the provider enterprise headquarters will have to travel to the location abroad to deliver services, hire personnel, and supervise the personnel. This necessity for investment in and travel to the host country in order to deliver the services effectively means that the delivery process is more complex than is true for the simple shipment of goods across a border. In turn, the complexity of delivery provides greater and more subtle opportunities for host country governments to impede the delivery of those services from abroad (and thus favor domestic providers). Restrictions on inbound FDI (including restrictions on ownership structures and arrangements), restrictions on immigration, and restrictions on commercial location and establishment can seriously impede international trade in services.

For example, White's (2001) examination of international trade in accounting services found the following types of restrictions present in one or more countries:

- nationality requirements with respect to who can offer local accounting services;

- residence or establishment requirements;
- restrictions on the international mobility of accounting personnel;
- restrictions as to the use of the brand names of firms or requirements that only local names be used;
- restrictions on advertising or other promotional efforts; restrictions on price competition;
- quantitative limits on the provision of services;
- restrictions on the services that accounting firms can and cannot provide;
- restrictions on who can be an owner of an accounting firm; for example, requirements that all or a specified number or fraction of the owners of an accounting firm be local citizens; be residents; be active in the business of the firm; be locally licensed; be members of an approved professional organization;
- restrictions as to the legal form or structure that an accounting firm must have (e.g., prohibitions on a corporate form);
- discriminatory arrangements with respect to the licensing of foreign accountants, including applications, testing, assessments of educational qualifications, relevant experience;
- differential taxation treatment;
- restrictions on international payments for services;
- restrictions on cross-border flows of information;
- inadequate protections for the intellectual property related to accounting services, such as computer software;
- “buy national” practices of governments with respect to their purchases of accounting services.

These types of regulatory restrictions on trade are found generally across the services sector.[6]

This last discussion provides a convenient segue to the second distinctive feature of services.

2.1.2 REGULATION

Services are often subject to extensive regulation by governments. As an example, consider the list of specific industries in the United States that were usually described as “heavily regulated” in the 1960s and 1970s and that were the target of substantial deregulation or regulatory reform efforts in the United States during the last quarter of the twentieth century (Joskow and Rose, 1989; Noll, 1989; Winston, 1993; Joskow and Noll, 1994):

- air transport;
- rail transport;
- truck transport;
- water transport;
- natural gas pipeline transport;
- petroleum pipeline transport;
- telephone;

- broadcasting;
- electricity;
- banking;
- securities;
- insurance.

All of these industries are services. Also, though the list was drawn from the US experience, similar lists of industries in the 1960s and 1970s in other countries were either heavily regulated or under government ownership and have subsequently been subject to deregulation or privatization.

Another broad category of services – professional services, such as those provided by doctors and other health care providers, lawyers, accountants, architects – are also subject to direct or indirect government regulation. Sometimes a governmental body directly regulates the profession; sometimes that regulation is delegated to a professional association, with the government retaining oversight.

This regulation adds yet another layer of complexity to international trade in services and to efforts to liberalize this trade. Regardless of form, regulation is almost always justified as governmental activity that is being pursued so as to advance "the public interest." Ideally, this pursuit of "the public interest" would involve government efforts to address market failures: specifically, the presence of significant market power (addressed through "economic" regulation, such as price limitations); the presence of significant externalities or spillover effects (addressed through health-safety-environment regulation, involving changes in production processes or product/service design); and/or the presence of significant asymmetric information problems (addressed through information regulation, involving required information attached to goods or services).[7] In addition, it is clear that governments often use regulation for outright "consumer protection," exercising paternalistic actions, and also use it in attempts to redistribute income.

Even good regulatory intentions can go awry, however, and regulatory bodies and procedures can be "captured" by groups – often producers – that care strongly about favorable outcomes for themselves; or the regulatory process may have begun as a regulatory capture effort. With capture usually comes protection: regulatory barriers to entry that favor incumbents. Among the entrants that are shut out are potential entrants from abroad. And thus regulation creates and maintains barriers to international trade in services.

The regulatory structure and procedures themselves add to the complexity of liberalizing trade. The regulatory process is imbued with a "protect the public interest" or outright "consumer protection" ethos. Incumbents can argue to government regulatory officials that it is the incumbents that are the embodiment of that public interest and that entry – especially entry from abroad, by "foreign" entities that (arguably) may care less about the domestic public interest and that are inherently harder to regulate – would undermine or weaken the regulation and would thus be antithetical to the public interest. And there are usually rules and laws that must be rescinded or modified to permit entry, giving incumbents ample opportunity to make their case and delay or impede entry. Again, international trade in services suffers.[8]

We return to these issues in Section 4 when we discuss the trade negotiations that have occurred in the services area.

2.2 The Similarities

Though, as the previous section argued, trade in goods and trade in services exhibit substantial dissimilarities, they also exhibit a fundamental similarity: The opening, and expansion, of international trade in services can bring the same type of economic gains to countries – an improved allocation of resources and higher living standards – as is brought about by trade in goods. Trade allows a country to focus its resources on what it does relatively well, importing the goods and services that it does less well and paying for those imports with the export of the goods and services that it does especially well. At its heart, the argument for the benefits of trade, whether for services or goods, is fundamentally an argument for the benefits of competition, applied in an international framework.[9]

This standard advantages-of-trade argument is based on reallocations of resources with given or static technologies. An additional "kick" to expanded output may be provided by the improved technologies and new ideas that accompany the expanded goods and services from abroad. The commercial presence/physical establishment that is important for many services may be an especially fortuitous vehicle for the introduction and transmission of these new ideas.

Further, just as expanded trade in goods that are important inputs into the production of other goods can have a multiplicative effect in reducing "downstream" distortions, improving resource allocations, and encouraging further trade flows, so too can expanded trade in crucial input services have multiplicative effects. As Deardorff (2000) points out, services such as freight transportation, telecommunications, finance, and law are all important inputs into the flow of goods in international trade. Greater efficiency in the provision of these services, brought about by expanded trade in these services, should encourage expanded goods trade. Also, financial services, accounting services, and legal services are all essential inputs into commercial transactions in any economy. Again, greater efficiency in the provision of these services, brought about by expanded trade, should encourage greater efficiency throughout an economy.

Over the past decade an important empirical literature has supported these latter notions.[10] The evidence clearly shows that greater development of a country's financial sector and of the laws and rules that surround financial and commercial transactions have significant positive influences on that country's rate of economic growth. Further, greater involvement of foreign financial firms in a country's financial services sector tends to encourage greater development of that sector and thus encourage more rapid economic growth (Levine, 1996; Claessens et al., 2000; Claessens and Jansen, 2000). Also, as White (2001) suggests, greater involvement of the major (worldwide) accounting firms in Asian economies in the 1990s could well have led to more rigorous accounting standards in those countries and an amelioration of the financial crises that beset the region in the late 1990s.

In sum, trade is generally beneficial; and, in that respect, the distinct characteristics of trade in services may well reinforce this general message.

3 The Expansion of Trade in Services: Predictions and the Data

In the Introduction we mentioned the growing importance of international trade in services. In this section we document that claim.

3.1 Why Would We Expect Trade in Services to be Growing?

There are at least four reasons why we would expect to see a rising trend in international trade in services – rising in absolute terms, rising relative to world GDP, and rising relative to trade in goods.

First, services appear generally to have an income elasticity of demand that is substantially greater than one, especially in developed economies (Kang, 2000b). As countries grow, they tend to produce and consume more-than-proportionately greater levels of services. Accordingly, to the extent that any services can cross borders, we would expect generally to see the growth of these cross-border services as well, and they would be documented as a growth in international trade in services.

Second, the technologies that underlie important categories of services have experienced rapid improvements in the past few decades. Included in this list are telecommunications engineering, data processing, transportation engineering, and biological sciences. The consequences for services have been profound. Better quality services, with greater variations and qualities, can be offered by more firms at lower costs over greater distances, generating more effective competition among providers. The real (quality-adjusted and inflation-adjusted) prices of many services have decreased.[11] The heightened quality/variety with lower prices has reinforced the income-elasticity effects to encourage greater consumption of services. And, again, some of this should spill over to international trade in services.

Third, the technological improvements in telecommunications, data processing, and transportation have had a direct beneficial effect on those services that are delivered internationally through the cross-border mode and an indirect beneficial effect on other services that are delivered through the other three modes. Further, because these improved technologies allow senior managers to manage better multiple-establishment enterprises that are located at geographically separated locations, they directly benefit the delivery of services that occur through the commercial presence mode.

Finally, to the extent that governments have eased their barriers to trade in services, that too would encourage greater trade.

In sum, there are strong reasons to predict a rise in the importance of international trade in services. We now turn to the data.

3.2 The Data and What They Show

There are three categories of data on international trade in services that is presented: (1) world data, including major regions; (2) US data; and (3) US data on sales of services from overseas establishments.

3.2.1 World trade data

World data on trade for 2000 can be found in WTO (2001a) and at the WTO website http://www.wto.org. The services data cover all "commercial services;" that is, all services except noncommercial services provided by governments (but commercial services provided by, say, a government airline would be included). However, the level of detail and extent of coverage for services are far less than are provided for goods; this statistical neglect is an indication of the general neglect – until recently – of attention to services: in national income accounts and balance of payments statistics, as well as in the multilateral negotiation arena.

The data for trade in services cover only cross-border, consumption abroad, and temporary-presence forms of trade. The commercial-presence form is not covered.

Table 17.1 provides data for trade in goods and in services in 2000, as well as average annual percentage increases since 1980. As can be seen, services trade constituted somewhat less than a fifth of total world trade in 2000. Since 1980, trade in services has expanded appreciably faster than trade in goods; but most of that relative expansion took place during the 1980s; since 1990 trade in services has grown only slightly faster than trade in goods. To this author's knowledge the reasons for this slowdown in the relative growth of trade in services in the 1990s have not been explored or explained.[12] However, it is noteworthy that the decade of rapid growth in services trade, the 1980s, was also the decade that saw the Uruguay Round propel services trade to a prominent place in the multilateral negotiating arena.

Both trade series are value-based measures. For trade in goods, a volume-based measure (which controls for price changes) is available as well, but no such measure is available for trade in services. Table 17.1 shows the rates of growth of real (constant-price) world GDP and the volume of goods trade for these same decades. The (real) volume of goods traded grew appreciably faster than real world GDP, especially for the 1990s. Although it seems likely that real services trade also exceeded real GDP growth, that claim must (unfortunately) remain a supposition until better data are available.

Table 17.2 presents the trade data classified by major regions.[13] Trade in services constituted over a fifth of total exports in 2000 for North America and for Western Europe and over a fifth of total imports for Western Europe and Africa. The worldwide slowdown in the relative growth of services trade in the 1990s was clearly due to the slower growth rates for services exports and imports in the 1990s for North America,

Table 17.1 Total World Trade, 1980–2000

	Year	**Average annual % change**		
	2000	**1980–2000**	**1980–90**	**1990–2000**
Value of goods ($B)	6,186	5.73	5.26	6.20
Value of services ($B)	1,435	7.08	7.92	6.25
Services as a % of total trade	18.83			
Real GDP		2.73	3.20	2.26
Volume of goods		5.44	3.93	6.96

Source: *WTO (2001a), and data tables found at http://www.wto.org.*

Table 17.2 World trade by regions, 1980–2000

	Year	Average annual % change		
	2000 (%)	1980–2000	1980–90	1990–2000
North America[a]				
Goods exports (val.)		6.62	5.92	7.33
Services exports (val.)		10.14	12.83	7.51
Services as a % of total exports	22.8			
Goods imports (val.)		8.04	7.19	8.89
Services imports (val.)		9.53	12.39	6.74
Services as a % of total imports	14.1			
Latin America & Caribbean				
Goods exports (val.)		6.10	2.95	9.35
Services exports (val.)		6.42	5.43	7.41
Services as a % of total exports	14.0			
Goods imports (val.)		5.90	0.55	11.54
Services imports (val.)		4.71	1.99	7.51
Services as a % of total imports	15.9			
Western Europe				
Goods exports (val.)		5.64	7.23	4.08
Services exports (val.)		5.72	6.93	4.53
Services as a % of total exports	21.2			
Goods imports (val.)		5.23	6.27	4.20
Services imports (val.)		6.07	7.56	4.61
Services as a % of total imports	20.6			
Africa				
Goods exports (val.)		0.91	−1.48	3.35
Services exports (val.)		4.48	4.00	4.96
Services as a % of total exports	18.5			
Goods imports (val.)		1.75	−0.20	3.74
Services imports (val.)		1.67	−0.26	3.64
Services as a % of total imports	23.2			
Asia				
Goods exports (val.)		9.04	9.37	8.72
Services exports (val.)		9.41	10.13	8.70
Services as a % of total exports	14.3			
Goods imports (val.)		8.08	8.41	8.10
Services imports (val.)		8.62	9.85	7.40
Services as a % of total imports	18.7			

Source: *WTO (2001a), and data found at http://www.wto.org.*
[a] *For the United States and Canada only.*

Western Europe, and Asia. The slower growth of services trade in North America in the 1990s is especially striking.

Table 17.3 shows the 10 leading country exporters and importers for goods and for services in 2000. The United States was the leading trader in all four categories;

Table 17.3 Ten leading country exporters and importers of goods and services, 2000

Goods exports	Value ($B)	Services exports	Value ($B)	Goods imports	Value ($B)	Services imports	Value ($B)
1 US	781.1	US	274.6	US	1,257.6	US	198.9
2 Germany	551.5	UK	99.9	Germany	502.8	Germany	132.3
3 Japan	479.2	France	81.2	Japan	379.5	Japan	115.7
4 France	298.1	Germany	80.0	UK	337.0	UK	82.1
5 UK	284.1	Japan	68.3	France	305.4	France	61.5
6 Canada	276.6	Italy	56.7	Canada	244.8	Italy	55.7
7 China	249.3	Spain	53.0	Italy	236.5	Netherlands	51.1
8 Italy	237.8	Netherlands	52.3	China	225.1	Canada	41.9
9 Netherlands	212.5	Hong Kong	42.1	Hong Kong	214.2	Belg.–Lux.	38.3
10 Hong Kong	202.4	Belg.–Lux.	42.0	Netherlands	198.0	China	35.9
Leading 10 as a % of world total	56.14		59.24		58.49		56.68

Source: *WTO (2001a).*

US services exports accounted for almost a fifth (19.16 percent) of the world trade in services, and US goods imports accounted for over a fifth (20.34 percent) of the world's trade in services. The US share in both categories was more than double the share of the runner-up country. The aggregated share of the leading ten traders for each of the four categories of trade was quite similar. And the leading trader countries tended to be dominant in all four categories: Only twelve countries were needed to account for the forty (top 10×4) leading positions in the four categories, and seven countries (the United States, the United Kingdom, Japan, Germany, France, Italy, and the Netherlands) appeared on all four lists.

Table 17.4 provides the three major categories of trade in services; no finer classifications are available on a worldwide basis. Transportation services (passenger and freight transportation and seaport/airport charges) accounted for 23 percent of world trade in services; travel services (tourism) accounted for 32 percent of the world total; and other services (everything else) accounted for 45 percent of the world total. The United States was the leading exporter and importer in all three categories. The aggregated share of the leading 10 trading countries in each category showed substantially greater variation than was true for Table 17.3. But, as was true for the broad goods and services trade categories, the leading countries in the specific services trade categories tended again to be dominant in all categories: only 16 countries were needed to account for the 60 (top 10×6) leading positions in the 6 categories, and 4 countries (the United States, the United Kingdom, Germany, and France) appeared on all six lists.

3.2.2 US SERVICES TRADE DATA

Substantially finer, detailed data are available for US trade in services, as is shown in Table 17.5.[14] In aggregate, the United States was a substantial net exporter of services in 2000; this net export surplus position has been true since at least 1970

Table 17.4 Major categories of trade in services and ten leading country exporters/importers, 2000

Exporters	Value ($B)	Importers	Value ($B)
A. Transportation services; world total = $330B			
1 US	51.2	1 US	64.6
2 Japan	25.6	2 Japan	35.1
3 Netherlands	20.8	3 Germany	25.0
4 Germany	19.4	4 UK	22.0
5 France	19.4	5 France	19.5
6 UK	18.2	6 Netherlands	14.6
7 Hong Kong	13.7	7 Italy	13.7
8 S. Korea	12.8	8 S. Korea	11.0
9 Denmark	11.0	9 China	10.4
10 Belg.–Lux.	10.4	10 Denmark	9.4
Leading 10 as a % of world total	61.36	Leading 10 as a % of world total	68.27
B. Travel services; world total = $465B			
1 US	100.5	1 US	67.3
2 Spain	30.9	2 Germany	47.1
3 France	30.9	3 UK	38.0
4 Italy	27.4	4 Japan	31.9
5 UK	21.7	5 France	17.9
6 Germany	17.6	6 Italy	15.4
7 China	16.2	7 China	13.1
8 Canada	10.6	8 Hong Kong	12.5
9 Austria	10.0	9 Netherlands	12.2
10 Australia	8.4	10 Canada	12.1
Leading 10 as a % of world total	58.97	Leading 10 as a % of world total	57.53
C. Other services; world total = $640B			
1 US	122.9	1 US	67.0
2 UK	60.0	2 Germany	60.2
3 Germany	43.0	3 Japan	48.7
4 Japan	39.3	4 Italy	26.6
5 France	30.8	5 Netherlands	24.3
6 Netherlands	24.4	6 France	24.1
7 Belg.–Lux.	24.2	7 Ireland	23.5
8 Hong Kong	20.9	8 UK	22.1
9 Italy	20.4	9 Canada	20.4
10 Canada	19.4	10 Belg.–Lux.	19.8
Leading 10 as a % of world total	63.33	Leading 10 as a % of world total	52.61

Source: *WTO (2001a).*

Table 17.5 US trade in services, 1992–2002

Category	Exports		Imports	
	Value in 2002 ($B)	Avg. ann. % ch. 1992–2002	Value in 2002 ($B)	Avg. ann. % ch. 1992–2002
Travel	66.5	1.97	58.0	4.18
Passenger fares	17.0	0.03	20.0	6.54
Other transportation	29.2	3.08	38.5	4.95
Royalties and license fees	44.1	7.79	19.3	14.07
Services to affiliates	43.5	9.97	32.4	12.88
Education	12.8	7.51	2.5	12.39
Financial services	15.9	14.67	3.7	14.03
Insurance, net	2.8	17.05	15.3	18.19
Telecommunications	4.1	3.67	4.2	−3.63
Computer and data processing ser.	3.0	14.49	1.1	31.00
Constr., eng., arch., and mining ser.	2.9	4.26	0.8	12.01
Data base and other info. ser.	2.4	14.24	0.2	12.61
Industrial engineering services	0.7	13.45	0.2	5.15
Install., maint., and repair of equip.	5.0	6.17	0.8	15.57
Legal services	3.3	9.19	0.8	9.46
Mgt. consult. and p.r. services	1.7	8.83	1.2	17.20
Medical services	1.9	10.04	n.a.	n.a.
Misc. disbursements	0.6	20.44	1.5	14.44
Operational leasing	3.6	15.39	0.2	−5.57
R&D and testing services	1.1	5.92	1.0	16.54
Other services[a]	17.4	6.93	3.5	11.30
Total	279.5	5.50	205.2	7.24

Source: *Borga and Mann (2003).* Courtesy of the Bureau of Economic Analysis
[a] *Includes accounting, auditing, bookkeeping, advertising, mailing, reproduction, commercial art, personnel supply, sports and performing arts, training, film and tape rentals, etc.*

(Landefeld and Young, 1988; WTO, 2001a). The comparative advantage of the United States varied within the services sector, however, and it was a significant net importer of some services: passenger and other (freight) transportation, insurance, telecommunications, and miscellaneous disbursements.

3.2.3 US AFFILIATES DATA

As was discussed above, services are often best provided through a commercial (physical) presence of a branch, subsidiary, or affiliate. The data presented thus far for the world and for the United States do not capture such services sales. They are available, however, for the United States.[15] Table 17.6 provides these data for 2001;[16] Table 17.7 provides the annual aggregates for 1986–2001.

Table 17.6 Sales of services through nonbank majority-owned affiliates, 2001

Industry category of affiliate	Sales by US-owned affiliates abroad ($B)	Sales by foreign-owned affiliates in the US ($B)
Manufacturing	9.8	30.6
Wholesale trade	21.3	9.9
Retail trade	0.6	0.2
Publishing industries	15.3	14.2
Motion picture and sound recording industries	9.8	6.6
Broadcasting and telecommunications	23.2	19.6
Information and data processing services	16.7	7.5
Insurance	65.4	85.1
Real estate and rental and leasing	12.1	18.5
Professional, scientific, and technical services	66.9	37.4
Mining	9.7	5.9
Utilities	74.6	28.6
Transportation and warehousing	20.1	23.9
Employment services	10.0	9.3
Food services and drinking places	12.9	10.9
Educational services	1.2	0.3
Other industries	62.6	58.4
Total	432.2	366.9

Source: *Borga and Mann (2003).* Courtesy of the Bureau of Economic Analysis

The data in these tables indicate that the United States was a net exporter of services provided through these sales in 2001 and has been a net exporter in every year since 1987, the first year for which complete data are available.[17] The strongest comparative advantage of the United States has been in the following industrial categories: wholesale trade; publishing industries; motion picture and sound recording; broadcasting and telecommunications; information and data processing; professional, scientific, and technical services; mining; utilities; food services and drinking places; and other industries (which include non-depository financial institutions).

3.2.4 A SUMMING UP

The data show that services constitute a substantial fraction of the world trade total and that this fraction has been growing; this relative growth was especially robust in the 1980s. The United States has been and continues to be a net exporter of services, measured both as direct flows and as sales through affiliates. It is not surprising that the United States has generally been an ardent advocate of reduced barriers to international trade in services.[18]

Table 17.7 Aggregate sales of services through nonbank majority-owned affiliates, 1986–2001

Year	Sales by US-owned affiliates abroad ($B)	Sales by foreign-owned affiliates in the US ($B)
1986	60.5	n.a.
1987	72.3	62.6
1988	83.8	73.2
1989	99.2	94.2
1990	121.3	109.2
1991	131.6	119.5
1992	140.6	128.0
1993	142.6	134.7
1994	159.1	145.4
1995	190.1	149.7
1996	223.2	168.4
1997	255.3	223.1[a]
1998	286.1	245.5
1999	353.2[a]	289.3
2000	413.5	344.4
2001	432.2	366.9

Source: *Borga and Mann (2003).* Courtesy of the Bureau of Economic Analysis

[a] *Major redefinition of sales of services through affiliates.*

4 Multilateral Trade Negotiations with Respect to International Trade in Services

4.1 Background

Through the mid-1970s there was no widespread perception that trade in services was a significant enough phenomenon to warrant multilateral attention that was comparable to the attention that had been focused on trade in goods since the 1940s through the GATT.[19] Beginning in the late 1970s, however, the United States began actively pursuing bilateral "open-skies" agreements to reduce the impediments that US airlines faced abroad; and the United States also was increasingly concerned about how other countries were treating the overseas branches of US banks, securities firms, and a few other services providers. Further, the European Common Market was recognizing that the flow of services among its member countries was a major task to tackle. During the early 1980s the United States began a more concerted diplomatic effort to bring trade in services into the realm of multinational negotiation and bargaining.

During the Uruguay Round of GATT negotiations in the late 1980s and early 1990s, there was a more general international realization that trade in services needed to be brought into the formal multilateral negotiating arena.[20] The eventual result was the creation of the WTO as a successor to the organization that had administered GATT, the negotiation of a general agreement on services – GATS – and the parallel

placement of GATT and GATS within the WTO.[21] GATS came into effect on January 1, 1995.

4.2 GATS[22]

GATS is a broad and expansive document that applies to all of the 144 (as of early 2002) country members of the WTO, and to virtually all services. The two exceptions are government services that are of a noncommercial nature (e.g., social insurance, public safety, national defense) and air traffic rights (which are reserved for separate bilateral negotiations).[23] It is a broad set of principles that has the goal of encouraging freer trade in services. GATS establishes general commitments toward most-favored-nation (MFN) treatment, transparency in domestic regulation of services, market access, and national treatment of services suppliers. It encompasses the commercial-presence mode of delivery, as well as the cross-border, consumption-abroad, and temporary-presence modes. A Council for Trade in Services, composed of all WTO (GATS) members, administers GATS.

Unfortunately, as is discussed below, GATS is riddled with loopholes and exceptions – a formal recognition of the sensitivities and difficulties that come with services in general and the regulatory processes that surround many services. The most important loophole is that a country can choose which of its services sectors, through which mode, will be subject to which of the specific provisions of GATS. It does so through a specific list of "bindings" (commitments) that apply only to specifically named sectors and modes on a schedule of commitments. It can list as many or as few sectors as it chooses, with as many or as few adherences to the GATS provisions for each sector as it wishes.[24]

GATS is encompassed within a set of 29 articles. The following are the major articles and their important features:

1. Article I describes the four major modes of trade in service that were discussed above: cross-border; consumption abroad; commercial presence; and temporary presence (presence of natural persons).
2. Article II establishes a general obligation for MFN treatment. *But* a country can specifically choose to exempt any sector from MFN treatment, and Article V specifically exempts members of formal multi-country economic-integration areas (such as the European Union, the North American Free Trade Agreement, or Mercosur) from MFN obligations. Also, Article VII (which encourages mutual recognition of qualifications) does not require MFN treatment.
3. Article III establishes a general obligation for transparency with respect to laws, regulations, administrative guidelines, and similar measures that are relevant to trade in services. The transparency should include opportunities for inquiry and notification of significant change.
4. Article V permits countries to participate in regional economic-integration arrangements.
5. Article VI applies to domestic regulation and requires members to ensure that qualification requirements and procedures, technical standards, and licensing

requirements do not constitute unnecessary barriers to trade. Domestic regulatory requirements should be based on objective and transparent criteria, such as competence and the ability to provide the service; should not be more burdensome than is necessary to ensure the quality of the service; and, with respect to licensing procedures, should not in themselves be a restriction on the supply of the service. Members must provide an objective and impartial means of reviewing and providing appropriate remedies for administrative decisions that affect trade in services. In addition, when a country has made a binding with respect to a service, measures that affect trade in that service should be administered in a reasonable, objective, and impartial fashion; also, member countries should provide adequate procedures for the mutual recognition of the competence of the providers of other member countries.

6. Article VII encourages mutual recognition of the qualifications of services providers. Member recognition should not constitute a means of discrimination among countries in standards or criteria. Multilateral processes are encouraged.
7. Article VIII addresses domestic services provided by monopolies and restricts the ability of the monopoly providers to interfere with international trade in ways that are inconsistent with the other provisions of GATS.
8. Article XIII exempts government procurement from the general GATS obligation, while requiring further negotiations that began in 1997.[25]
9. Article XIV permits exceptions where important national goals, such as public safety and national security, are involved.
10. Article XVI applies to market access by countries that have made a binding with respect to a service. The article prohibits specific limitations (such as quotas) on the number of suppliers, the total value of service transactions or assets, the total number of service operations or quantity of service output, the total number of people employed, the types of legal entity or joint venture through which the service is supplied, and the participation of foreign capital.
11. Article XVII applies to market access by countries that have made a binding with respect to a service. The article requires "national treatment" of service suppliers: Countries should accord to foreign services and service suppliers treatment that is no less favorable than is applied to domestic services and suppliers. Specific treatment can be different, so long as the resultant conditions of competition do not favor domestic services or service suppliers.

4.3 An Assessment

Despite almost 20 years of attention, actual efforts at liberalization of trade in services are still in their infancy, and actual results have thus far been modest.

The signing of GATS by the members of the WTO brought it into effect in 1995. With the establishment of GATS came: (1) the submission of schedules by member countries as to the specific services sectors that they were prepared to commit (bind) to GATS-based liberalization rules; (2) further negotiations with respect to more specific guidelines for telecommunications, financial services, marine transportation, and movement of persons; (3) the development of detailed rules (disciplines) for accounting

services; and (4) a commitment to begin specific negotiations for widespread services liberalization in 2000. We address each in turn.

4.3.1 SUBMISSION OF SCHEDULES

The submission of schedules by member countries in 1995 did not yield much that was new or newly liberalized. Most countries simply bound themselves to the status quo with respect to the state of their services sector; some countries even bound themselves to less openness than their status quo, so as to give themselves more flexibility for subsequent negotiations and bargaining. Developed countries' bindings tended toward greater degrees of openness; developing countries' bindings tended toward less openness.[26]

Since any binding became a GATS/WTO commitment, which would give a country's trading partners the right of retaliation in the event that the country subsequently tightened its restrictions, the bindings were largely a "standstill" agreement that froze the status quo (or a little less) and discouraged backsliding. In that sense the bindings represented some "progress."

4.3.2 SECTOR NEGOTIATIONS

Sector negotiations for telecommunications and for financial services concluded in 1997, yielding further specific agreements and more openness-oriented bindings in these two areas, which became annexes to GATS.[27] Negotiations for maritime transport services and movements of natural persons also concluded in 1997, but with no agreement.

4.3.3 THE ACCOUNTING DISCIPLINES

Accounting services was selected by the Council for Trade in Services as a lead sector among professional services (which also encompass doctors, lawyers, architects, etc.) for developing greater specificity for the general principles of GATS (WTO, 1998; Ascher, 1999; White, 2001). This selection was not by chance. Accounting was already a field with a substantial level of international activity. All of the major accounting firms had had well-established international branch offices and affiliates. But it was (and continues to be) also a field with substantial domestic regulation, in virtually every country, and major impediments to trade (Colecchia, 2000; White, 2001), as the examples described in Section 2 indicated.

In May 1997 the Council adopted nonbinding guidelines for mutual recognition agreements (MRAs), which amplify the framework of Article VII of GATS. And in December 1998 the Council adopted disciplines (rules) on the domestic regulation of accountancy that amplify the provisions of Article IV of GATS, so as to provide greater detail as to transparency, licensing requirements and procedures, qualification requirements and procedures, and technical standards. The disciplines will not come into force, however, until the *conclusion* of the round of services negotiations that began in 2000.

4.3.4 THE GATS 2000 NEGOTIATING ROUND

As part of GATS, member countries agreed to start a round of detailed negotiations with respect to services in 2000. Those negotiations began in 2000, and, as of early 2002, are continuing. A lengthy term can be expected. After all, the Tokyo Round of GATT negotiations lasted from 1973 to 1979, and the Uruguay Round from 1986 to 1994. The issues that are to be negotiated in the "GATS 2000 Round" will be at least as complex as those that occupied the earlier trade negotiation rounds.

There are at least four additional reasons to expect lengthy and difficult negotiations. First, as Hoekman and Messerlin (2000) point out, any detailed negotiations concerning trade in services inevitably will be more complicated than those that apply to trade in goods. Country negotiators invariably begin their "request-offer" trade negotiations from a mercantilist perspective: "My country wants your country to lower your trade barriers on the following items; in return (and only in return) we would be willing to lower our barriers on the same or other items." For trade in goods, the barriers are often in the form of tariffs, so that the bargaining as to reciprocal reductions can occur in terms of readily comparable numerical reductions. Even when the barriers are in the form of quotas, a tariff-equivalent measure can usually be computed.

As was described above, however, the barriers to international trade in services are rarely in the form of tariffs or simple quotas. Instead, the regulatory barriers for services are far more complex and subtle, requiring extensive efforts to distill the true protective nature of the barriers, as well as the gains that could be achieved through their removal.[28] The determination of what constitutes an acceptable *quid pro quo* in such circumstances, especially when the barriers surrounding different services are being requested and offered, will surely create an additional layer of bargaining frictions and delay.

Second, for federal countries such as the United States, some of the regulatory impediments (e.g., banking, insurance, professional services) occur at the state level, yet it is the national negotiators who are engaged in the bargaining process. Again, this must add to bargaining frictions and delay.

Third, recall that the GATS schedules of country bindings consist of "positive lists" of those services sectors that a country has chosen to bind in some fashion to the GATS process. Though the bound sectors are thereby transparent and open to further negotiations with respect to any remaining barriers, the unbound sectors are simply unlisted and therefore opaque. This lack of transparency as to the unbound sectors will surely increase the difficulties of negotiations, since negotiators will have a less explicit knowledge base for their bargaining requests with respect to these opaque sectors.

Fourth, with only services "on the table" for negotiations, the room for maneuvering and bargaining will be restricted, since four-fifths of trade flows – that is, goods – will not be explicitly available for the reciprocal counteroffers that will be part of the bargaining process.[29] Again, this will surely impede the negotiations.

4.3.5 A SUMMING UP

The negotiation process for trade in services has been, and will continue to be, tentative and slow. One indicator of this slow progress has been the *absence* of any major

disputes in services trade that would require WTO adjudication. Countries are not yet seriously confronting each others' policies and impediments. The slow progress is probably the unavoidable reflection and consequence of the complexity and subtlety of services and their regulation and the national sensitivities that accompany this regulation, as is indicated by the "positive list" approach, the widespread exemptions, and the like.

Unfortunately, the disruptions at the WTO "ministerial" meeting in Seattle in late 1999 have surely slowed the process yet further. The disruptions have thrown the WTO on the defensive and have made it more tolerant of – and even endorsing – governments' protectionist tendencies and more tentative in its defense of freer trade in services. For example, in an April 2, 2001, WTO press release announcing the adoption of the GATS negotiating guidelines and procedures and the conclusion of an interim assessment of the negotiations, WTO Director General Mike Moore was prominently quoted (in the third paragraph of the press release) as follows: "One significance of the guidelines should not go unnoticed . . . Governments have unequivocally endorsed some of the fundamental principles of the GATS: Governments' right to regulate and to introduce new regulations on the supply of services in pursuit of national policy objectives; their right to specify which services they wish to open to foreign suppliers and under which conditions; and the overarching principle of flexibility for developing and least-developed countries" (WTO, 2001b).

Similarly, a WTO (2001c) publication, "GATS – Fact and Fiction," repeatedly emphasizes what GATS isn't, with the definite flavor of tolerance for protectionism:

"– Member Governments choose those service sectors or subsectors on which they will make commitments guaranteeing the right of foreign suppliers to provide the service. Each Member must have a schedule of commitments, but there is no minimum requirement as to its coverage – some cover only a small part of one sector";

"– For those services that are committed, Governments may set limitations specifying the level of market access and the degree of national treatment they are prepared to guarantee";

"– Governments were able to limit commitments to one or more of the four recognized 'modes of supply' through which services are traded. They may also withdraw and renegotiate commitments";

"– In order to provide more favourable treatment to certain trading partners, Governments may take exemptions, in principle limited to 10 years' duration, from the MFN principle, which is otherwise applicable to all services, whether scheduled or not" (WTO, 2001c, pp. 6–7).

A page later in the same document, "There is no obligation on any WTO Member to allow foreign supply of any service . . ." (WTO, 2001c, p. 8).

All of these descriptions of negative features of GATS are, of course, accurate; and they are duly accompanied by positive statements about the benefits and virtues of expanded trade and reduced barriers. Nevertheless, the prominence of these negative descriptions in WTO pronouncements and publications indicates a defensiveness that may well encourage the ever-present mercantilist/protectionist tendencies

within member countries and thereby cause greater delays in the negotiations, greater rigidities in the bargaining, and lesser reductions in barriers in the outcomes.

Excessive pessimism is probably not warranted. The processes of international trade in services have surely benefited, and will continue to benefit, from the existence of the WTO and GATS and the attention that has thereby been focused on both the existence of international trade in services and the existence of barriers that restrict that trade. The GATS 2000 negotiating round will surely – eventually – bring greater liberalization than existed before.[30] But the cause of freer trade usually needs all of the help that it can get to withstand the forces of mercantilism and protectionism. Backsliding by the WTO, even for understandable tactical reasons, does not help.

5 Conclusion

International trade in services is still a relatively new phenomenon: conceptually, statistically, and diplomatically. There is still much to be explored in all of these dimensions. It is likely that the early-twenty-first century will see substantial progress in all of them.

Acknowledgments

I would like to thank Gary Hufbauer and James Hartigan for valuable comments on an earlier draft.

Notes

1 See Kang (2000b) for recent cross-country data that show the generally rising proportion of services in GDP over time and across higher-income countries.
2 A complete bibliography is beyond the scope of this chapter. A fairly comprehensive list of papers and books, with links, can be found at http://www.SITrends.com, a website maintained by the Mark Twain Institute. Especially informative anthologies include Aharoni (1997), Aharoni and Nachum (2000), Claessens and Jansen (2000), Findlay and Warren (2000), Stephenson (2000), Sauve and Stern (2000), and Stern (2001).
3 Smuggling, of course, is an exception.
4 These modes are specified in Article I of GATS.
5 There is, of course, the question of whether the commercial presence abroad should be operated through a directly owned subsidiary or branch, or whether instead a licensing or franchising arrangement with an overseas partner should be arranged. This question applies equally strongly to goods or services. How it is resolved is dependent partly on the regulatory issues that are discussed below and partly on the nature of the product or service, the managerial capabilities of the provider, the contractual and monitoring possibilities, and so on. It is worth noting that GATS directly deals with issues related to commercial presence, whereas GATT does not.
6 See, e.g., Hoekman and Braga (1997) and the essays in Aharoni (1997) and Findlay and Warren (2000).
7 See Joskow and Rose (1989); Gruenspecht and Lave (1989); White (1996).
8 However, where a genuine market imperfection is present (and has motivated domestic regulation), then the theory of "second-best" tells us that the amelioration of another market imperfection (a reduction in trade barriers) will not necessarily improve welfare; see Lipsey and Lancaster (1956–7).

9 A more detailed analytical framework to support this position, along with analyses of exceptions and caveats, can be found in any international economics text; see, e.g., Yarbrough and Yarbrough (1997) or Pugel and Lindert (2000).
10 See, e.g., King and Levine (1993a, 1993b); Levine (1996, 1997, 1998, 1999); Levine and Zervos (1998); La Porta et al. (1997, 1998); Rajan and Zingales (1998); Beck et al. (2000).
11 This decrease is in contradiction to the predictions of Baumol (1967). But the Baumol predictions were based on a very limiting set of assumptions about the nature of the production and consumption of services.
12 It may be the case that these observed trade flows – which represent the cross-border, consumption-abroad, and temporary-presence modes – are being supplanted by the commercial-presence mode. This is only a hypothesis, however, which should be examined in future research.
13 The WTO (2001a) does not present services data aggregations for Eastern Europe and for the Middle East, although trade in goods aggregates were presented. This is another indicator of the lack of attention to services and to trade in services that has prevailed until recently.
14 Additional data and information can be found in USITC (2001).
15 The data in Tables 17.6 and 17.7 do not cover sales by bank affiliates, apparently because banks were not included in the surveys that were conducted to generate the data (Borga and Mann, 2003).
16 This is the most recent year available; the affiliates sales data appears consistently to lag behind the services trade flows data by a year.
17 Landefeld and Young (1988) provide earlier data that show net surpluses extending back to at least 1977.
18 The United States has, however, been selective in its advocacy and in its openness. Though the United States has relatively low barriers with respect to financial services and especially banking (McGuire and Schuele, 2000; Skipper, 2001; White, 2002) and telecommunications (Warren, 2000a, 2000b; Warren and Findlay, 2000), it has high barriers with respect to maritime shipping services (White, 1988; Fox and White, 1997; Kang, 2000a; Kang and Findlay, 2000; McGuire, Schuele, and Smith, 2000), to domestic air service (Kasper, 1988), and even to accountancy services (Colecchia, 2000). For example, US citizen ownership requirements for the provision of coastal maritime shipping services (and US citizenship for staffing as well), domestic air services, and over-the-air broadcasting and telephony constitute absolute barriers to entry.
19 A more extensive treatment of the diplomatic history of GATS can be found in Nicolaidis (1989).
20 Also, the 1988 international agreement among bank supervisory agencies (the "Basel Accord") on minimum capital requirements for banks was, in essence, an agreement to restrict the national levels of subsidies for banks (White, 1996).
21 An "Agreement on Trade-Related Aspects of Intellectual Property Rights" (TRIPS) was also reached in the Uruguay Round and included in the WTO, with a separate council to administer it.
22 More detailed descriptions of GATS can be found in, e.g., WTO (1999b, 2001d) and Feketekuty (2000a, 2000b).
23 However, aircraft repair and maintenance services, air transport marketing services, and computer-reservations systems are covered.
24 This is commonly described (somewhat perversely) as a "positive list" arrangement, as compared to the "negative list" arrangement that applies to goods, whereby only when a country lists a goods category and specific restrictions do the presumptions of free trade in the GATT not apply. But the GATS lists also have a "negative list" aspect: When a country

does bind itself with respect to a specific service category and mode, it then is committed to the GATS obligations with respect to that service and mode, unless it specifically exempts a GATS provision.

25 A separate "Government Procurement Agreement" (GPA) applies to GATT and GATS but is voluntary. Its goal is MFN, national treatment, and transparency in procurement. See Evenett and Hoekman (2000) for further discussion.

26 See, e.g., USITC (1996a, 1996b, 1997, 1998); Altinger and Enders (1996); Hoekman and Braga (1997); WTO (1999a); Hoekman (2000).

27 Skipper (2001) documents some examples of greater openness in financial services; Cowhey and Klimenko (2001) discuss the telecommunications agreement.

28 For examples of such efforts, see the essays in Findlay and Warren (2000).

29 The four-fifths figure is an overestimate, since a significant fraction of trade flows in goods is now unrestricted and thus unavailable for counteroffers. Also, there may be some implicit promises as to future concessions with respect to goods barriers that may be possible. Still, the unavailability of goods categories for bargaining will unavoidably restrict the bargaining process.

30 Forward-looking assessments of the GATS and recommendations for the future can be found in, e.g., Feketekuty (2000a, 2000b), Hoekman (2000), Hoekman and Messerlin (2000), Low and Mattoo (2000), Thompson (2000), and WTO (2001d).

References

Aharoni, Yair 1997: *Changing Roles of State Intervention in Services in an Era of Open International Markets*. Albany, NY: State University Press of New York.

Aharoni, Yair and Lilach Nachum 2000: *Globalization of Services: Some Implications for Theory and Practice*. London: Routledge.

Ascher, Bernard 1999: Trade disciplines for regulation: lessons from the accountancy sector. Paper delivered at the Conference on (R)Evolution of Quality and Competency Assurance in the Global Marketplace, Santa Fe, N.M., June 2–4. Washington, DC: USTR, draft 7/28/99, mimeo.

Altinger, Laura and Alice Enders 1996: The scope and depth of GATS commitments. *The World Economy*, 19, 307–32.

Baumol, William 1967: Macroeconomics of unbalanced growth: the anatomy of urban crisis. *American Economic Review*, 57, 415–26.

Beck, Thorsten, Asli Demirguc-Kunt, Ross Levine, and Vojislav Maksimovic 2000: Financial structure and economic development: firm, industry, and country evidence. Policy Research Working Paper No. 2423, World Bank, Washington, DC, June 14.

Borga, Maria and Michael A. Mann 2001: U.S. international services: cross-border trade in 2000 and sales through affiliates in 1999. *Survey of Current Business*, 81, 49–95.

Borga, Maria and Michael A. Mann 2003: U.S. international services: cross-border trade in 2002 and sales through affiliates in 2001. *Survey of Current Business*, 83, 58–118.

Claessens, Stijn and Marion Jansen (eds.) 2000: *The Internationalization of Financial Services: Issues and Lessons for Developing Countries*. The Hague: Kluwer Law International.

Claessens, Stijn, Asli Demirguc-Kunt, and Harry Huizinga 2000: How does foreign entry affect the domestic banking market? In Stijn Claessens and Marion Jansen (eds.), *The Internationalization of Financial Services: Issues and Lessons for Developing Countries*. The Hague: Kluwer Law International, chapter 6.

Colecchia, Alessandra 2000: Measuring barriers to market access for services: a pilot study on accountancy services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 245–66.

Cowhey, Peter F. and Mikhail M. Klimenko 2001: Implementing telecommunications liberalization in developing countries after the WTO Agreement on basic telecommunications services. In Robert M. Stern (ed.), *Services in the International Economy*. Ann Arbor: University of Michigan Press, 349–68.

Deardorff, Alan V. 2000: International provision of trade services, trade, and fragmentation. Discussion Paper No. 463, Research Seminar in International Economics, School of Public Policy, University of Michigan, Ann Arbor, Mich., August 30.

Evenett, Simon J. and Bernard M. Hoekman 2000: Government procurement of services and multilateral disciplines. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 143–64.

Feketekuty, Geza 2000a: Assessing and improving the architecture of GATS. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 85–111.

Feketekuty, Geza 2000b: Improving the architecture of the general agreement on trade in services. In Sherry M. Stephenson (ed.), *Services in the Western Hemisphere: Liberalization, Integration, and Reform*. Washington, DC: Brookings Institution, 19–42.

Findlay, Christopher and Tony Warren (eds.) 2000: *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge.

Fox, Nancy Ruth and Lawrence J. White 1997: U.S. ocean shipping policy: going against the tide. *Annals of the American Academy of Political and Social Science*, 553, 75–86.

Gruenspecht, Howard K, and Lester B. Lave 1989: The economics of health, safety, and environmental regulation. In Richard Schmalensee and Robert D. Willig (eds.), *Handbook of Industrial Organization*, vol. 2. Amsterdam: North Holland, 1507–50.

Hoekman, Bernard M. 2000: Towards a more balanced and comprehensive services agreement. In Jeffrey J. Schott (ed.), *The WTO after Seattle*. Washington, DC: Institute for International Economics, 119–35.

Hoekman, Bernard M. and Carlos A. Primo Braga 1997: Protection and trade in services: a survey. *Open Economies Review*, 8, 285–308.

Hoekman, Bernard M. and Patrick A. Messerlin 2000: Liberalizing trade in services: reciprocal negotiations and regulatory reform. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 487–508.

Joskow, Paul L. and Roger G. Noll 1994: Economic regulation: deregulation and regulatory reform during the 1980s. In Martin Feldstein (ed.), *American Economic Policy in the 1980s*. Chicago: University of Chicago Press, 367–440.

Joskow, Paul L. and Nancy L. Rose 1989: The effects of economic regulation. In Richard Schmalensee and Robert D. Willig (eds.), *Handbook of Industrial Organization*, vol. 2. Amsterdam: North Holland, 1449–506.

Kang, Jong-Soon 2000a: Price impact of restrictions on maritime transport services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 189–200.

Kang, Jong-Soon 2000b: The services sector in output and in international trade. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 18–41.

Kang, Jong-Soon and Christopher Findlay 2000: Regulatory reform in the maritime industry. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 152–71.

Kasper, Daniel M. 1988: *Deregulation and Globalization: Liberalizing International Trade in Air Services*. Cambridge, Mass.: Ballinger.

King Robert G. and Ross Levine 1993a: Finance and growth: Schumpeter might be right. *Quarterly Journal of Economics*, 108, 717–37.

King Robert G. and Ross Levine 1993b: Finance, entrepreneurship, and growth: theory and evidence. *Journal of Monetary Economics*, 32, 513–42.

Landefeld, J. Steven and Kan H. Young 1988: US trade in services: 1970–85. In Wray O. Candilis (ed.), *United States Service Industries Handbook*. New York: Praeger, 91–126.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andre Schleifer, and Robert W. Vishny 1997: Legal determinants of external finance. *Journal of Finance*, 52, 1131–50.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andre Schleifer, and Robert W. Vishny 1998: Law and finance. *Journal of Political Economy*, 106, 1133–55.

Levine, Ross 1996: Foreign banks, financial development, and economic growth. In Claude Barfield (ed.), *International Financial Markets: Harmonization versus Competition*. Washington: American Enterprise Institute, 224–54.

Levine, Ross 1997: Financial development and economic growth: views and agenda. *Journal of Economic Literature*, 35, 688–726.

Levine, Ross 1998: The legal environment, banks, and long-run economic growth. *Journal of Money, Credit and Banking*, 30, 596–613.

Levine, Ross 1999: Law, finance, and economic growth. *Journal of Financial Intermediation*, 8, 8–35.

Levine, Ross and Sara Zervos 1998: Stock markets, banks, and economic growth. *American Economic Review*, 88, 537–58.

Lipsey, Richard G. and Kelvin Lancaster 1956–7: The general theory of the second best. *Review of Economic Studies*, 24, 11–32.

Low, Patrick and Aaditya Mattoo 2000: Is there a better way? Alternative approaches to liberalization under GATS. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 449–72.

McGuire, Greg and Michael Schuele 2000: Restrictiveness of international trade in banking services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 201–14.

McGuire, Greg, and Michael Schuele, and Tina Smith 2000: Restrictiveness of international trade in maritime services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 172–88.

Nicolaidis, Kalypso 1989: Learning while doing: how services got on the Uruguay Round agenda. In Albert Bressand and Kalypso Nicolaidis (eds.), *Strategic Trends in Services: An Inquiry into the Global Service Economy*. New York: Ballinger, 161–79.

Noll, Roger G. 1989: Economic perspectives on the politics of regulation. In Richard Schmalensee and Robert D. Willig (eds.), *Handbook of Industrial Organization*, vol. 2. Amsterdam: North Holland, 1253–87.

Pugel, Thomas G. and Peter H. Lindert 2000: *International Economics*, 11th edn. New York: McGraw-Hill.

Rajan, Raghuram H. and Luigi Zingales 1998: Financial dependence and growth. *American Economic Review*, 88, 559–86.

Sauve, Pierre and Robert M. Stern (eds.) 2000: *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution.

Skipper, Harold D., Jr. 2001: *Insurance in the General Agreement on Trade in Services*. Washington, DC: AEI Press.

Stephenson, Sherry M. (ed.) 2000: *Services in the Western Hemisphere: Liberalization, Integration, and Reform*. Washington, DC: Brookings Institution.

Stern, Robert M. (ed.) 2001: *Services in the International Economy*. Ann Arbor: University of Michigan Press.

Thompson, Rachel 2000: Formula approaches to improving GATS commitments. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 473–86.

U.S. International Trade Commission 1996a: *General Agreement on Trade in Services (GATS): Examination of Major Trading Partners' Schedules of Commitments*, Publication 2940, January.

U.S. International Trade Commission 1996b: *General Agreement on Trade in Services (GATS): Examination of South American Trading Partners' Schedules of Commitments*, Publication 3007, December.

U.S. International Trade Commission 1997: *General Agreement on Trade in Services (GATS): Examination of Asia/Pacific Trading Partners' Schedules of Commitments*, Publication 3053 August.

U.S. International Trade Commission 1998: *General Agreement on Trade in Services: Examination of the Schedules of Commitments Submitted by Eastern Europe, the European Free Trade Association, and Turkey*, Publication 3127, September.

U.S. International Trade Commission 2001: *Recent Trends in U.S. Services Trade: 2001 Annual Report*, Publication 3409, May.

Warren, Tony 2000a: The impact on output of impediments to trade and investment in telecommunications services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 85–100.

Warren, Tony 2000b: The identification of impediments to trade and investment in telecommunications services. In Christopher Findlay and Tony Warren (eds.), *Impediments to Trade in Services: Measurement and Policy Implications*. London: Routledge, 71–84.

Warren, Tony and Christopher Findlay 2000: Measuring impediments to trade in services. In Pierre Sauve and Robert M. Stern (eds.), *GATS 2000: New Directions in Services Trade Liberalization*. Washington, DC: Brookings Institution, 57–84.

White, Lawrence J. 1988: *International Trade in Ocean Shipping Services: The U.S. and the World*. Cambridge, Mass.: Ballinger.

White, Lawrence J. 1996: Competition versus harmonization: an overview of international regulation of financial services. In Claude Barfield (ed.), *International Trade in Financial Services*. Washington: American Enterprise Institute, 5–48.

White, Lawrence J. 2001: *Reducing the Barriers to International Trade in Accounting Services*. Washington, DC: AEI Press.

White, Lawrence J. 2002: Unilateral trade openness: the experience of the U.S. financial services sector. In Jagdish Bhagwati (ed.), *Going Alone: The Case for Relaxed Reciprocity in Freeing Trade*. Cambridge, Mass.: MIT Press, 445–81.

Winston, Clifford 1993: Economic deregulation: days of reckoning for microeconomists. *Journal of Economic Literature*, 31, 1263–89.

World Trade Organization, Council for Trade in Services 1998: Accountancy services, December 4.

World Trade Organization, Council for Trade in Services 1999a: Structure of commitments for modes 1, 2, and 3.

World Trade Organization, Secretariat, Trade in Services Division 1999b: An introduction to the GATS.

World Trade Organization 2001a: *International Trade Statistics 2001*. Geneva: World Trade Organization.

World Trade Organization 2001b: WTO services talks press ahead. Press release.

World Trade Organization 2001c: GATS – fact and fiction.

World Trade Organization 2001d: *Guide to the GATS: An Overview of Issues for Further Trade Liberalization in Services*. The Hague: Kluwer Law International.

Yarbrough, Beth V. and Robert M. Yarbrough 1997: *The World Economy: Trade and Finance*, 4th edn. New York: Dryden Press.

18

The Dynamic Effects of Trade Liberalization and Environmental Policy Harmonization

Larry Karp and Jinhua Zhao

Chapter Outline

We study a general equilibrium model in which trade is driven by institutional differences (differences in environmental distortions) and differing factor endowments (differences in production-related environmental stocks). The endogenous evolution of environmental stocks depends on the trade regime and on the extent of environmental distortions in the two countries. The economy can have multiple steady states that depend on the policy regime and the initial stock levels. We investigate the effects of harmonization of the distortions and of trade liberalization, in the short and the long run. The model illustrates the ambiguous effects of these reforms, and it explains why the predictions of environmentalists or economists are more likely to be borne out, depending upon whether the environment is fragile or resilient.

1 Introduction

Mainstream economists and environmentalists frequently find themselves on opposite sides of the "trade versus environment" debate. Economists tend to support liberal trade for reasons having little to do with the environment. Environmentalists often oppose trade liberalization because they think that it is likely to harm the natural environment. Economists argue that in the long run, trade will help the environment as income rises, promoting stricter environmental regulation, as environmentally friendly technologies

are diffused, or as imports of resource-intensive goods relieve environmental pressures in resource-scarce countries. Environmentalists worry that even if nations gain from trade in the short run, they may suffer as trade eventually exhausts the natural environment in the resource exporting country. Economists are skeptical of efforts to harmonize environmental regulations across nations at different stages of development. Many environmentalists think that this kind of harmonization becomes more essential with liberalized trade.

The exchange between Daly (1993) and Bhagwati (1993) exemplifies the debate between environmentalists and economists in academic circles. The disagreement over whether to grant US presidents fast-track negotiating authority for trade agreements reflects a wider debate over the merits of more liberal trade. The riots surrounding recent WTO and IMF meetings show the intensity of opposition, in some circles, to increased globalization in general, and to more liberal trading rules in particular.

The differences in opinion between the two groups might be explained by differences in values: perhaps economists and environmentalists have different assessments of the relative importance of economic growth and of the natural environment. However, the widening difference between the two groups may also be due to the polarization that occurs when each side attempts to present its case using sound-bites. The resulting debate becomes a repetition of "talking points" that only makes each side appear more entrenched.

This chapter incorporates environmental concerns in a framework for analyzing the effects of trade liberalization and policy harmonization. Economists know that these kinds of policy changes may have ambiguous welfare effects. Many noneconomists think that economic models are vacuous because the models' conclusions are predetermined in a trivial and biased manner. We show that a simple model can support contradictory positions, depending in a sensible manner on parameter values in the model. Thus, the arguments of different camps can be viewed as the expressions of different assumptions about these parameter values, rather than as worldviews that are radically different and mutually exclusive. For example, our model predicts different outcomes depending on whether the environment is "fragile" or "resilient," on whether institutional differences between trading partners are large or small, and on whether we consider the short or the long run.

The underlying arguments in favor of free trade and opposed to mandatory harmonization are familiar, but worth mentioning. By enabling countries to exercise their comparative advantage, trade makes it possible to produce more goods with the same amount of resources, increasing wealth. Trade may also create dynamic benefits by encouraging the spread of new technologies; by enlarging markets, trade can lead to the exploitation of economies of scale, again increasing wealth; increased wealth might promote policies that protect the environment. Cross-country differences in environmental policies might reflect differences in income, tastes, capital stocks, resource endowments, or a variety of other factors that contribute to interindustry trade. In this case, the mandatory harmonization of environmental policies thwarts the efficient workings of the market and lowers welfare.

These arguments assume that markets work efficiently, and therefore that the goal of policy reform should be to reduce market frictions such as trade restrictions. Mainstream economic theory recognizes that in the presence of multiple market imperfections,

trade liberalization or some other market reform can either increase or lower welfare. This ambiguity, which is known as the "Theory of the Second Best," plays a central role in (academic) economic policy analysis, but is often swept under the carpet when presenting the economic arguments for trade liberalization or against mandatory environmental harmonization. For purposes of analyzing trade reform, economists sometimes assume that non-trade distortions are of secondary importance, or are not closely related to trade.

Environmentalists tend to think that other market imperfections are of primary importance, and that these imperfections are closely related to trade. We adopt this view in our model. In particular, we assume that many standard reasons for trade, such as tastes, technology, nonenvironmental endowments, and other characteristics are the same in the two countries. The countries differ only in their endowment of the natural resource, and in the degree of market imperfections. These two differences are the only reasons for trade.

Market imperfections arise for many reasons. We examine the situation where an institutional failure causes excessive exploitation of an environmental resource. For example, there may be imperfect property rights to the environmental resource. Whatever the exact cause of the market failure, society attempts – with limited success – to reduce the resulting economic distortion. For example, it might tax the use of the resource or impose quantitative restrictions. We assume that society is not able to perfectly correct the distortion. Hereafter, we consider the market failure to be imperfect property rights. Strengthening these property rights leads to a reduction in the socially excessive level of exploitation of the environmental resource. We use "environmental reform," or merely "reform" to describe an institutional or policy change that would lead to a smaller exploitation of natural resources if other variables (such as prices) were held constant. Harmonization means that the magnitudes of the distortions in the two countries are made more similar.

Within our model, the short-run effects of trade and harmonization, that is, the instantaneous effects given the environmental stocks, are standard. For example, the difference in the endowment of the natural resource is a standard source of gains from trade. To the extent that this difference is significant, trade liberalization improves economic efficiency and welfare. The difference in property rights (or some other environmental market failure) is an institutional reason for trade. When this difference is important, relative to the difference in the resource endowment, trade liberalization or the harmonization of policies have ambiguous effects.

The long- and short-run effects can be significantly different. The long-run effects arise from changes in the steady states as the environmental stocks change endogenously. The manner in which the stock changes depends on the trade regime (autarky or free trade), and on the levels of the environmentally related market failure. It is possible that both countries gain from trade or harmonization in the long run, even when they lose in the short run. The critical insight is that, depending on the environmental resilience, trade can either enable a resource-rich country to "pull up" a resource-poor country (so both resources are preserved), or cause the latter to "drag down" the former (so that both resources are depleted).

We emphasize the long-run effects and the role of environmental resilience in determining these effects. This focus is motivated by the concern of many environmentalists

about the possible depletion of the environmental resource. Different resources have different regenerating capabilities, so environmental resilience naturally plays a significant role in our long-run analysis.

A number of papers, including Chichilnisky (1993, 1994), Brander and Taylor (1997a, 1997b, 1998), and Copeland and Taylor (1994, 1995) study market failures and the relationship between trade and the environment. Our approach is different in that we emphasize the possibility of multiple steady states, which enables us to investigate qualitatively different effects of policy changes. For example, autarkic and free trade economies may reach different kinds of steady states, depending on the environmental distortions, resilience, and the starting stock levels; environmental reform can lead to qualitatively different effects in the short and the long run, and under free trade and autarky. The model that we describe in Sections 4–7 follows the analysis in Karp, Sacheti, and Zhao (2001) and Karp, Zhao, and Sacheti (2003). Those two papers contain all of the technical details that we omit here.

2 Empirical Review

There is a growing body of empirical literature on the relationship between trade and the environment. Although no consensus has been reached regarding this relationship, there is compelling evidence that overall trade may improve the environment through raising the income of the trading countries (Antweiler, Copeland, and Taylor 2001). Economists have conjectured that when countries have different degrees of environmental regulation, those with loose regulation may export more pollution-intensive goods and even attract polluting industries (the pollution heaven hypothesis). However, the empirical evidence on this conjecture is far from clear. Econometric studies find little evidence of a relation between aggregate trade flows and differing levels of environmental protection, although Mani and Wheeler (1998) present evidence that trade may create transitory pollution havens.

There is some anecdotal evidence, mostly at a commodity-specific level, indicating that differing levels of market failure do influence trade flows. For example, in response to serious deforestation, China restricted logging in 12 provinces in 1998, and in 18 provinces in the year 2000. This logging ban, together with continued economic growth and a reduction in tariffs, has caused China to become one of the world's largest importers of logs. Burma, where logging is controlled by warlords, and where the environmental market failure is probably more severe than in prereform China, has become a primary source of supply. The environmentally beneficial policy in China could worsen the regional environmental problems by increasing the pressure on Burmese resources (Pomfret, 2001).

The empirical literature has also failed to identify the *dynamics* of the trade pattern in response to changes in the resource and environmental bases of the trading countries. A resource-rich country may exhaust its resource if it keeps exporting the resource-intensive good. This possibility is a major concern of anti-trade environmentalists who fear that even if the country may benefit from trade in the short run, it may lose in the long run as the stock is driven down. Again, there is only anecdotal evidence supporting this conjecture. For example, Thailand and the Philippines were major timber exporters

in the 1970s and early 1980s. However, they eventually exhausted many of their forest stocks and became net importers of roundwood (FAO, 1994). Trade and loosely defined property rights contributed to the exhaustion (McDowell, 1989).

One of our major conclusions is that the long-run effects of trade on the environment and on welfare depend in large part on the resilience of the resource stock, that is, on how fast the resource can regenerate. In our model, trade is likely to have positive (negative) long-run effects when the resource is relatively resilient (fragile). There is no empirical investigation in the literature on how the relationship between trade and the environment depends on the nature of the environment. To the extent that trade leads to expanded economic activity and increased resource extraction in the resource-exporting country, studies on the time path of economic growth in resource-based economies offer indirect evidence. Diamond (1999) compares the effects of technological advance and population growth in the Fertile Crescent (the Near East) and other Eurasian regions, over a period of thousands of years. Both regions experienced rapid growth with the advent of agriculture, but the Fertile Crescent "had the misfortune to arise in an ecologically fragile environment," due possibly to low rainfall (Diamond, 1999, p. 411). Eventually agriculture destroyed the resource base, leading to devastating desertification and salinization. On the other hand, Europe and China had adequate rainfall and a resilient resource base. The development of agriculture and the subsequent economic growth spurred further growth, which continues today. Brander and Taylor (1998b) document the rise and fall of Easter Island, again due to its fragile resource base.

3 Three Ideas from Economic Theory

In order to assess the model and the results described in following sections, the reader should understand three important ideas from economic theory, discussed in this section. The Principle of Comparative Advantage helps in understanding why trade tends to increase welfare; the Theory of the Second Best shows why there may be circumstances where trade lowers welfare; and the Principle of Targeting explains why economists generally favor liberal trade even in a "second best" environment.

In many respects, the model in this chapter is standard: trade allows countries to exercise their comparative advantage, leading to possible efficiency gains. However, the existence of an environmentally related distortion can cause trade to lower welfare. The reader might leave with the message that the theory implies only that "anything can happen," not a particularly helpful insight, and not one that promotes the adoption of liberal trading rules. The fact that liberalized trade or environmental reform can lead to a wide range of possible outcomes does not imply that there is a theoretical argument for discouraging either trade or environmental reform. The purpose of our model is to improve our intuition about the likely effects of trade and different types of reform in different settings, and in particular to investigate differences in the short and long run and to emphasize the role of environmental resilience. Given the assumptions that we need to make in order to study these issues, our model does not provide a basis for policy recommendations. In this section we hope to give the noneconomist a sufficient grounding in economic theory to avoid leaving with the wrong message.

The theoretical argument in favor of liberal trade is based on ideas concerning the efficiency of market outcomes, and on the Principle of Comparative Advantage. This principle implies that under free trade and given that domestic markets are not distorted, a country exports commodities that it can produce relatively more cheaply than its partners. "More cheaply" is understood in terms of "opportunity costs" rather than in terms of dollars or labor hours. It is convenient to explain the meaning of opportunity cost using a two-commodity example, since our model below involves only two commodities.

Suppose that food and steel are the only two commodities. In this setting, the opportunity cost of steel is simply the number of units of food that the economy must sacrifice in order to obtain one more unit of steel. The economy converts food into steel by reallocating factors of production from the food to the steel sector. Firms that use the factors of production, and workers and landowners who supply these factors, maximize their profits, utility, or rent. In an undistorted, competitive equilibrium, it is not possible to increase output of one commodity without decreasing output of the other commodity: the allocation of factors is efficient. The opportunity cost of steel equals the equilibrium relative price of steel, $p^{\mathrm{s}}/p^{\mathrm{f}}$, where p^{s} is the nominal price of steel and p^{f} is the nominal price of food.

If two economies have different equilibrium relative prices in autarky, then one economy necessarily has a lower equilibrium relative price of steel. That economy has a lower opportunity cost of steel – a comparative advantage in steel – and it exports steel in a free trade equilibrium. When the equilibrium trade price differs from the autarkic prices, trade increases total income in both the countries. The income of owners of any particular factor of production might fall as a consequence of trade, but the magnitude of this fall is less than the gain in income of owners of other factors of production, so national income increases.

This conclusion depends on the assumption that the economies are "undistorted." Anything that causes the economy to be at an inefficient equilibrium can be viewed as a distortion, including imperfect competition, missing markets (e.g., absence of insurance markets), or government policies that restrict trade (e.g., tariffs). The Theory of the Second Best states that if there are two or more market imperfections (distortions), correcting one of them may either increase or decrease welfare. For example, if there are two tariffs, eliminating one may not increase welfare. A pessimistic interpretation of this conclusion is that it implies that economic theory allows us to reach no conclusion about real-world markets, since we know that these are subject to many imperfections. A more moderate interpretation is that we cannot uncritically use economic theory to conclude that a particular reform, such as trade liberalization, necessarily improves efficiency.

We provide a simple example of the theory of the second best and then discuss its application to our setting. Imagine an economy in which there are only two market failures, both of which are present in a particular sector. The first failure is that production of the commodity damages the environment, but the producer does not pay for this damage (i.e., there is a negative environmental externality). The second failure is that the industry is oligopolistic rather than competitive. These two market imperfections cut in opposite directions. The first causes the market outcome to result in excessive production, from the standpoint of society. The second causes the market outcome to

result in too little production, since oligopolists (typically) sell at a point where price is greater than marginal cost. At this level of generality, we do not know whether there is too little or too much production on balance. We cannot conclude that welfare would be higher if we remove one of the imperfections, for example, by forcing the oligopolists to produce where price equals marginal cost in order to mimic the competitive outcome. The salient feature of this example is that each distortion affects the welfare cost of the other distortion.

Our model allows for trade restrictions and an environmental distortion in each country. The Theory of the Second Best implies that a reduction in one of these distortions, such as the elimination of the trade restriction and the switch from autarky to free trade, might increase or lower welfare. Similarly, under free trade, environmental reform in one country leads to changes in commodity prices that can affect the economic demands on the environment in the other country. Since there are two distortions under free trade – the imperfect property rights in each country – reform in a particular country might improve or worsen welfare, either at the national or global level. These possibilities arise in our model. However, our objective is not to illustrate the Theory of the Second Best, but to understand why short- and long-run effects may differ.

The Theory of the Second Best may appear to undercut economic arguments for liberal trade or for improving property rights to environmental goods. Policy decisions in the real world always involve many distortions that policymakers reasonably regard as fixed. A set of theoretical results known as the Principle of Targeting (Bhagwati, Panagariya, and Srinivasan, 1998) explain why economists remain broadly united in favor of liberal trade and strong property rights. This principle merely states that distortions, or market failures, should be "targeted" as directly as possible.

For example, suppose that policymakers believe that domestic agricultural production provides security that benefits society. Consumers do not pay for, nor are producers compensated for, this security. The competitive equilibrium results in too little domestic production, from the standpoint of society, because producers do not take the social benefit of improved security into account when making their production decisions. In this situation, an agricultural import tariff (or export subsidy) improves the competitive equilibrium. Such a tariff or subsidy is a second-best policy: it ameliorates the distortion caused by producers' failure to internalize the positive production externality, but in the process it creates a consumption distortion. On balance, the second-best tariff improves welfare, but it does so at a cost. The first-best policy in this example is an agricultural production subsidy. The "target," after all, involves the level of production. A production subsidy achieves this target without generating other distortions.

In our model, the underlying market failure is an environmental distortion. The Principle of Targeting implies that the optimal policy should address this distortion directly. When environmental policy fails to eliminate the market distortion – as we assume – trade restrictions can lead to net welfare gains, but they do so by introducing other distortions.

Our model has only two types of departure from an efficient equilibrium, the environmental distortions and trade restrictions. We study only trade liberalization and a particular kind of environmental reform. Consequently, our analysis does not provide a basis for policy recommendations – and certainly it does not support an anti-trade or anti-property rights agenda.

4 A General Model

The long-run effects of trade and environmental reform depend on the steady state that the economies reach. In our model, economies may have multiple steady states, and these may differ qualitatively. Here we provide a graphical analysis that illustrates why there might be multiple steady states under both autarky and free trade, and the differences between the steady states. Moving from autarky to free trade can lead to qualitatively different steady states, and the effect of environmental reform can be qualitatively different under autarky and free trade.

In our model, the environmental problem is national rather than global. Production decisions within a country directly influence the environment only in that country. Thus, under autarky, decisions within a country have no effect on the other country. International trade links the markets in the two countries. With trade, production decisions within a country affect world prices, and these affect production decisions and the environment in other countries. Trade causes national environmental problems to become multinational problems.

In the following three subsections, we explain why there may be multiple steady states under both trade and autarky. We show how trade can change the steady state that an economy approaches; then we explain how environmental reform can change the equilibrium when we hold the trade regime fixed.

4.1 The Existence of Multiple Steady States

Suppose that the production of final goods requires environmental services (E), the supply of which is endogenous. The cost of producing E (a flow variable) decreases with the environmental stock Z. The E-producing industry has a market failure that leads to an inefficiently high exploitation of the environmental stock and an inefficiently high supply of environmental services, for a given stock level. The magnitude of this distortion is measured by δ; a larger value of δ implies a greater market failure.

The equilibrium supply of E depends on both the market failure and the current environmental stock $Z : E = E(Z, \delta)$. A larger environmental stock decreases the cost of supplying environmental services but at a decreasing rate, so that $E_Z \geqslant 0$ and $E_{ZZ} \leqslant 0$. A larger market failure increases the equilibrium supply of environmental services for a given stock: $E_\delta \geqslant 0$. Environmental reform means that the distortion is reduced, that is, δ is reduced.

In order to obtain a specific functional form for $E(Z, \delta)$ we need to specify the market failure and the nature of the producer's optimization problem (among many other things). At this stage, we do not need that degree of detail. The intuition for our results depends on the curvature of the extraction function, described above.

To complete the description of the model we assume that the natural growth rate of the environmental stock (absent extraction) is a strictly concave function $G(Z)$ that increases for small Z, reaches a maximum, and then decreases to zero (at the natural carrying capacity). The steady state of the autarkic economy depends on the relation between $G(Z)$ and $E(Z, \delta)$. Figure 18.1(a) illustrates three possible configurations.

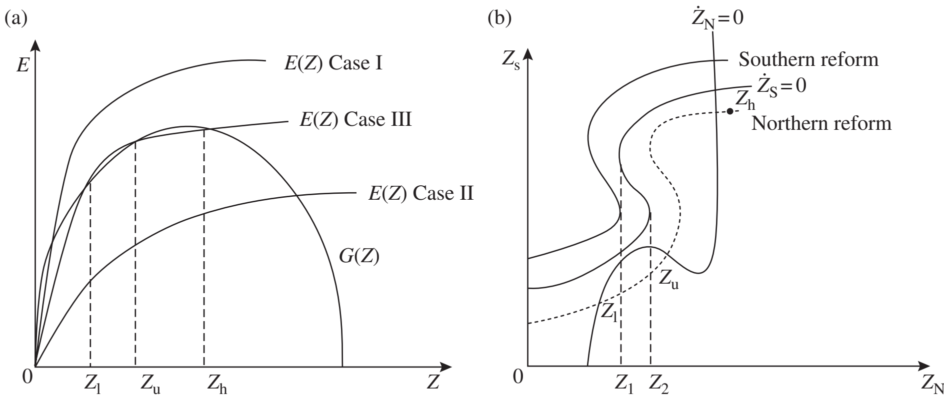

Figure 18.1 Possibilities of steady states. (a) Autarky and (b) Trade

In Case I, there is a unique low steady state and in Case II, there is a unique high steady state. Case III shows a situation where $G(Z)$ and $E(Z, \delta)$ intersect at three points, a low steady state, Z_l, a high steady state, Z_h and an intermediate (unstable) steady state, Z_u. The economy moves toward either the high or the low steady state, depending on whether the initial level of Z is above or below Z_u.

In a trade equilibrium, two countries, North and South, exchange commodities that use environmental services as inputs. At a point in time, their environmental stocks are Z_N and Z_S. As was the case under autarky, E_i (the equilibrium supply of environmental services in country i) depends on Z_i via its direct effect on production costs, and on the market failure, δ_i; E_i also depends on the price of environmental services, which depends on the aggregate (world) supply of services. With trade, the supply of environmental services in country i (E_i) therefore depends on the environmental stocks and the market failures in both countries. An increase in Z_N, for example, decreases North's relative costs of producing environmental services. Under plausible circumstances, higher Z_N increases the equilibrium supply of E_N and decreases the equilibrium supply of E_S.

Figure 18.1(b) shows the relation between the two countries' steady states. The heavy line labeled $\dot{Z}_S = 0$ shows the set of stocks in $Z_N - Z_S$ space at which South's stocks are in long-run equilibrium. This curve can be non-monotonic, as shown. For $Z_N < Z_l$ (i.e., for low levels of Z_N), South produces environmental services not only for domestic use, but also for export (embodied in final products). Thus, under trade, a low level of Z_N implies that the steady state for Z_S is also low. (This situation corresponds to the case where the graph of E_S as a function of Z_S is high, as in Case I in Figure 18.1(a); there $\dot{Z}_S = 0$ has a unique solution at a low steady state.) Increases in Z_N shift the supply of environmental services from South to North, and therefore shift down the graph of E_S (as a function of Z_S). This change increases South's low steady state. When North's environmental stocks are high ($Z_N > Z_2$), North provides the bulk of environmental services and South's extraction is low. Here, Case II in Figure 18.1(a) applies; there is a unique solution to $\dot{Z}_S = 0$, the high steady state. Again, increases in Z_N raise this steady state. For intermediate levels $Z_l < Z_N < Z_2$, Case III in Figure 18.1(a) applies; there are two stable and one unstable solutions to $\dot{Z}_S = 0$. Over this region, an increase

in Z_N increases both of South's stable steady states and decreases the unstable steady state.

The curve labeled $\dot{Z}_N = 0$ shows the set of steady states for North, as a function of South's stock. Figure 18.1(b) illustrates the case where there is a unique high steady state for both countries under trade: the heavy solid curves intersect at a single point, involving high stocks in both countries. There are, however, several other possibilities. For example, the graphs might intersect more than once (giving rise to multiple steady states) or they might intersect only at low steady states.

4.2 The Effects of Trade on the Steady States

Figure 18.1(b) indicates the possibilities that both countries reach high or low steady states under trade. Although at this level of generality it is impossible to precisely link the autarky and trade steady states, it is conceivable that under autarky the two countries may reach different steady states due to their differences in either the initial stock levels or the environmental distortions. For example, it is possible that North is described by Case II while South is described by Case I in Figure 18.1(a). Then trade may cause a *qualitative* change in the steady state for at least one country: If under trade the relevant $\dot{Z}_N = 0$ and $\dot{Z}_S = 0$ curves are the bold lines in Figure 18.1(b), both countries reach high steady states and trade increases long-run aggregate welfare. But if both countries reach the low steady states, trade might reduce long-run aggregate welfare.

The particular process through which the two countries move to the high or low steady states can only be studied using a more specific model. Here we outline two possibilities. Suppose that under autarky North is described by Case II (a unique high steady state) and South is described by Case I (a unique low steady state) in Figure 18.1(a), and that the two countries start at similar stock levels. South has higher extraction due to its worse environmental distortion, making it likely that South exports the environmentally intensive good. Trade initially further depletes South's stocks and leads to an increase in North's stocks. Eventually, comparative advantage in the environmentally intensive good shifts from South to North. This change can enable South's stocks to recover, and eventually both countries can reach the high steady states. In this case, North "pulls up" the South under trade. Another possibility is that the higher North extraction eventually depletes the North's stock without enabling South's stocks to recover. In this case, both countries approach the low steady states: trade causes South to "drag down" North. The actual outcome depends in part on how quickly Southern stock can recover, that is, on the resource's resilience. The "pull up" scenario is more likely when the resource is more resilient.

This example merely shows that, in the presence of multiple distortions (environmental distortions in both countries and trade restrictions), the elimination of one distortion (here, the trade restriction) can lower welfare. The example illustrates the Theory of the Second Best. At this level of generality we can only conclude that the effects of trade are ambiguous. In order to understand how the economic and environmental fundamentals are likely to determine these effects, we need a less general model.

4.3 The Effects of Environmental Reform on the Steady States

We first consider the case of autarky in Figure 18.1(a). Reform corresponds to a decrease in δ, leading to a downward shift in the curve $E(Z, \delta)$. The market failure means that the environmental sector absorbs more inputs than is socially optimal. Reform reduces the market failure and decreases the inputs used in extraction, increasing welfare in the short run.

The long-run welfare effects of the reform depend on the change in the steady states. A large reform (a sufficiently large downward shift in $E(Z, \delta)$) can eliminate the low steady state, leaving the high steady state as the unique long-run equilibrium (Case II). When this happens, reform can lead to a qualitative change in the steady state, and a correspondingly large welfare change.

With trade (Figure 18.1(b)), environmental reform in one country, say North, increases its costs of producing the environment-intensive good. This reform decreases North's equilibrium extraction level and increases extraction in South. At the initial relative prices, a reduction in North's environmental distortion leads to a more efficient allocation of resources in North and a rise in its welfare. The decrease in North's supply of environmental services increases the price of the commodity that uses these services relatively intensively. If North is an importer of the environmentally intensive good, reform causes its terms of trade to deteriorate. The net short-run welfare impact for North depends on the relative magnitudes of the direct effect and the price effect. At the initial relative prices, North's reform increases the relative severity of the distortion in South. By making South's allocation of resources less efficient, this reform tends to lower South's welfare. However, if South exports the environmentally intensive good, North's reform improves South's terms of trade. Again, the net short-run welfare effect depends on the relative magnitudes of the direct effect and the price effect.

To study the long-run effects of reform, we again consider the changes in the steady states as one or both countries reform. For any level of Z_N, reform in South increases its stable steady states. The mechanism is the same as under autarky. Southern reform therefore shifts the $\dot{Z}_S = 0$ isocline up to the left (the thin solid curve in Figure 18.1(b)). Reform in North shifts production of environmental services to South, lowering South's stable steady states. North's reform shifts the $\dot{Z}_S = 0$ isocline down to the right (the dashed curve in Figure 18.1(b)). (These two reforms have similar effects on the curve labeled $\dot{Z}_N = 0$. We do not show these effects in order to avoid cluttering the graph.)

Suppose North reduces its environmental distortion. This reform shifts the $\dot{Z}_N = 0$ curve to the right, to the dotted curve. In addition, as explained above, North's reform shifts down the $\dot{Z}_S = 0$ curve. Figure 18.1(b) illustrates a situation in which North's reform leads to two stable (and one unstable) steady states. (The two post-reform dotted curves intersect at three points.) If the initial condition is such that the equilibrium moves toward the low steady state, North's reform leads to lower stocks and a likely welfare reduction in the long run in both countries. In this case, even if Northern reform improves the short-run welfare of one country, in the long run the welfare in both countries can fall because reform lowers the steady state environmental stocks.

Comparing (a) and (b) in Figure 18.1 also shows that the effects of reform can differ much in closed and open economies. Reform increases the steady state stock in a closed economy and thereby improves the country's welfare. In an open economy, reform might harm a country in both the short and the long run. Again, this ambiguity arises because with trade there are two relevant distortions, the market failure in each country.

5 A Special Model

This section describes a model that is simple enough to yield a closed-form solution, and yet general enough to admit the kinds of possibilities illustrated in the previous section. This model helps to identify the circumstances that make one outcome or another more likely. Here we describe the static model, explain the equilibrium, and discuss some of the special features of the model. Section 7 embeds this model in a dynamic setting.

5.1 Description of the Model

Figure 18.2 shows a flow chart of the autarkic economy. The top arrow shows that the stock and flow in the previous period (Z_{-1}, E_{-1}) affect the current stock, Z. We return to this dynamic relation in Section 7, and here consider only the static aspects of the model. There are two goods: the "subsistence good" F, which we choose as the

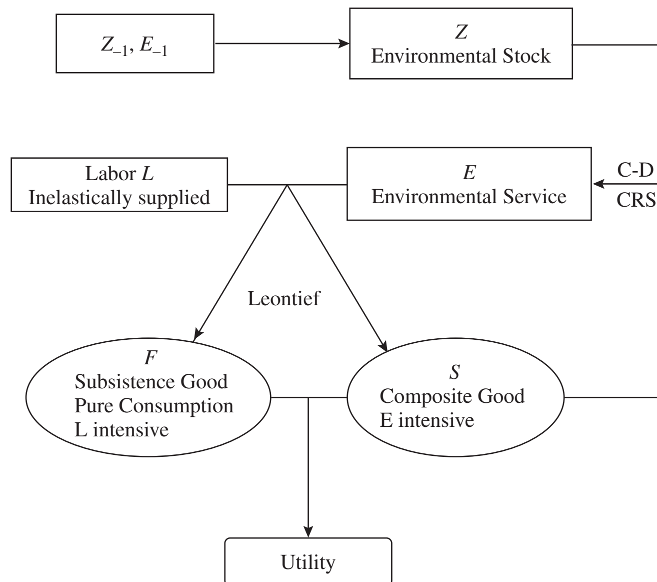

Figure 18.2 Structure of the economy

numeraire (i.e., its price is normalized to one), and the "composite good" S that has price p.

For clarity, we can think of good F as food, good S as steel, Z as the stock of water in lakes, and E as the flow of water used in production. Food is a pure consumption good, and its income elasticity falls as income increases. Steel can be consumed (in the form of cars) or used for pipes to transport water from lakes to agricultural and steel production. A poor economy uses steel only for pipes, but a richer economy also consumes cars. Water in lakes is a renewable resource that provides benefits only as a factor of production. A larger stock of water means that supplies are closer to production, so less steel is needed to obtain usable water.

Food and steel are competitively produced using labor L and environmental services E with Leontief technologies:

$$F^{\mathrm{p}} = \min\left\{\frac{E_{\mathrm{F}}}{a_{\mathrm{F}}}, \frac{L_{\mathrm{F}}}{b_{\mathrm{F}}}\right\}; \quad S^{\mathrm{p}} = \min\left\{\frac{E_{\mathrm{S}}}{a_{\mathrm{S}}}, \frac{L_{\mathrm{S}}}{b_{\mathrm{S}}}\right\}. \tag{1}$$

This technology implies, for example, that production of F^{p} units of food requires the use of at least $a_{\mathrm{F}}F^{\mathrm{p}}$ units of environmental services, and at least $b_{\mathrm{F}}F^{\mathrm{p}}$ units of labor. We assume that S is relatively environment-intensive, which means that, compared with food, it requires relatively more E to produce: $a_{\mathrm{S}}/a_{\mathrm{F}} > b_{\mathrm{S}}/b_{\mathrm{F}}$. Our results do not change if we reverse this inequality.

The representative consumer attempts to purchase F^* units of F. If her income, y, is less than F^*, she spends everything on good F, receiving utility y (equal to the consumption of F). If her income exceeds F^*, she buys F^* units of good F and $(y - F^*)/p$ units of S, resulting in utility $F^* + (y - F^*)/p$. These preferences provide a simple way to describe a situation where the share of income spent on food falls as income rises, for income above the critical level $y = F^*$. We assume that the representative consumer's income exceeds this critical level.

The supply of labor is exogenously fixed at $\bar{L}$. Environmental services, E, are "extracted" from the environmental stock Z using S^{e} units of good S with a decreasing returns to scale technology. Larger stocks decrease the costs of producing E. Let p^{e} be the price of E. We assume that the equilibrium supply function for environmental services is

$$E = \frac{\delta Z p^{\mathrm{e}}}{p}. \tag{2}$$

This relation can be derived from the assumptions that (i) the aggregate production function for environmental services is $E = (S^{\mathrm{e}}Z)^{0.5}$ and (ii) n (a fixed number) producers share total output, with each producer's share proportional to his share of inputs. The non-cooperative Nash equilibrium to this game implies equation (2), with $\delta = 1 - 1/(2n)$. For $n > 1 (\delta > 0.5)$ – as we assume – the value of marginal product of S in the extraction sector is less than the price of steel. The environmental sector absorbs too much of the input S. Production is inefficient, and the degree of inefficiency increases with δ. We refer to δZ as the *apparent stock* of this economy. A larger distortion or a larger physical stock both increase the apparent stock and therefore increase extraction.

The assumption that income exceeds F^* implies that the consumption of F is fixed at F^*. In this case, the economy's welfare is measured by the consumption of S, which equals the production of S minus the amount used in the extraction industry S^e (and the net export in the case of trade).

The two economies, North and South, are identical except for their values of δ and (possibly) their stock levels. We assume that $\delta_S > \delta_N$, so the environmental distortion is worse in South. When the countries trade we assume that both countries produce both goods, which implies that factor prices (the prices of E and L) are equal across the countries.

We use the following conventions:

Definition 1. Environmental reform *in country i means a reduction in* δ_i. Harmonization *of environmental policies means a reduction in* δ_S/δ_N. Upward harmonization *means a reduction in* δ_S/δ_N *caused by a decrease in* δ_S. Downward harmonization *means a reduction in* δ_S/δ_N *caused by an increase in* δ_N.

Downward harmonization is consistent with a "race to the bottom," while upward harmonization is consistent with a "race to the top."

5.2 Description of the Equilibria

Our assumptions about preferences and technology imply that there is a simple relation, depending on δ, between the stock level Z and the equilibrium extraction level of E. Figure 18.3 graphs this relation for two levels of distortion: $\delta_1 > \delta_2$. Given δ, the equilibrium level of extraction increases with the stock, for low stock levels. When the stock exceeds a critical level – denoted $Z^c(\delta)$, a decreasing function of δ – further increases in the stock have no effect on the equilibrium extraction of E.

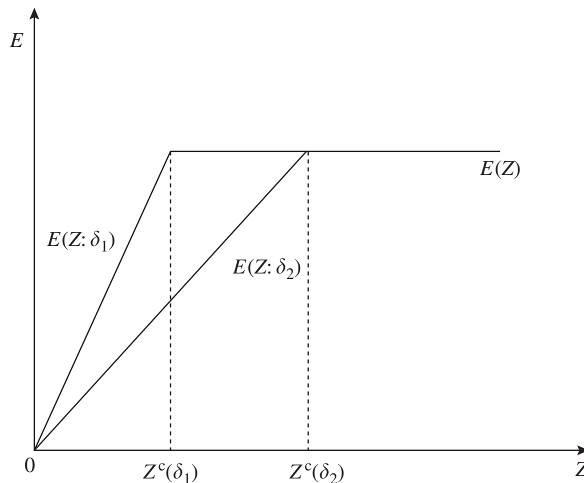

Figure 18.3 The extraction function: $\delta_1 > \delta_2$

The reason for this result is that the fixed labor supply becomes fully employed when $Z > Z^c(\delta)$. Our assumptions about preferences and technology, together with our assumption that the economy is rich enough to produce F^*, mean that a fixed amount of labor is used in sector F. Consequently, there is a fixed amount of labor available for Sector S. Since labor and environmental services are used in fixed proportions, the demand for environmental services is also fixed when labor is fully employed. Then further increase in the stock does not raise the extraction level, although it reduces the costs of extraction.

If the stock of Z is sufficiently large ($Z \geqslant Z^c(\delta)$), the equilibrium supply does not vary with the stock of the resource. Thus, for $Z > Z^c(\delta)$, reform affects neither the extraction of E nor the flow of welfare in autarky. If $Z < Z^c(\delta)$, reform reduces E and raises the critical level Z^c.

Under free trade, we assume that both countries are incompletely specialized (in equilibrium), so factor prices are equal. Consequently, labor is unemployed (and its price is zero) either in both countries or in neither country. Since the equilibrium depends on stocks in the two regions, the assumption of incomplete specialization restricts our analysis to a certain region of the state space – a region where the two stocks are not extremely different. This restriction greatly reduces the number of types of static equilibria that we need to study. Unemployment (in both countries under autarky and trade) corresponds to a region where both stocks are fairly small (but still large enough to support consumption of F^*); full employment (in both countries under autarky and trade) corresponds to a region where both stocks are quite large.

When labor is unemployed, there is only one constraining factor of production, E, resulting in the standard Ricardian model. In view of the assumption that countries have the same technology, the autarkic and free trade equilibria are identical. In this case, the aggregate supply of E is the same under free trade and autarky. This supply is increasing in both δ_S and δ_N: reform in either country reduces the supply of E in that country.

If labor is fully employed, the technologies and utility function imply that the total amount of E used in the world production of F and S is fixed. In addition, aggregate E under trade equals the sum of the autarky full employment levels of E in the two countries. However, the distribution of this aggregate level depends on the apparent stocks, and thus on the property rights. Using (2) and the assumption that factor prices are equal, we have $E_S/E_N = \delta_S Z_S/\delta_N Z_N$. Reform in either country affects extraction levels in both countries, but not aggregate extraction. However, reforms in both countries that leave the relative property rights δ_S/δ_N unchanged do not affect extraction in either country.

The distinction between "real" and "apparent" comparative advantage is straightforward in this model. The countries have the same technologies and preferences, that is, they are the same except for δ and Z. The country with higher δZ has a lower autarkic price of S and therefore has an apparent comparative advantage in S (and exports S). However, the social opportunity cost of S is lower in the country that has the higher value of Z. For example, South's opportunity cost of producing the resource-intensive good is lower than North's if and only if $Z_N < Z_S$, but its autarkic price of S is lower than North's if and only if $\delta_N Z_N < \delta_S Z_S$. Thus, for $Z_S < Z_N < (\delta_S/\delta_N)Z_S$, South has an "apparent" but not a "real" comparative advantage in the production of S.

5.3 Discussion of the Model

Our assumptions about preferences and technology make this model tractable. Here we discuss whether those assumptions are reasonable, and describe their important implications.

Any general equilibrium model that admits a closed-form solution requires simple functions for preferences and technology. The Cobb–Douglas functional form is an obvious alternative to our choices. This function implies that consumers always spend the same fraction of their income on a particular commodity, regardless of price and income; the elasticity of demand with respect to income is a constant, unity. In addition, a particular input always accounts for the same fraction of a firm's costs, regardless of prices and output; the elasticity of substitution is constant with the Cobb–Douglas form. In contrast, our assumptions about utility and technology imply that there is no substitutability in either consumption or production. Neither extreme assumption – zero or constant positive elasticity of substitution – is realistic. Our utility function implies that the elasticity of demand for the subsistence good falls from unity to zero as income crosses a threshold. Although this discontinuity is not plausible, the fact that the elasticity falls is reasonable.

The lack of substitutability in our model leads to a piece-wise linear extraction function, and this makes it possible to obtain some clear results when we introduce dynamics. This extraction function implies that under autarky the environmental distortion has no effect on the equilibrium when environmental stocks are large, but it has a significant effect when stocks are small. The implication that the distortion is less important to an economy when stocks are large is both reasonable and important. The discontinuity that occurs as the stock passes a threshold is not particularly plausible, but it does help to emphasize one effect of trade. When the environmental stocks in both countries are large, the distortion matters under trade, but not under autarky.

Some environmentalists are concerned that trade, by increasing aggregate world production, will increase environmental pressures. In our model, trade has no effect on aggregate production, or on the aggregate pressure on the environment when stocks are large. It does, however, affect how this aggregate environmental demand is allocated over countries. Standard trade models emphasize that in moving from autarky to trade, production is reallocated over countries. Our model shows how this reallocation leads to both short- and long-run environmental consequences.

6 Short-run Effects of Trade and Reform

Here we discuss the short-run effects of a movement from autarky to trade, and the short-run effects of environmental reform under autarky and under trade. Our welfare comparisons use the utility function described in the previous section. In particular, a country's welfare increases with the consumption of commodity S; aggregate (i.e., world) welfare increases with the aggregate consumption of S.

6.1 The Effects of Trade

The short-run effect of trade on resource extraction and welfare depends on the stock levels in both countries. Figure 18.4 divides the (Z_N, Z_S) plane into six regions, bounded by the horizontal and vertical lines at the critical levels Z_S^c and Z_N^c and by the line labeled FEL (the full employment line).[1] FEL is defined as the set of points that satisfy $\delta_N Z_N + \delta_S Z_S = \delta_N Z_N^c + \delta_S Z_S^c$. In the free trade equilibrium, labor is fully employed if the stock levels lie above this line.

For the purpose of describing the six regions identified in Figure 18.4, the reader should ignore the 45° line and the line labeled NTL. For stocks in Region I, labor is fully employed in both economies under trade and autarky. In Region IV labor is unemployed in both economies under trade and autarky. In Regions II or VI labor is unemployed in one economy under autarky, but fully employed in both economies under trade. In Regions III and V labor is unemployed in both economies under trade, but is fully employed in one economy under autarky. We focus on Region I, but the analysis can be extended to other regions.

The line labeled NTL (the "No Trade Line"), defined as $\delta_N Z_N = \delta_S Z_S$, shows the stock levels at which the autarkic relative prices are equal in the two countries (i.e., the countries have the same apparent stocks). For these stock levels, no trade occurs. For stocks in Region I below this line, North has the apparent comparative advantage in, and exports, commodity *S*; above this line, South exports *S*. However, North has the real comparative advantage in commodity *S* if and only if stocks lie below the 45° line ($Z_S = Z_N$). For stocks within the cone defined by NTL and the 45° line in Figure 18.4, the apparent and real comparative advantages are reversed. For stocks in this cone, the "wrong" country exports commodity *S*, and trade lowers aggregate welfare.

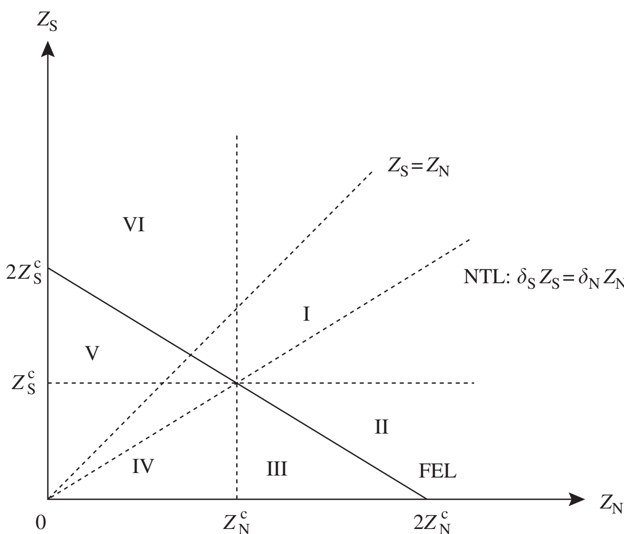

Figure 18.4 Possibilities of going from autarky to trade

Even outside this cone, trade can reduce aggregate welfare due to the inefficient volume of trade. For example, for stocks slightly above the 45° line, the two countries are similar in their endowments. Allowing them to trade creates small welfare gains if $\delta_S = \delta_N$. But because $\delta_S > \delta_N$, trade exacerbates the negative effects of Southern distortion, as now South produces more S. The negative welfare impacts in this case exceed the positive effects. However, aggregate welfare increases under trade if the Southern stock is sufficiently high, that is, when the stocks in Figure 18.4 lie far above the 45° line. In this case, the countries are sufficiently different in their endowments that the gain from trade can overcome the loss from exacerbating South's distortion.

Trade affects the welfare of individual countries differently. The S importing country always gains from trade. This country exercises its comparative advantage and enjoys the standard benefits of trade; in addition, its imports reduce domestic extraction of E, ameliorating the environmental distortion. For the country that exports S, trade increases extraction of E and exacerbates the environmental distortion. Welfare in the S exporting country increases if and only if the benefits of exercising its comparative advantage outweigh the costs resulting from the more serious environmental distortion (the worsened allocation of factors of production). For stocks in Region I near NTL, a country's apparent comparative advantage is negligible, but the welfare effect caused by the reallocation of demand for environmental services is significant. Thus, a sufficient condition for the S exporting country to have lower static welfare under trade is that the stock levels lie close to the NTL line.

6.2 Environmental Reform in Open and Closed Economies

We noted that under autarky, environmental reform affects the economy only when $Z < Z^c(\delta)$. In this case, reform reduces the supply of E and improves welfare. If stocks in both countries are small (e.g., in Region IV), the autarkic relative prices are equal in the two countries and they have no incentive to trade. In this case, reform has the same effects under trade as under autarky.

If stocks are large (e.g., in Region I), upward or downward harmonization (smaller δ_S/δ_N) improves aggregate welfare equally, while an increase in δ_S/δ_N decreases aggregate welfare. Reform narrows the gap between the price of steel and the value of marginal product of steel in the reforming country's E sector, leading to an efficiency gain there. However, this reform induces the trading partner to produce more E, and lowers efficiency there. The efficiency loss under trade arises from the *difference* in (or the ratio of) the distortions in North and South, rather than their absolute magnitudes; therefore, harmonization in either direction improves world welfare.

For example, Northern reform decreases its production of E, increases South's production of E, and leaves unchanged the aggregate supply of E, F ($=2F^*$) and S. Since $\delta_S > \delta_N$, the value of marginal product of S (in the production of the input E) is lower in South. The price of steel is equal in the two countries under trade, so the gap between the value of marginal product of steel and the price of steel is greater in South than in North. As North reforms, the total production of E is unchanged, but production has become less efficient because more E is produced in South; the total amount

of S used in the E sectors has increased. Consequently, the amount of S available for consumption has decreased, and aggregate world welfare falls. Reform in South, on the other hand, increases world welfare.

Consider now an individual country's welfare when stocks are in Region I. We already described the effect of reform in one country on the efficiency of the E sectors in both countries. Environmental reform also increases the equilibrium price of steel, the commodity that uses E relatively intensively. Therefore, reform in the steel exporting country necessarily improves its welfare and harms its trading partner. The direct effect on production efficiency and the terms of trade effect are positive for the reforming exporter, and both are negative for the trading partner. When an importing country reforms its domestic environmental distortion, the welfare effects for both countries are ambiguous. In both countries the production efficiency effect and the terms of trade effect have the opposite signs.

Equal-proportionate reform in the two countries that leaves relative distortions unchanged does not alter E_i or aggregate welfare. However, this reform reduces world apparent resource stocks and raises the price of food. Equal-proportionate reform thus benefits the exporter of the environmentally intensive good and harms the importer.

7 Long-run Effects of Trade and Reform

To study the long-run effects, we assume the evolution of Z is given by the logistic growth function $dZ_i/dt = \dot{Z}_i = \eta Z_i - \gamma Z_i^2 - E_i$, for $i = \mathrm{N}, \mathrm{S}$. In keeping with our assumption that tastes and technology are the same in the two countries, we assume that this function is also the same. The parameter γ captures the congestion effect of the stock; $\gamma > 0$ insures that Z is bounded. The non-congested growth rate of the environment, η, provides a measure of environmental resilience. When η is large, the environmental stock recovers quickly from low levels. We associate a large value of η with a resilient environment, and a small value of η with a fragile environment. The carrying capacity of the stock is η/γ and the stock that maximizes the sustainable yield is $\eta/2\gamma$. The level of E_i is the amount of extraction (the flow of environmental services) at a point in time.

The dynamic equilibria for autarky and trade are sequences of the static equilibria studied in the previous section. The static equilibria describe what happens under autarky and trade given the stock levels. As the countries extract E, the stock levels evolve, changing the associated autarky and trade equilibria. In a steady state, the stock levels and the associated autarky and trade equilibria remain constant. (The economies approach the steady state asymptotically, i.e., as $t \to \infty$.) In studying the long-run effects, we investigate how the steady states change as the economies move from autarky to free trade, or as they improve property rights (reduce δ).

Under both autarky and trade, there could be unique or multiple steady states, depending on the magnitude of η relative to critical values of that parameter. Figure 18.5 describes the possible steady states in our special model, and is therefore a specialization of Figure 18.1. The figure shows two kinds of steady states: a low steady state is one that is less than $\eta/2\gamma$ (the stock level that produces the maximum sustainable yield), and a high steady state is one above $\eta/2\gamma$. Given the equilibrium extraction function $E(Z)$,

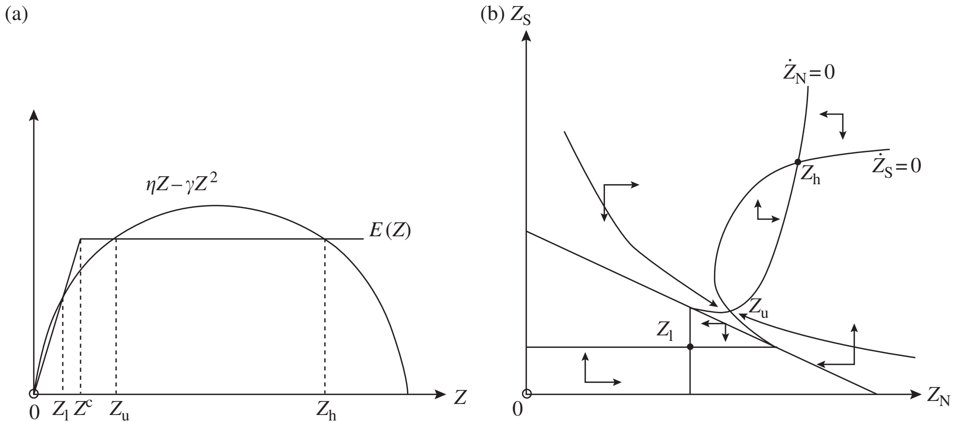

Figure 18.5 Resource dynamics with multiple steady states. (a) Autarky: $\hat{\eta}^a < \eta < \eta^{*a}$ and (b) Trade: $\hat{\eta} < \eta < \eta^*$

which is independent of η, the economy is more likely to move to the high steady state when η is higher (i.e., as the natural growth curve rises in Figure 18.5(a)).

In particular, under autarky there exist two critical values of η, denoted as η^{LA} and η^{HA}, with $\eta^{\mathrm{LA}} < \eta^{\mathrm{HA}}$. There is a unique low steady state, Z_{l}, if $\eta \leqslant \eta^{\mathrm{LA}}$, a unique high steady state Z_{h} if $\eta \geqslant \eta^{\mathrm{HA}}$, and both low and high steady states (together with an unstable steady state Z_{u}) if $\eta^{\mathrm{LA}} < \eta < \eta^{\mathrm{HA}}$. (The superscripts L and H denote critical values for low and high steady states, and the superscript A denotes autarky.) Similarly, under trade, the critical values are: η^{LT} such that there is a unique low steady state $Z_{\mathrm{l}} = (Z_{\mathrm{Nl}}, Z_{\mathrm{Sl}})$ if $\eta \leqslant \eta^{\mathrm{LT}}$; η^{HT}, such that there is a unique high steady state $Z_{\mathrm{h}} = (Z_{\mathrm{Nh}}, Z_{\mathrm{Sh}})$ if $\eta \geqslant \eta^{\mathrm{HT}}$. If $\eta^{\mathrm{LT}} < \eta < \eta^{\mathrm{HT}}$ there exist both low and high steady states and an unstable steady state $Z_{\mathrm{u}} = (Z_{\mathrm{Nu}}, Z_{\mathrm{Su}})$. (The superscript T denotes trade.)

The value of η is determined by physical systems, and does not respond to policy changes. However, a policy change such as a move from autarky to trade, or an environmental reform that decreases δ, can change some of the critical values of η defined above. Policy reform can therefore change the relation between the fixed value of η and the endogenous critical levels of η. Such a change can cause a *qualitative* change in the steady states, that is, it can change the types of steady state (high or low) that the economies reach.

Some policy changes have different types of effects on different critical values. This fact explains why the actual value of η is important in determining the long-run effects of the policy. This relation is essential to the rest of the analysis in this chapter. Suppose, for example, that a particular policy change increases the low critical value and decreases the high critical value of η. If the actual value of η is low (i.e., if the environment is fragile), this policy change is likely to be harmful, because it makes it “more likely” that the economy reaches a low steady state. The potentially beneficial effect of the reduction in the high critical value is irrelevant, since the actual value of η is not close to that critical value. In contrast, the increase in the low critical value is likely to be important, since the economy is close to that value. On the other hand,

if the actual value of η is high (i.e., the economy is resilient), this policy change is likely to be beneficial, because it makes it "more likely" that the economy reaches a high steady state. Thus, the long-run effect of the policy change depends on whether the actual value of η, or the resilience of the environment, is close to the low or the high critical value.

Figure 18.5 shows examples of autarkic (panel a) and trade (panel b) equilibria. We show only examples of the situation where there are multiple steady states under autarky and free trade. Under autarky, the economy moves to Z_h if the initial stock $Z_0 > Z_u$, and it moves to Z_l if $Z_0 < Z_u$. Under trade, the economies move to Z_h if their initial stock combination is above the saddle path – the solid curve going through the unstable steady state Z_u; for initial stocks below the saddle path, the economies move to the low steady state. In panel (b), the curves labeled $\dot{Z}_S = 0$ and $\dot{Z}_N = 0$ divide the plane into several regions. The arrows in the figure show the direction of motion of the stocks in each of these regions.

7.1 The Long-run Effects of Trade

To identify the critical values for a particular country under autarky, we use subscripts, η_S^{iA} and η_N^{iA} for South and North, $i = \mathrm{L, H}$. The critical value above which only a high steady state exists is different in North and South: $\eta_N^{HA} < \eta_S^{HA}$. North is more likely to avoid the low steady state given its less severe distortion. The autarkic critical values below which only a low steady state exists is independent of the domestic distortion. These two critical values are therefore the same for North and South: $\eta_N^{LA} = \eta_S^{LA} \equiv \eta^{LA}$. Figure 18.6 shows the two possible rankings of the critical values of η defined above. With both of these rankings, two important inequalities hold.

The first of these is $\eta^{LT} > \eta^{LA}$. This inequality means that there are values of η for which both autarkic economies might reach a high steady state (depending on the initial level of the stock), but under trade these two economies necessarily reach a low steady state (regardless of initial stock levels). If $\eta^{LT} > \eta > \eta^{LA}$ and the initial environmental stocks in both countries are high ($Z_i > Z_i^c$ for $i = \mathrm{N, S}$), the countries would reach high steady states under autarky. When they begin to trade, Southern extraction increases, leading to a decline in its stock. Due to the low growth rate, Southern stocks are unable to recover. Eventually, North begins to extract more to compensate for low Southern extraction. In the process it drives its stocks to a low level, and both countries move to

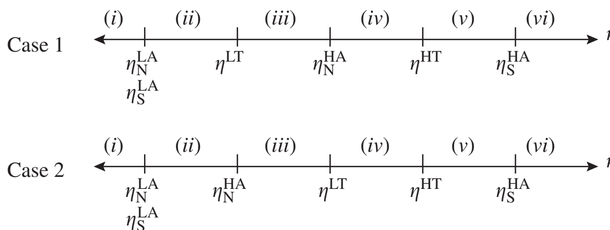

Figure 18.6 Possible ranking schemes of critical η values

the low steady state. This circumstance illustrates the outcome environmentalists fear. Trade causes the two countries to drag each other down.

The other important inequality is $\eta_N^{HA} < \eta^{HT} < \eta_S^{HA}$. This inequality is consistent with two possibilities. First, if $\eta_N^{HA} < \eta < \eta^{HT}$, then North certainly reaches a high steady state under autarky, but it might reach a low steady state under trade. This outcome is similar to the situation described in the previous paragraph. The second, and more optimistic case is where $\eta^{HT} < \eta < \eta_S^{HA}$. In this situation, there is a unique high steady state under trade. Under autarky, South reaches a low steady state if its initial stock is low (if $Z_S < Z_S^c$). For example, suppose Southern resource stocks are slightly below Z_S^c. After trade begins, Northern stocks will eventually be large enough (regardless of their initial condition) so that North exports the resource-intensive good. Southern stocks recover and eventually reach their high steady state. At some point South begins to export the resource-intensive good. The relevant comparison for South is between a high and a low steady state, and South is better off at the former. The relevant comparison for North is between autarkic and free trade high steady states, and we can show that North prefers the latter. North does well by doing good. In this scenario, trade causes North to "pull up" South, after which South returns the favor.

These scenarios illustrate the range of possible outcomes in moving from autarky to trade. The resilience of the environment, measured by the growth parameter η, is one determinant of which outcome occurs. Our model implies that "environmental resilience" (a high growth rate) reduces the likelihood of the scenario in which trade exacerbates market failures and lowers welfare in both regions in the long run (even if it might improve welfare in the short run). The level of Southern stocks at the time of the policy change is the other factor that determines whether trade is Pareto inferior or superior in the long run. The scenario in which trade lowers long-run welfare requires large Southern stocks, and the scenario in which it improves long-run welfare requires small stocks. Thus, trade may offer a helping hand to the South, especially when South experiences severe environmental problems. A resource-poor South may benefit most from international trade, contrary to the popular arguments of environmentalists.

7.2 Long-run Effects of Environmental Reform

In this section we consider separately the long-run effects of environmental reform under autarky and under trade. We then summarize the differing effects of reform in the short and the long run, and under free trade and autarky.

7.2.1 Long-run Effects of Reform under Autarky

Our assumptions about the utility and production functions imply that η^{LA} is independent of the environmental distortion; thus $d\eta^{LA}/d\delta = 0$. As a country's distortion δ increases, it is more likely that it reaches a low steady state, so $d\eta^{HA}/d\delta > 0$. These conclusions imply that if $\eta < \eta^{LA}$, the economy moves to the low steady state; reform under autarky cannot enable the economy to escape and move to a high steady state. Similarly, if $\eta > \eta^{HA}$, the economy reaches the high steady state with or without reform (since reform reduces η^{HA}). Reform has no long-run effects.

For the intermediate case, $\eta^{\mathrm{LA}} < \eta < \eta^{\mathrm{HA}}$, the magnitude of the reform is important. If the reform is large, the reduction in η^{HA} can be great enough to reverse the inequality $\eta < \eta^{\mathrm{HA}}$; in this case, reform removes the low steady state as a possibility. This large reform qualitatively improves the long-run welfare for small initial stocks, since the stock after reform approaches a high rather than a low steady state. But for large initial stocks, the stock would have approached a high steady state even without reform, so the reform does not have any qualitative long-run effects.

7.2.2 LONG-RUN EFFECTS UNDER TRADE

Recall from Section 5 that under free trade, the extraction levels of the two countries at the high steady state depends only on the *relative* distortions of the two countries, $\delta_{\mathrm{S}}/\delta_{\mathrm{N}}$. Consequently, the critical value η^{LT} also depends only on $\delta_{\mathrm{S}}/\delta_{\mathrm{N}}$, rather than the absolute values of the two distortions. Since differences between the two countries' distortions lead to inefficient allocation of the extraction efforts, it is less likely that they will reach the low steady state when they harmonize their environmental distortions. That is, η^{LT} decreases as the two distortions become less different:

$$\frac{d\eta^{\mathrm{LT}}}{d(\delta_{\mathrm{S}}/\delta_{\mathrm{N}})} > 0. \tag{3}$$

Harmonization of policies, achieved by either an improvement in Southern standards, or a deterioration in Northern standards, reduces $\delta_{\mathrm{S}}/\delta_{\mathrm{N}}$ and thus η^{LT}. If $\eta < \eta^{\mathrm{LT}}(\delta_{\mathrm{S}}/\delta_{\mathrm{N}})$ prior to harmonization, there is only a low steady state. Harmonization can reverse this inequality, thereby creating a high steady state. If the initial stocks, $\boldsymbol{Z}_0$, are sufficiently large, harmonization then causes the economies to move toward the high steady state. In this case, harmonization benefits both North and South in the long run. Here, harmonization shifts production of the resource-intensive good away from South, possibly altering the nations' apparent comparative advantage and reversing the direction of trade for a time. The lower level of exploitation enables South's stocks to recover. In the long run, South exports steel and North's stocks also recover.

Unilateral reform in North, which represents a movement away from harmonization, increases η^{LT}. Suppose, for example, that prereform $\eta > \eta^{\mathrm{LT}}$ and $\boldsymbol{Z}_0$ lies above the convergent saddle path through $\boldsymbol{Z}_{\mathrm{u}}$; in this case the economies move toward $\boldsymbol{Z}_{\mathrm{h}}$. If after Northern reform, $\eta < \eta^{\mathrm{LT}}$, the economies approach the low steady state $\boldsymbol{Z}_{\mathrm{l}}$. Northern reform harms both countries in the long run.

The critical value η^{HT} depends on both the relative and absolute values of δ_{S} and δ_{N}. Southern reform decreases both the absolute distortion in South and its distortion relative to North. This reform has an unambiguously beneficial long-run effect, since it reduces the range of values of η under which the low steady states exist (i.e., it increases the range of values for which the unique steady state is high):

$$\frac{d\eta^{\mathrm{HT}}}{d\delta_{\mathrm{S}}} > 0. \tag{4}$$

Southern reform may cause the regime to change from the case of multiple steady states to the case with only a high steady state. If this occurs, Southern reform causes the

economies to move to a high steady state even if they were previously trapped at a low steady state.

The long-run effect of Northern reform is ambiguous because it increases the relative distortions but decreases an absolute distortion. The effect of this reform on η^{HT} depends on which of the two influences is stronger. In particular, it depends on the initial difference between δ_S and δ_N and some parameters indicating the "severity of the environmental problem," such as the congestion parameter γ. If the resource is not extremely congested (i.e., if γ is low), the absolute effect of Northern reform always dominates the relative effect. In this case, Northern reform decreases the critical value η^{HT}. If, on the other hand, the environment is very congested, then either the absolute or relative effect may dominate. If the difference between the property rights is large, the relative effect dominates, and Northern reform increases the critical value η^{HT}. If the difference between the distortions is small, the absolute effect dominates, and Northern reform decreases the critical value η^{HT}.

The fact that upward harmonization (through reducing δ_S) certainly decreases η^{HT}, but downward harmonization (through increasing δ_N) may increase η^{HT}, argues in favor of upward rather than downward harmonization. This argument is based on the long-run effects of reform.

7.2.3 The different effects of reform

The dynamic and static effects of reform differ under free trade. If stocks are large enough to support full employment of labor (so that "trade matters") and if factor prices are equalized, instantaneous aggregate welfare depends only on relative distortions. Harmonization, whether achieved by upgrading Southern standards or degrading Northern standards, has the same instantaneous effect on aggregate welfare. In the long run, however, absolute as well as relative levels of standards are important. Harmonization upward is more likely than harmonization downward to increase long-run stocks and welfare.

The dynamic effects of reform depend on the trade regime. In a closed economy, reform does not alter the critical value η^{LA}, below which only a low steady state exists. If a country under autarky is trapped in a steady state with low environmental stocks, technology and preferences determine its destiny. Environmental reform cannot lead to high steady state stocks. In contrast, if open economies are trapped in a low steady state, harmonization of environmental policies (which reduces η^{LT} may enable them to escape to a high steady state.

In a closed economy, reform always reduces the critical value η^{HA}, above which only a high steady state exists. Therefore, if both the high and low steady states exist in an autarkic economy, reform might eliminate the low steady state, ensuring that the economy reaches the high steady state. With open economies, reform in the more distorted economy reduces the critical value η^{HT} above which only a high steady state exists. Reform in the less distorted economy, which reduces harmonization, may increase this critical value. Therefore, in open economies, reform against harmonization can either increase or decrease the danger that environmental stocks move to a low steady state.

In addition to illustrating these possibilities, the model shows how the plausibility of a particular outcome is related to the intrinsic growth rate of the environment.

If η is small, we consider the environment to be "fragile," in the sense that it regenerates slowly. In this situation, changes in η^{LT} and η^{LA} are more important than are changes in η^{HT} or η^{HA} (since the actual value of η is more likely to be close to the first pair of critical values). Thus, "fragility" of the environment makes it more likely that harmonization – even if achieved by lower standards in North – improves the environment and welfare in both the short and the long run. If, on the other hand, the environment is "resilient" (η is close to η^{HT} or η^{HA}), unilateral reform in North may lead to long-run improvements, at the cost of short-run welfare losses.

8 Discussions and Conclusion

We have examined the effects of trade liberalization and environmental reform in a simple dynamic general equilibrium model in which two countries, North and South, differ in two respects. These two differences cause autarkic equilibrium prices to differ, and therefore lead to trade. The first difference is institutional. The two countries have different degrees of environmental distortion or different property rights for an environmental resource. The distortion is greater in South, which tends to give South an apparent comparative advantage in the environment-intensive good. The second difference is that the environmental stocks, and thus the costs of producing the environment-intensive good, may differ in the two countries. This difference – unlike the institutional difference – is a standard motivation for trade. At a point in time the stock levels are predetermined. However, environmental stocks evolve endogenously, depending on how heavily the resource is exploited. Over time, factor endowments are endogenous and the pattern of trade can change.

We used the model to address two related but distinct issues, about which environmentalists and economists have a history of disagreement. The first concerns the welfare effects of trade liberalization, which we modeled by comparing autarky and free trade. The second concerns the effect of unilateral environmental reform and of harmonization. We defined harmonization as a change that makes the distortions in the two countries more similar; harmonization is consistent with a reduced environmental distortion in South, or an increased distortion in North.

We emphasized at the outset that in a world with more than one distortion, the welfare effect of reducing a single distortion is likely to be ambiguous. We also emphasized that trade restrictions are not the appropriate policy for dealing with environmental problems, and we cautioned against using the results of our analysis to promote an anti-trade or anti-property rights agenda. Our model – like any analytic model – ignores too many factors that we know to be important, to lead to unequivocal policy conclusions. For example, we ignore difference in tastes and technology and factor endowments other than the environmental stocks; we ignore other dynamic variables, such as technology and capital stocks; we assume that the level of the environmental distortion is exogenous, when in fact it might respond to changes in income or other endogenous (e.g., political) variables. Finally, we used specific functional forms that limit substitutability in consumption and production.

On balance, we think that our model has more of an "environmentalist's slant" than an "economist's slant." The model emphasizes an environmentally related market

failure and ignores many of the reasons for trade that are standard in economic models. Therefore, it is not surprising that the model supports certain positions favored by environmentalists, such as the beneficial effects (at least in the short run) of harmonization of environmental policies; the model assumes away the reasons why economists are skeptical of the benefits of harmonization. In evaluating this conclusion, the reader has to balance the relative importance of the different reasons for trade in the real world – the reasons that our model emphasizes and those that it ignores.

Despite this environmental slant, the model has implications that are familiar to economists. For example, it explains why, under some conditions, "downward harmonization" is as beneficial as "upward harmonization". This extreme result is due to the assumed lack of substitutability in our model, but it would survive (in a weaker form) if there was some substitutability. Also, the model demonstrates that trade is certainly beneficial in the short run if the stock differences are sufficiently large relative to the institutional differences.

However, we believe that the main contribution of our analysis is the light it sheds on differences between short- and long-run policy effects, and on the manner in which these effects depend on the "fragility" of the environment. There are many ways that one could measure environmental fragility; in our setting, it is natural to consider a resource that regenerates slowly as being fragile.

Our model contains the possibility of multiple steady states; with this possibility, the long-run equilibrium depends on the initial environmental stocks. A policy reform – liberalized trade or a reduced environmental distortion – can change the nature of the steady state to which the economy converges, leading to long-run qualitative effects that are very different from the short-run effects.

Although the relation between long-run policy effects and the fragility of the environment is rather complex (even in our simple model), one clear and important tendency emerges. When the environment is fragile, it is more likely that a movement from autarky to free trade reduces long-run welfare, and it is more likely that policy harmonization is beneficial under trade. Similarly, when the environment is resilient, it is more likely that the movement from autarky to trade improves welfare; it is also more likely that environmental reform – even when this increases the difference between environmental policies – improves welfare.

The model's implications are consistent with the beliefs of the "typical" environmentalist or economist; the implications depend on whether the environment is fragile or resilient. Since it seems likely that environmentalists (relative to economists) are more apt to view the environment as being fragile, the model provides a partial explanation for the differing views of the two groups. Even such a partial explanation is useful in promoting dialog between the groups. The groups' differing policy conclusions are not necessarily the result of irreconcilable worldviews, but a consequence of different beliefs about unknown facts of the world – facts about which we can hope to learn more.

There can, of course, be no definitive answer to questions such as "Is trade good for the environment?" or "Is policy harmonization a desirable goal?". However, if opposing groups are less dogmatic about their answers to these kinds of questions, there is greater potential for designing policies that liberalize trade and protect the environment.

Note

1 This figure is a simplification of the full taxonomy of equilibria. It ignores the constraints needed to insure that $F \geqslant F^*$ and that both economies are diversified in the free-trade equilibria.

References

Antweiler, Werner, Brian R. Copeland, and M. Scott Taylor 2001: Is free trade good for the environment? *American Economic Review*, 91, 877–908.

Bhagwati, Jagdish 1993: Trade and the environment: the false conflict? In Durwood Zaelke, Paul Orbuch, and Robert Houseman (eds.), *Trade and the Environment*. Washington, D.C.: Island Press, chapter 13, 159–90.

Bhagwati, Jagdish, Arvind Panagariya, and T. N. Srinivasan 1998: *Lectures on International Trade*, second edn. Cambridge, MA: MIT Press.

Brander, James A. and M. Scott Taylor 1997a: International trade and open access renewable resources: the small open economy case. *Canadian Journal of Economics*, 30, 526–52.

Brander, James A. and M. Scott Taylor 1997b: International trade between consumer and conservationist countries. *Resource and Energy Economics*, 19, 267–97.

Brander, James A. and M. Scott Taylor 1998a: Open access renewable resources: trade and trade policy in a two-country model. *Journal of International Economics*, 44, 181–209.

Brander, James A. and M. Scott Taylor 1998b: The simple economics of Easter Island: a Ricardo–Malthus model of renewable resource use. *American Economic Review*, 88, 119–38.

Chichilnisky, Graciela 1993: North–south trade and the dynamics of renewable resources. *Structural Change and Economic Dynamics*, 4, 219–48.

Chichilnisky, Graciela 1994: North–south trade and the global environment. *American Economic Review*, 15, 851–74.

Copeland, Brian and M. Scott Taylor 1994: North–south trade and the environment. *Quarterly Journal of Economics*, 3, 755–87.

Copeland, Brian and M. Scott Taylor 1995: Trade and transboundary pollution. *American Economic Review*, 85, 716–37.

Daly, Herman E. 1993: Problems with free trade: neoclassical and steady-state perspectives. In Durwood Zaelke, Paul Orbuch, and Robert Houseman (eds.), *Trade and the Environment*. Washington, D.C.: Island Press, chapter 12, 147–58.

Diamond, Jared 1999: *Guns, Germs, and Steel: The Fates of Human Societies*. New York, NY: Norton.

FAO 1994: *FAO Yearbook*. Rome: UN FAO.

Karp, Larry, Jinhua Zhao, and Sandeep Sacheti 2003: The uncertain benefits of environmental reform in open economies. *Journal of Environmental Economics and Management*, 45, 246–264.

Karp, Larry, Sandeep Sacheti, and Jinhua Zhao 2001: Common ground between free-traders and environmentalists. *International Economic Review*, 42, 617–47.

Mani, M. and D. Wheeler 1998: In search of pollution heavens? Dirty industry in the world economy 1960–1995. *Journal of Environment and Development*.

McDowell, Mark A. 1989: Development and the environment in ASEAN. *Pacific Affairs*, 62, 307–29.

Pomfret, John 2001: China's globalizing economy ravages border forests. *Washington Post*, March 26, A19.

19

Do Bilateral Tax Treaties Promote Foreign Direct Investment?

Bruce A. Blonigen and Ronald B. Davies

Chapter Outline

We explore the impact of bilateral tax treaties on foreign direct investment using data from OECD countries over the period 1982–92. We find that recent treaty formation does not promote new investment, contrary to the common expectation. For certain specifications we find that treaty formation may actually reduce investment as predicted by arguments suggesting treaties are intended to reduce tax evasion rather than promote foreign investment.

1 Introduction

Economists have long been concerned over the effect of taxation on foreign direct investment (FDI). A plethora of studies have examined whether and to what extent FDI responds to tax incentives. While the specific results vary, the general consensus is that firms do indeed respond to a variety of tax policies.[1] This issue is of primary concern to economists because it can result in an inefficient allocation of investment across countries. As governments use their tax policies to affect the rates of return on capital, provide public goods, or simply capture part of the profits that would otherwise be repatriated to other countries, this can divert investment from its most productive locations. One potential method of relieving this inefficiency is a bilateral tax treaty

on FDI. These treaties adjust the tax environment for investment between treaty partners by specifying the applicable tax base, the withholding taxes that can be applied, and other measures affecting the taxation of FDI. Worldwide, over 2,000 of these treaties are in force and they govern the taxation of the large majority of FDI (Radaelli, 1997).

Since treaties indicate cooperative taxation by treaty partners, many economists assume that treaties increase investment. However, it is by no means certain that treaties do so. In particular, since treaties can reduce tax avoidance and other tax-saving strategies by firms, they might actually have a dampening effect on FDI. Furthermore, legal researchers have questioned whether FDI promotion is even a primary goal of treaty formation. In light of these conflicting arguments, the effect of tax treaties on FDI is an open question. This chapter presents empirical results on treaty formation by members of the Organization for Economic Cooperation and Development (OECD). Our findings suggest that, at least for recent treaties, treaty formation most likely does not increase FDI activity between members and may in fact decrease it. Thus, our results stand at odds with the FDI promotion rationale for treaty formation.

Before delving into our data analysis, it is instructive to consider the conflicting opinions about tax treaties. The FDI promotion view of treaties springs from the idea of a tax distortion to investment. To illustrate this concept, consider the following simple model of FDI. An investor in the parent country has an amount of capital K that she can invest either in her parent country, a host country, or both. The amount of her exported capital is Z, thus Z represents the level of FDI. The rate of return in the parent country is a decreasing function of capital invested in the parent country $r(K - Z)$, while rate of return in the host country is a decreasing function of FDI $r^*(Z)$. Efficiency requires that capital be allocated between the two countries such that the rates of return are equal (or that all the capital is invested in the high-return country). However, the investor is not concerned with the total return on capital, but rather her share of it. In other words, the investor bases her decisions on the after-tax rate of return in each country, not the gross rate of return. As a result, if t and t^* are the marginal effective tax rates on investment in the parent country and in the host, the investor will compare $(1-t)r(K-Z)$ with $(1-t^*)r^*(Z)$.[2] Unless the marginal effective tax rates are equal, the equilibrium distribution of capital will be inefficient because efficiency calls for capital export neutrality in which only real, that is non-tax, variations govern capital flows. Therefore, as governments use their tax policies, it can quite easily lead to differential tax rates across locations and inefficient levels of FDI. Note that differential tax rates do not require different statutory taxes since the effective tax depends on many factors, including the definition of the tax base, accelerated depreciation rules, research and development tax credits, double taxation relief, and the like. It is generally believed that the effective tax rate against host investment exceeds that on investment in the parent country, implying that FDI is inefficiently low. Hines (1988) and Wilson (1993) support this concept by illustrating how typical parent country tax policies such as accelerated depreciation for domestic investments effectively result in tax rates that discriminate against overseas investment. Tax treaties can help to alleviate this problem by coordinating tax policies between treaty partners. This idea is mirrored in the introduction of the OECD's model tax treaty, which states that a primary goal of treaty formation is "removing the obstacles that double taxation presents," thus reducing its "harmful effects on the exchange of goods and services and

movements of capital, technology, and persons" (OECD, 1997, p. I-1).[3] If treaties do indeed reduce these tax barriers to FDI, one would expect that FDI activity would rise after a treaty is enforced.

Although individual treaties include a wide range of specific investment incentives, overall treaties reduce the barriers to FDI in two ways.[4] First, by harmonizing the tax definitions and the tax jurisdictions of treaty partners, a treaty can reduce the double taxation of investment. For example, income is typically taxed in a host country when it is generated through a permanent establishment. However, without a treaty each country can form its own definition of a permanent establishment. If this definition differs between countries, it can lead to double taxation of overseas profits (Hamada, 1966). Janeba (1996) discusses how these definition differences can then result in inefficient capital flows. This idea is confirmed empirically by Hines (1988), who finds that the 1986 Tax Reform Act, which revised US tax definitions, led to an increase in US outbound investment. Since treaties standardize tax definitions and jurisdictions (often by matching them to those provided by the OECD's model tax treaty (OECD, 1997)), they have a similar potential to increase FDI.

Second, tax treaties affect the actual statutory taxation of multinationals. They do so through the rules affecting double taxation relief and the withholding taxes levied on repatriations by FDI. Following the OECD model treaty guidelines, most tax treaties specify that both countries must either exempt foreign-earned profits from domestic taxation or offer foreign tax credits when calculating the domestic tax bill.[5] Although most countries already offer their investors credits or exemptions, certain treaties do alter the relief method applied by one or both treaty partners.[6] In addition to the provisions for double taxation relief, treaties usually reduce maximum allowable withholding taxes on three types of remitted income: dividend payments, interest payments, and royalty payments.[7] Some treaties lower these withholding rates to as low as zero. Most treaties specify that the same maximum rates apply to both treaty partners.[8] If these reductions in the withholding tax reduce the tax burden on overseas investment, equivalent to reducing t^* above, this should increase FDI.[9] Note that even though withholding tax rates fall under a treaty, this does not imply that tax receipts from inbound investment must decline. Since withholding taxes can be tailored to the specific investment from a treaty partner, it may be possible to set tax rates that encourage tax-sensitive inbound investment and actually raise total tax receipts.[10] In addition, treaties are accompanied by improved information exchange between partner governments.[11] Because of this, tax evasion may fall under a treaty, leading to increased tax revenue.[12]

Combining these arguments, it is easy to understand the common expectation that tax treaties serve to increase the amount of FDI activity between treaty partners. Nevertheless, there exist several economic and legal arguments that suggest treaties may have no effect on FDI. For instance, Dagan (2000) asserts that the use of treaties to promote FDI is "a myth" (p. 939). She claims that since a parent country could unilaterally adjust its tax policy to eliminate distortions caused by differing parent and host country tax policies, promotion of efficiency plays little role in treaty formation. Instead, Dagan suggests that treaties are intended to reduce administration costs, reduce tax evasion, and to extract tax concessions from treaty partners. Radaelli (1997) also dismisses the double taxation objective in favor of the view that tax treaties are geared toward reducing tax evasion. Furthermore, there is the concern that tax treaties arise due

to lobbying efforts by profit-seeking investors. If this is the case, then treaties may be geared toward maximizing investor profits rather than promoting efficient investment.[13] In addition to uncertainties about government objectives in treaty formation, it is by no means clear that firms' investment activities will necessarily respond to reductions in withholding tax rates. Hartman (1985) and Sinn (1993) argue that withholding taxes are irrelevant for expanding multinational firms since it is cheaper to expand an overseas affiliate through retained earnings than through repatriated and reexported funds. This is because retained earnings avoid the withholding taxes applied to repatriated funds. As a result, they suggest that only mature, nonexpanding foreign affiliates will repatriate earnings, implying no effect of withholding taxes on FDI.

Finally, it is even possible that tax treaties may actually increase the tax barriers for certain kinds of investment. As noted above, through information exchange treaties can reduce the firm's ability to engage in transfer pricing. This is the practice by which, through manipulation of the price of goods traded between their various subsidiaries, firms can shift profits to low tax locations and minimize their global tax revenues. As argued by Casson (1979), transfer pricing provides firms with an incentive to invest in those low tax locations in order to shield profits from taxes.[14] Since treaties streamline and promote the exchange of tax information by governments, this reduces firms' ability to avoid taxes through misrepresentation of costs. As a result, treaties may reduce the incentive to engage in investment for tax minimization reasons, leading to a decrease in FDI activity. In addition, recent tax treaties have sought to eliminate treaty shopping. Treaty shopping is a practice in which investments are funneled through a treaty country by a third nation for the purpose of avoiding or reducing taxes. According to Radaelli (1997), concerns over treaty shopping have been a primary focus of many new treaties and have prompted the US to renegotiate many of its older treaties. In addition, certain so-called "tax haven" countries (particularly Aruba, Malta, and the Netherlands-Antilles) have seen several of their treaties cancelled due to perceived insufficient efforts to prevent treaty shopping. If a treaty is revised to close this possibility, then this could easily reduce the investment activity between treaty partners as third nation investors choose to simply send their capital directly to the ultimate host. While there are many variations in the regulations that address treaty shopping, the most common rules restrict treaty benefits if more than 50 percent of a corporation's stock is held by a third, non-treaty country's residents (Doernberg, 1997).

With these conflicting arguments in mind, it is by no means certain that bilateral treaties will increase the amount of FDI between partner countries. This issue is of importance because of the sizable costs to treaty formation. For most countries, tax treaties are like any other international agreement in that they must be drawn up and then ratified by the appropriate governing body. Ratification is no mean feat since the treaty cannot conflict with other national policies. For the US, this is not a severe problem since tax treaties are federal instruments and thus supersede state or local laws.[15] This is not always true elsewhere. Anders (1997), for example, gives an in-depth discussion of the difficulties the US–German treaty faces since it may violate the antidiscrimination rules of the European Union. Even ignoring such possible roadblocks, the simple conflicts of bilateral negotiation require much time and effort. Conflicts over the terms of the treaty can lead to failure during the development stage or even during ratification. For example, the US–Cyprus treaty required three attempts before it was finally ratified in 1988. Thus,

in light of these large costs, it is important that we gauge the potential gains from treaty formation. We seek to provide one possible measure of the gains from treaty formation by estimating the impact of treaty formation on the FDI of OECD members.[16]

Using recent models developed by Carr, Markusen and Maskus (2001) and Markusen and Maskus (2001), we test whether or not treaty formation by OECD members is associated with changes in FDI activity. Using data from 1982 to 1992, we examine the behavior of FDI stocks and FDI flows. Initial results indicate that treaties seem to increase FDI. However, this result is suspect because the sample includes many older treaties, which were enacted well before our data series begins. Therefore, we also split our sample into "old" and "new" treaties. Here, we find that new treaty formation is not significantly correlated with FDI activity. In fact, when we restrict ourselves to just those countries that enacted a treaty during our sample, we find that FDI stocks are significantly decreased after treaty formation. These results are consistent with Dagan (2000) and Radaelli (1997) who claim that recent treaties are not geared toward the promotion of FDI but rather toward reductions in tax evasion. While these results do not mean that treaties cannot be used to increase FDI they do suggest that the FDI promotion argument seems suspect.

The chapter proceeds as follows. In Section 2, we present our empirical methodology. We discuss our data in Section 3. Section 4 presents our results and Section 5 concludes.

2 Empirical Framework for Statistical Analysis

In order to examine how treaties affect FDI, we require a framework that describes the determinants of FDI. Over the past couple of decades, James Markusen and coauthors have developed formal general equilibrium theory models of multinational enterprise (MNE) activity.[17] These theory models lead to predictions of equilibrium FDI activity across bilateral pairs of countries in terms of a few observable factors. Recently, Carr et al. (2001) and Markusen and Maskus (2001) tested the "knowledge–capital" version of the Markusen model of MNE activity using data on affiliate sales of US firms in other countries and foreign affiliate sales in the United States over the period 1986–94. They find substantial empirical fit of the knowledge–capital model to the data. The Carr, Markusen, Maskus (CMM) empirical framework is specified by the following equation:

$$\mathrm{FDI}_{ij} = f(\mathrm{SUMGDP}_{ij}, \mathrm{GDPDIFSQ}_{ij}, \mathrm{SKDIFF}_{ij}, \mathrm{SKDIFF}_{ij}{}^{*}\mathrm{GDPDIFF}_{ij}, (\mathrm{SKDIFF}_{ij})^{2*}\mathrm{T_OPEN}_{j}, \mathrm{Z}_{ij}, \mathrm{TREATY}_{ij}) \quad (1)$$

The dependent variable, FDI_{ij} is a measure of FDI activity from a parent country (i) to a host country (j). CMM use affiliate sales data as their "FDI" measure, though Blonigen and Davies (2000, 2001) find that the CMM empirical framework also fits US FDI stock (and to some extent, FDI flow) data reasonably well.

The first five independent variables on the right-hand side of (1) are the variables specific to the CMM framework. The first two terms are relatively straightforward, with SUMGDP defined as the sum of the two countries' real gross domestic products (GDPs), and GDPDIFSQ defined as the squared difference between the two countries' real GDP.

There is an expected positive correlation between SUMGDP and FDI activity and an expected negative correlation between GDPDIFSQ and FDI activity. The intuition is that with some positive level of trade frictions, larger and more similar-sized markets better support the higher fixed costs associated with setting up production across countries (versus exporting) and lead to greater FDI activity.

The third, fourth, fifth terms on the right-hand side of (1) are more complicated terms related to differences in the two countries' relative endowments of skilled labor to unskilled labor. Skilled labor is important in the theory of MNE activity because firms that have firm-specific assets (which are developed using skilled labor) will have the greatest incentives to expand their operations across borders. In some theories of MNE activity, MNE firms are additionally attracted to place production activities in less-skilled countries because production activity is less skill-intensive than the headquarter activities conducted in the parent country. These are generally models that emphasize vertical reasons for MNE activity; that is, to locate its production activities in countries with lower wages for low-skill-intensive activities. These models would suggest that FDI activity should be stronger between countries with greater differences in skilled labor abundance. Other models that emphasize MNEs' motives to locate production in large markets to avoid tariffs and transport costs (horizontal MNE models) often predict that FDI activity should increase with greater similarities in skilled labor abundance. In (1) the SKDIFF variable is the parent country's skilled labor abundance minus the host country's skilled labor abundance. Given the vertical MNE motives that exist in the CMM knowledge–capital model, CMM predict a positive correlation between SKDIFF with FDI activity from the parent to the host country. The fourth term is an interaction term between SKDIFF and GDPDIFF, the parent country's GDP minus the host country's GDP. CMM suggest that the knowledge–capital model predicts a negative correlation between this variable and FDI from the parent to the host country, since the theory predicts larger effects of SKDIFF on FDI when the parent country is a lot smaller than the host country. Finally, CMM include a fifth term that interacts the square of SKDIFF multiplied by trade openness in the host country. The predicted sign from their model on this term is positive.[18]

In addition to the variables just described, additional control variables (Z_{ij}) have typically been included in the CMM empirical framework. First, distance ($DISTANCE_{ij}$) is included to proxy for transport and other trade costs that may affect a firm's decision about whether to become an MNE. Second, trade openness for both the parent and host countries (T_OPEN_i and T_OPEN_j) affects the MNE's ability to trade intermediates and final goods, which then affects the location of MNEs. Greater openness in the host country should lower FDI activity because it lowers trade friction and makes exporting a relatively more attractive method of serving the host market than FDI. Greater openness in the parent country should increase FDI, since it makes it easier to ship goods back to the parent country from foreign affiliates. FDI openness of the host country (F_OPEN_j) proxies for the costs of setting up an MNE, with greater openness expected to increase FDI activity.

The last independent variable on the right-hand side of (1) is our focus variable: a measure of bilateral investment tax treaty activity. There are substantial measurement issues that determine how we define this variable. In particular, we can observe when countries make bilateral investment tax treaties with each other, but these treaties

certainly differ from each other along many dimensions, which are very difficult to quantify. In addition, the same treaty on paper can have vastly different consequences for different pairs of countries depending on the unilaterally adopted tax practices of the countries before entering the treaty. There is little that can be quantified across treaties with the exception of agreed-upon tax withholding rates. Unfortunately, these withholding rates are typically maximum allowable rates and there is evidence that countries sometimes set rates below these maximum allowable rates after the treaty, making them uninformative for our purposes of analysis (Blonigen and Davies, 2000, 2001). Because of these difficulties, we primarily measure tax treaty activity in this chapter as a binary variable taking the value of "1" if two countries have signed a bilateral tax treaty that was eventually entered into force and "0" if they do not. As a result, we will only be able to estimate an average total impact of tax treaties across our sample of countries.

While we only report results on the effect of bilateral tax treaties using the empirical CMM framework described above, we note that we get qualitatively similar results for the impact of tax treaties when using alternative empirical frameworks. This includes variations of the CMM framework found in Markusen and Maskus (1999) and a "gravity-type" framework that posits FDI activity as only a function of the size of the countries (proxied by real GDP) and distance. In addition, Blonigen et al. (2002) show that the CMM framework misspecifies the SKDIFF and GDPDIFF terms, which should be expressed in absolute values. While this correction strongly affects the implications of the relationship between skill differences and FDI activity, it has virtually zero quantitative effect on our estimates of treaty effects. Although results from these alternative empirical specifications have been omitted for space, they are available upon request from the authors.

3 Data

One of the major hurdles in the analysis of the effect of bilateral tax treaties on FDI is data constraints. As noted above, this begins with measuring the treaty activity. However, there are significant measurement issues with respect to the data on bilateral FDI activity as well. For this study we use OECD data on bilateral FDI stocks and flows, as reported by OECD-member countries.[19] Such data were not even compiled into a publicly available form until 1993 with the first annual OECD *International Direct Investment Statistics Yearbook*.[20] Since the data are collected from national sources in each country, there is substantial variation in coverage by country source and by year, and there is variation in measurement of FDI activity itself. Almost all countries report inflows and outflows for at least some select bilateral pairings, with about half also reporting measures of inward and outward stocks of FDI. The earliest data available begin in 1982. Table 19.1 provides further details on data coverage across OECD countries and years in our sample.

There are definite comparability issues of FDI measures across countries. For example, a number of OECD countries do not include reinvested earnings by firms in their measures of FDI. Countries can also differ in what percentage of foreign-owned shares of a firm is necessary for it to be classified as FDI rather than portfolio

Table 19.1 FDI data coverage by OECD country

Country	Type of FDI measure	Number of partner countries reported[a]	Years covered[b]
Australia	Inflows	18	1982–
	Outflows	17	1982–
	Inbound stock	17	1982–
	Outbound stock	12	1982–
Austria	Inflows	11	1982–
	Outflows	13	1982–
	Inbound stock	9	1982, 1986–
	Outbound stock	11	1982, 1986–
Belgium–Luxembourg	Inflows	44	1982–
	Outflows	45	1982–
Canada	Inflows	3	1982–
	Outflows	3	1982–
	Inbound stock	29	1982–
	Outbound stock	30	1982–
Denmark	Inflows	20	1982–
	Outflows	22	1982–
Finland	Inflows	10	1982–
	Outflows	17	1982–
	Inbound stock	1	1989–
France	Inflows	23	1982–
	Outflows	23	1982–
	Inbound stock	43	1987–
	Outbound stock	34	1987–
Germany	Inflows	19	1982–
	Outflows	18	1982–
Greece	Inflows	14	1987–
Iceland	Inflows	0	1987–
Ireland	Inflows	0	1983–
Italy	Inflows	20	1982–
	Outflows	20	1982–
	Inbound stock	22	1985–
	Outbound stock	23	1985–
Japan	Inflows	9	1982–
	Outflows	36	1982–
	Inbound stock	9	1982–
	Outbound stock	37	1982–
Netherlands	Inflows	9	1982–
	Outflows	9	1982–
	Inbound stock	14	1984–
	Outbound stock	14	1984–
New Zealand	Inflows	4	1984–
	Outflows	4	1984–

Table 19.1 Continued

Country	Type of FDI measure	Number of partner countries reported[a]	Years covered[b]
Norway	Inflows	14	1986–
	Outflows	19	1987–
	Inbound stock	18	1987–
	Outbound stock	24	1988–
Portugal	Inflows	18	1982–
	Outflows	10	1982–
Spain	Inflows	20	1982–
	Outflows	18	1982–
Sweden	Inflows	15	1982–
	Outflows	36	1982–
Switzerland	Inflows	0	1983–
	Outflows	0	1986–
	Outbound stock	0	1986–
Turkey	Inflows	18	1982–
	Outflows	6	1989–
United Kingdom	Inflows	18	1982–
	Outflows	36	1984, 1987–
	Inbound stock	19	1982–
	Outbound stock	35	1984, 1987–
United States	Inflows	39	1982–
	Outflows	41	1982–
	Inbound stock	28	1982–
	Outbound stock	41	1982–

[a] *Figures for 1990.*
[b] *Not all reported countries are necessarily reported each year during range indicated.*
"0" indicates that data is available only by region and not for specific countries.

investment.[21] However, with only a couple of exceptions, we note that FDI definitions are fairly consistent for the same country over time. This is important because in our statistical analysis we will be able to use techniques that estimate relationships between variables using only the sample data variation within individual countries over time, avoiding cross-country data measurement consistency problems that arise for other statistical techniques.

Besides our FDI and tax treaty measures, our statistical analysis employs the additional CMM empirical framework control variables. Our data on real GDP, which is used to construct the SUMGDP and GDPDIFSQ variables, come from the well-known Penn World Tables, which are described by Summers and Heston (1991) and are available online at http://datacentre.chass.utoronto.ca:5680/pwt/. To construct our measure of differences in skilled labor endowments, we use World Bank data on the average country-level education attainment as a proxy for such endowments. Nehru and Dhareshwar (1993) provide further details on these data. DISTANCE was measured as the distance between capital cities as reported by the Bali Online

Corporation. This distance calculator can be found at http://www.indo.com. Trade openness measures of the two countries were obtained from the Penn World Tables and are defined as a country's total trade flows (exports plus imports) divided by its GDP. The F_OPEN_j was constructed in a similar manner as a country's total FDI flows divided by its GDP using the UN *World Investment Directory*. Details on this variable construction can be found in Blonigen and Davies (2001).

The Penn World Tables run only through 1992, while our FDI data begin in 1982 at the earliest. These data availability issues limit our sample and analysis to the period from 1982 through 1992.

Table 19.2 provides descriptive statistics of our dependent and independent variables. We have 3,276 observations on FDI outflows by OECD countries, and 2,235 observations on FDI outbound stock. Average annual bilateral FDI outflows are almost $284 million with substantial variation across observations (standard deviation of just more than $1 billion).[22] Average FDI outbound stock is $3.378 billion with a standard deviation of $8 billion.[23] About 77 percent of the sample observations are of bilateral country pairs with bilateral tax treaties in place. The majority of these (74 percent) are what we term "old" treaties that were in place before our sample begins, with 3 percent of the sample observations connected to "new" treaties that were enacted during our sample. We discuss the importance of this distinction for our statistical analysis in more detail in the next section.

Table 19.2 Descriptive statistics of variables

Variables	Number of observations	Mean	Standard deviation	Minimum	Maximum
Dependent variables					
FDI outbound stock	2,235	3,378.13	8,002.59	0.00	92,733.8
FDI outflows	3,276	283.82	1,056.61	−2,550.56	42,267.61
Independent variables					
$SUMGDP_{ij}$	3,276	1,259.53	1,260.96	61.79	6,449
$GDPDIFSQ_{ij}$	3,276	1,993,241	4,710,473	0.11	2.06e + 07
$SKDIFF_{ij}$	3,276	1.21	2.54	−6.65	8.1
$SKDIFF_{ij}*GDPDIFF_{ij}$	3,276	2,124.00	5,374.54	−7,494.69	31,011.47
$(SKDIFF_{ij})^2*T_OPEN_j$	3,276	475.35	821.85	0.00	8,850.6
$DISTANCE_{ij}$	3,276	5,695.07	5,129.36	174.00	19,007
T_OPEN_i	3,276	56.38	31.01	17.62	156.45
T_OPEN_j	3,276	68.22	55.26	8.96	386.23
F_OPEN_j	3,276	12.42	14.2	0.15	93.69
$TREATY_{ij}$	3,276	0.77	0.42	0.00	1.00
OLD $TREATY_{ij}$	3,276	0.74	0.44	0.00	1.00
NEW $TREATY_{ij}$	3,276	0.03	0.17	0.00	1.00

Note: *FDI outbound stock, FDI outflows, SUMGDP, and GDPDIFSQ are measured in millions of real US dollars. Distance is measured in miles between capital cities. Skill differences is measured in mean years of female and male education attainment. See text for variable definitions.*

4 Statistical Analysis and Results

Our statistical analysis will proceed in stages, beginning with a relatively naive estimation of (1), and then proceeding with more sophisticated estimation procedures to correct for potential statistical problems. We do this to show that it is not easy to estimate the true effect of tax treaties on FDI, primarily because many treaties by OECD countries were in place before our sample begins. Most of these "old" treaties are between OECD countries, with many beginning in the decade after World War II. If we get a positive correlation between our tax treaty variable and our dependent variable, FDI activity, it is not clear whether other unobservable characteristics of the tax treaty country pairings may be leading to both increased FDI activity and a tax treaty. This occurs because the tax treaty variable, a simple binary variable, will pick up any residual effects on FDI that are not measured by the other control regressors. This problem, known as simultaneity, makes identification of the treaty effect difficult to measure. For example, there are likely to be a number of underlying reasons beyond their bilateral tax treaty that explain why the United States and the United Kingdom have large FDI activity with each other. These other reasons, however, may not be observable. Because of this, statistical analysis will assign the influence of the unobserved factors to the observed existence of a treaty. As such, while statistical analysis would indicate a strong positive relationship between the existence of a treaty and FDI activity, one hesitates to say that the treaty *causes* the activity between the two countries.

However, our sample also includes a number of bilateral tax treaties completed by OECD countries after our sample data begins. These "new" treaties afford a much better opportunity to measure the impact of a tax treaty, as we have data on FDI activity both before and after the treaty is signed. Presuming that there are no other changes occurring at the time of the treaty that would affect FDI (besides those captured by our control regressors), we can estimate the effect of these new treaties on FDI more precisely by comparing the pre- and post-treaty information. Table 19.3 lists new treaties that were completed by OECD countries after the first year of our sample, 1983, through the last year of our sample, 1992.[24] As we show below, the estimated effects of old treaties versus new treaties on FDI activity are quite different. Given this discussion, we give much more weight to the credibility of our evidence for new treaties.

Column 1 of Table 19.4 provides statistical results when we use ordinary least squares (OLS) regression techniques to estimate (1) on our full sample of countries and years when FDI activity is measured as the parent country's FDI stock in the host country. Column 2 of Table 19.4 provides statistical results when we conduct the same procedure, but define our FDI activity measure as the parent country's FDI flow into the host country. With the exception of SKDIFF and the trade openness measures, the control variables have their predicted signs and are generally statistically significant in the FDI stock and FDI flow regressions. The SKDIFF variable is estimated with an incorrect negative sign in both regressions, but is not statistically significant at standard levels of statistical confidence. The general fit of the empirical framework to the data is better for the FDI stock regression, where the control variables explain about 34 percent of the variation in the dependent variable, as compared to the FDI flow regression, where

Table 19.3 New treaties by OECD countries from 1983–92

Country	Bilateral tax treaties (year of treaty in parentheses)
Australia	Italy (1983), Korea (1983), Norway (1983), Ireland (1984), Finland (1985), Austria (1987), China (1990), Thailand (1990), Hungary (1992), Poland (1992)
Austria	Thailand (1986), Australia (1987)
Belgium–Luxembourg[a]	Korea (1985), Turkey (1988)
Canada	Bangladesh (1983), Brazil (1985), Cyprus (1986), India (1986), China (1987), Kenya (1987), Poland (1989), Slovak Republic (1987)
Denmark	Cyprus (1984), Indonesia (1986), China (1987)
Finland	Australia (1985), New Zealand (1985), Turkey (1985), Thailand (1986), China (1987), Yugoslavia (1987), Indonesia (1988)
France	Bangladesh (1988), Trinidad and Tobago (1988), Nigeria (1991)
Germany	Philippines (1984), Turkey (1986), Indonesia (1988)
Greece	Hungary (1984), Switzerland (1984), Czechoslovakia (1987), Norway (1989)
Iceland	None
Ireland	Australia (1984), New Zealand (1987)
Italy	China (1987), India (1987)
Japan	Indonesia (1983)
Netherlands	Pakistan (1983), Romania (1983), China (1987), India (1989), Brazil (1991)
New Zealand	Norway (1983), Finland (1985), Ireland (1987), Indonesia (1988)
Norway	Australia (1983), New Zealand (1983), Yugoslavia (1984), China (1987), Pakistan (1987), Philippines (1988), Greece (1989)
Portugal	None
Spain	Yugoslavia (1983)
Sweden	Trinidad and Tobago (1985), China (1987), Cyprus (1989), Indonesia (1989), Turkey (1989)
Switzerland	Greece (1984), Egypt (1988), Indonesia (1989), China (1991)
Turkey	Finland (1985), Germany (1986), Belgium–Luxembourg (1988), Sweden (1989)
United Kingdom	India (1983), Thailand (1983), China (1985), India (1987)
United States	Barbados (1984), China (1985), Tunisia (1986), Cyprus (1988), Indonesia (1989), India (1990), Spain (1991)

[a] *Includes only treaties made by Belgium, not Luxembourg.*

only about 10 percent of the variation in the dependent variable is explained by the independent variables.

Both statistical regressions provide strong evidence for a positive effect of tax treaties on FDI activity. Everything else equal, the presence of a tax treaty means almost an extra \$2.5 billion of parent FDI stock in the host country versus a situation where there was no bilateral tax treaty. This is quite significant given an average of \$3.4 billion in FDI stock in the sample. The effect of treaties on FDI flows is estimated to be somewhat larger relative to its mean. Everything else equal, the presence of a tax treaty means an extra

Table 19.4 Estimated treaty effects on OECD outbound FDI stock and flows using OLS

	Empirical models and dependent variables			
	Ordinary least squares with treaty variable		**Ordinary least squares with separate old and new treaty variables**	
Regressors	**FDI stock 1**	**FDI flows 2**	**FDI stock 3**	**FDI flows 4**
Treaty variables				
TREATY	2,446.9**	233.81**		
	(0.00)	(0.00)		
NEW TREATY			−263.06	−20.586
			(0.74)	(0.86)
OLD TREATY			2,782.8**	257.81**
			(0.00)	(0.00)
CMM controls				
$SUMGDP_{ij}$	4.368**	0.426**	4.426**	0.431**
	(0.00)	(0.00)	(0.00)	(0.00)
$GDPDIFSQ_{ij}$	−0.0001	−0.00004**	−0.0001	−0.00004**
	(0.16)	(0.00)	(0.09)	(0.00)
$SKDIFF_{ij}$	2.948	−5.747	65.085	−1.017
	(0.97)	(0.51)	(0.41)	(0.91)
$SKDIFF_{ij}$* $GDPDIFF_{ij}$	−0.315**	−0.020**	−0.313**	−0.020**
	(0.00)	(0.00)	(0.00)	(0.00)
$(SKDIFF_{ij})^2$*T_OPEN_j	−0.697**	−0.049	−0.623**	−0.039
	(0.00)	(0.10)	(0.01)	(0.18)
$DISTANCE_{ij}$	−0.396**	−0.031**	−0.392**	−0.030**
	(0.00)	(0.00)	(0.00)	(0.00)
T_OPEN_i	−7.567	−0.43	−4.767	−0.335
	(0.42)	(0.54)	(0.61)	(0.63)
T_OPEN_j	−26.184**	−1.347*	−27.754**	−1.461**
	(0.00)	(0.02)	(0.00)	(0.01)
F_OPEN_j	220.58**	16.622**	226.06**	16.923**
	(0.00)	(0.00)	(0.00)	(0.00)
R-squared	0.34	0.11	0.34	0.11
F test	112.14**	38.31**	113.21**	35.45**
Sample size	2,235	3,276	2,235	3,276

Note: *P-values are in parentheses, with ** and * denoting statistical confidence levels at the 99 and 95 percent levels, respectively. P-values indicate statistical probability (in decimal form) that the true parameter value is zero (i.e., has no effect on the dependent variable). R-squared is the ratio of the variation in the dependent variable explained by the regressors. F test is a statistical test of whether we can reject the hypothesis that the coefficients of the regressors are jointly zero.*

$234 million of annual parent FDI flows into the host country, compared to an average annual flow of $284 million. Even setting issues of magnitude aside, these estimates suggest that we can be over 99 percent confident that these coefficient estimates are not zero.

As mentioned above, our sample includes both old treaties that occur before our sample period begins and new treaties that occur during our sample's time period. Because of this, it is not clear whether the strongly positive treaty effects in columns 1 and 2 are the result of the treaties themselves or some other unobserved factor. To deal with this, columns 3 and 4 of Table 19.4 provide results when we estimate the same empirical framework as in columns 1 and 2, but allow for separate effects for old and new treaties; that is, separate binary variables indicating whether an old or new treaty is present between the bilateral country pair in a given year or not. The estimated effects of old and new treaties are quite different for both of our specifications (FDI stock and FDI flows). Old treaties continue to show a positive effect on FDI activity, and that effect continues to be highly statistically significant and of slightly larger magnitude than our results in columns 1 and 2. Our estimates suggest that the presence of an old treaty increases outbound FDI stock by $2.8 billion (compared to $2.5 billion in column 1 estimates) and increases outbound FDI flows by $258 million (compared to $242 million in the column 2 estimates). In contrast, new treaties yield coefficient estimates that are *negative* in sign. However, given high standard errors relative to the small coefficient estimates, we cannot statistically reject the hypothesis that the new treaties have no impact on FDI activity.

As mentioned above, it is difficult to assign causation to the old treaty effects because they are in place before our sample period begins – the positive correlation may just suggest that countries that would naturally have high FDI activity due to other (unobserved) factors also make sure they have a bilateral tax treaty in place.[25] This is particularly worrisome since our (observed) control variables do not even account for half of the variation in the dependent variables. A common way to control for unobserved characteristics that affect the FDI activity between a bilateral pair of countries is to estimate what is known as a fixed-effects specification. This means that, in addition to our control variables, we also include a binary variable for each bilateral country pairing. These binary fixed-effect country-pair variables will estimate the aggregate effect of time-invariant characteristics (both observed and unobserved) that raise or lower the FDI activity for that bilateral pairing versus the average. In other words, it acts as an intercept term that is specific to the bilateral country pair. With fixed effects, our estimates then come only from the time series variation in our variables within each bilateral country pair, not the variation across bilateral country pairs. Note that now the effects of our time-invariant regressors, distance, and old treaties, will be subsumed into these fixed effects. In other words, we cannot identify the effects of old treaties separately from other time-invariant characteristics, such as historical conditions, that affect the overall FDI activity between a bilateral country pair. Because of this, we focus solely on the new treaties enacted at some point during the sample of years. Estimation using only the variation within bilateral country pairs provides an additional benefit since it can help to reduce the problems caused by different definitions of FDI activity across the countries in our sample, providing more credible estimates.

Columns 1 and 2 of Table 19.5 present our fixed-effects empirical estimation results. The fit of the equations increases dramatically: 94 percent of the variation in the dependent variable is explained in the FDI stock regression and 38 percent of the variation is explained in the FDI flow regression. Not surprisingly, statistical tests easily confirm that the fixed-effect variables are jointly statistically significant for explaining FDI activity. The main CMM framework control variables still perform well. In fact, the skilled-labor difference variable (SKDIFF), now has the correct sign and is statistically significant. However, the trade openness measures generally have the wrong sign in these regressions.

The surprising result in columns 1 and 2 of Table 19.5 is that the estimated effect of the new treaties on FDI activity is now strongly negative. The estimated impact has increased substantially and the effects are much more statistically significant: A new treaty leads to a $2.6 billion decrease in FDI stock and $351 million decrease in annual FDI flows. In column 3, we introduce a lagged dependent variable to control for dynamic adjustment in our FDI stock equation as there is probably persistence in the FDI stock over time that can lead to statistical problems with our estimates. The coefficient on the lagged dependent variable is statistically significant, but does not seriously affect our other regressor coefficient estimates, including the new treaty variable.[26]

In summary, we estimate very different correlations between old treaties and FDI activity versus new treaties and FDI activity. Given the simultaneity concerns described above, it is difficult to assign much weight to our old treaty evidence for positive effects on FDI, even though it is the best we can do given the data constraints. Because consistent recording of FDI activity began much later than the enactment of these old treaties, it is virtually impossible to identify the effect of the treaties on FDI activity from other country-pair characteristics that might affect both the inherent FDI activity and the incentives to have treaties. In contrast, we are much more confident in our new treaty estimates from a statistical standpoint and these results yield a very surprising conclusion: new treaties are not promoting FDI activity and the evidence suggests that they may even be decreasing FDI activity.

While this new treaty result is surprising, it is confirmed by Blonigen and Davies (2001) using a completely different database on US-only FDI activity. The US data, collected by the US Bureau of Economic Activity (BEA), is arguably the best-measured and most comprehensive data on FDI activity of any country in the world. The BEA has data on US bilateral FDI activity (both inbound and outbound) stretching back as far as 1966 for FDI flows and stock and the early 1980s for affiliate sales.[27] Using this US database and US bilateral tax treaty activity, Blonigen and Davies (2001) also find evidence for substantial negative effects of new treaties on FDI activity. In an earlier version of the paper, Blonigen and Davies (2000), had similarly addressed the issue of old versus new treaties by estimating the US-only sample without observations connected with Canada, Japan, and European countries. They found that the remaining sample still displayed positive effects of tax treaties on FDI activity, but with a very long lag. However, this result was driven by the fact that the reduced sample still contained FDI activity with a few remaining old treaty partners, specifically Australia and New Zealand. Once *all* of the old treaty partners are eliminated from the sample,

Table 19.5 Estimated treaty effects on OECD outbound FDI stock and flows using fixed effects

	Empirical models and dependent variables: fixed effects with new treaty variable only		
Regressors	**FDI stock 1**	**FDI flows 2**	**FDI stock with lagged dependent variable 3**
Treaty variables			
NEW TREATY	−2,597.6**	−350.83	−2,212.3**
	(0.00)	(0.06)	(0.00)
Control regressors			
LAGGED FDI STOCK			0.308**
			(0.00)
$SUMGDP_{ij}$	9.527**	1.379**	7.053**
	(0.00)	(0.00)	(0.00)
$GDPDIFSQ_{ij}$	−0.0004**	−0.0001**	−0.0003**
	(0.00)	(0.00)	(0.00)
$SKDIFF_{ij}$	2,677.6**	67.116	1,331.6**
	(0.00)	(0.63)	(0.00)
$SKDIFF_{ij}$ * $GDPDIFF_{ij}$	−0.819**	−0.008	−0.540**
	(0.00)	(0.89)	(0.00)
$(SKDIFF_{ij})^2$*T_OPEN_j	1.029*	0.01	0.61
	(0.01)	(0.94)	(0.09)
T_OPEN_i	−1.924	−3.164	−2.145
	(0.90)	(0.50)	(0.89)
T_OPEN_j	−47.163**	−0.552	−28.957**
	(0.00)	(0.85)	(0.00)
F_OPEN_j	124.66**	7.596	79.207**
	(0.00)	(0.19)	(0.00)
R-squared	0.94	0.38	0.96
F test	95.46**	12.94**	192.87**
Sample size	2,235	3,276	2,041

Note: *P-values are in parentheses, with ** and * denoting statistical confidence levels at the 99 and 95 percent levels, respectively. P-values indicate statistical probability (in decimal form) that the true parameter value is zero (i.e., has no effect on the dependent variable). R-squared is the ratio of the variation in the dependent variable explained by the regressors. F test is a statistical test of whether we can reject the hypothesis that the coefficients of the regressors are jointly zero.*

as effectively done in Blonigen and Davies (2001), results are consistent with this chapter's results for OECD countries: There is no credible evidence in the data that tax treaties have significant positive effects on FDI activity.

5 Conclusion

The majority of economic and legal texts stress the intuitive notion that bilateral tax treaties should promote FDI activity. As discussed in the introduction, there are a number of reasons why this may not be true in theory. This chapter adds to this debate by providing some of the first evidence on the effect of bilateral tax treaties on FDI activity. Using OECD data we find that new treaty activity (during the 1983–92 period) suggests strong negative impacts on FDI. While we find a positive correlation in the case of much older treaties, we cannot weight this evidence very heavily as we cannot observe FDI activity before these treaties were in place. These results are consistent with previous work by Blonigen and Davies (2001) using only US data. Thus, in conjunction with this earlier work, our results cast doubt upon the FDI promotion rationale for treaty formation, which stands in contrast to the conventional wisdom among many economists and lawyers.

One possible reason for the non-promotion effect of treaties on FDI activity is that treaties reduce firms' abilities to evade taxes through transfer pricing or treaty shopping. Our data on aggregate FDI activity are not well suited to address whether these issues connected with firm-level behavior are behind the overall result. An additional possibility for non-promotion of FDI activity by new treaties is that treaties may increase investment uncertainty, at least in the short run. Since a new treaty has yet to be tested in the courts of the partner countries, it may actually increase the perceived risk of investment between treaty partners until the legal interpretation of the treaty has been resolved. Thus, in the short run, the treaty may lead to a reduction in FDI activity. Over the long run, however, this uncertainty will be resolved, clearing the way for the treaty to promote investment. However, when we alter our new treaty dummy variable to only take the value of "1" a year (or even two years) after the treaty was signed, we get similar negative and statistically significant effects of new treaties on FDI activity. This would argue that the uncertainty issue is not behind the effects we find unless it takes many years to resolve such uncertainty. On the other hand, our work with US data in Blonigen and Davies (2001) found evidence that the negative drop in FDI activity seemed to occur mainly at the time of the treaty. We leave these and other important issues for future work and hope that these results serve as a guidepost for continued exploration of the relationship between tax treaties and foreign investment.

Acknowledgments

We thank James Hartigan for comments and Sarah Lawson for excellent research assistance. All remaining errors are our own.

Notes

1 See Wilson (1999) and Gresik (2001) for surveys of both the theory and empirical economic literature regarding the effects of taxation on foreign investment.

2 Since we are examining the marginal capital allocation, the appropriate tax rate is the marginal rate. Identification of the marginal tax rate is notoriously difficult. Graham (1996) provides discussion of the various proxies researchers have used for the marginal tax on corporate income.
3 For an excellent discussion of the workings of the OECD model tax treaty, see Baker (1994).
4 For specifics of the treaties, see the treaties themselves as reprinted by Diamond and Diamond (1998).
5 Under a tax credit, the domestic tax bill is calculated by applying the standard domestic tax rate to the pre-foreign tax level of overseas profits. A credit against this amount is then applied up to the amount of the foreign taxes paid. If this credit exceeds the domestic tax liability, the firm is in an "excess credit" position and pays no additional taxes on these overseas profits. If the parent country tax bill is greater than the amount of the credit, the firm is in an "excess limit" position and pays the remaining amount to the parent country's government.
6 For instance, Belgium applies a reduced domestic tax rate to the foreign-earned profits of its residents. Under the US/Belgium treaty, however, income earned by Belgian firms in the United States is exempt from Belgian taxation.
7 Note that a lower foreign tax rate does not guarantee a reduction in the firm's total tax bill. When a parent country offers foreign tax credits, only firms in an excess credit position will benefit by a reduction in host taxes (Altshuler and Newlon, 1991). Thus, a decrease in the overseas tax rate may not improve capital flows.
8 One exception is the US treaty with Pakistan in which US firms receive no tax break from Pakistan while Pakistani firms do receive reduced tax rates from the United States.
9 Several researchers, including Altshuler and Newlon (1991), Hines (1992), Altshuler et al. (1995), and Grubert and Mutti (1996) have found that firms do respond to changes in withholding tax rates by changing both the timing and method of repatriation. However, this evidence suggests that treaties can affect the profitability of overseas investment, and does not necessarily imply that treaties will induce new FDI.
10 As derived in Bond and Samuelson (1989) among others, the tax revenue maximizing tax rate on inbound FDI is $1/(1+\epsilon)$, where ϵ is the elasticity of inbound FDI supply. If ϵ varies across countries, the revenue maximizing tax rate will differ across countries. Since a treaty allows a country to lower its tax rate, this can actually raise tax revenues if the current tax rate is greater than the optimal tax on FDI from that particular country.
11 In addition to the bilateral treaties, the OECD has established the multilateral Convention on Mutual Assistance in Tax Matters, which provides for information exchange even between members without bilateral treaties (OECD, 1989).
12 Devereux et al. (2001) find that, even though corporate tax rates for the United Kingdom appear to have declined since the mid-1960s, tax revenues from corporate income have not. They suggest that part of this result may stem from a broadening of the incomes classified as corporate.
13 According to Radaelli (1997), treaty formation by the United States is free of such business lobbying efforts.
14 Caves (1993) provides an explanation of transfer pricing. Graham and Krugman (1995) provide case studies of firms prosecuted by the United States for engaging in transfer pricing.
15 Generally, the only limitation US tax treaties place on state taxation is that a foreign corporation must be treated the same as a firm incorporated in another state (White, 1991).
16 Hines and Willard (1992) empirically examined the number of treaties a country signs as well as the tax concessions dictated by a particular treaty. However, they do not include

the amount of FDI activity as an explanatory variable nor do they consider the effect of taxation on FDI. The United Nations (1998) has studied the effects of bilateral treaties for the promotion and protection of FDI. These treaties are generally geared toward increasing investment in developing nations by ensuring a favorable political and economic climate and do not address tax concerns.

17 For recent examples of this work, see Markusen et al. (1996) or Markusen and Venables (1997).

18 In CMM (2001) they interact the skill difference term with host trade costs, the opposite of host trade openness. Thus, we expect the opposite sign on this coefficient than they do. See CMM (2001) for further details on the knowledge–capital model.

19 As mentioned in the section above, Carr et al. (2001) and Markusen and Maskus (2001) use US data on MNEs' foreign affiliate sales, which is a theoretically preferable measure of MNE activity. However, to our knowledge, there are very few countries that keep track of affiliate sales, and there is no comprehensive cross-country database of foreign affiliate sales activity, even for OECD countries.

20 These data are available in print form in these annual yearbooks or in electronic form on the OECD Statistical Compendium CD-ROM, available for purchase from the OECD.

21 IMF and OECD guidelines specify investment as FDI when acquired shares are 10 percent or higher of target firm's outstanding stock, which many of the countries follow or eventually adopted. Graham and Krugman (1995) find that the foreign parent of an MNE in the United States on average owns 77.5 percent of the affiliates equity, suggesting that this problem may not be overwhelming.

22 These figures are in 1995 US dollars.

23 We remind the reader that these are figures for those bilateral pairs reported by each OECD country, which are typically skewed toward only reporting the countries with which it has the largest FDI activity.

24 We note that because of missing data, some of the bilateral pairs completing treaties in Table 19.3, are not covered by our sample. More specifically, 19 of the new bilateral treaties in Table 19.3 are covered in our sample in at least one of the two possible directions.

25 There is a similar, though milder, concern that the new treaties occur because FDI activity is increasing or expected to increase between the countries due to unobserved factors. However, this would be expected to bias us toward finding a positive correlation between new treaties and FDI activity. In contrast, our results below find evidence for a *negative* correlation.

26 A lagged dependent variable is not statistically significant in the FDI flow specifications.

27 Of course, the disadvantage of the US data versus the OECD data is that its observations are all tied to one country, making it difficult to comprehend whether the results from such a sample generalize to the rest of the world.

References

Altshuler, Rosanne and T. Scott Newlon 1991: The effects of US tax policy on the income repatriation patterns of U.S. multinational corporations. National Bureau of Economic Research Discussion Paper Series, No. 571.

Altshuler, Rosanne, T. Scott Newlon, and William Randolph 1995: Do repatriation taxes matter? Evidence from the tax returns of U.S. multinationals. In Martin Feldstein, James Hines, Jr., and R. Glenn Hubbard (eds.), *Effects of Taxation on Multinational Corporations*, Chicago: University of Chicago Press, 253–72.

Anders, Dietmar 1997: The limitation on benefit clause of the U.S.–German tax treaty and its compatibility with European law. *Northwestern Journal of International Law and Business*, 18.

Baker, Philip 1994: *Double Taxation Conventions and International Tax Law: A Manual on the OECD Model Tax Convention on Income and on Capital of 1992*, 2nd edn. London: Sweet & Maxwell.

Blonigen, Bruce A. and Ronald B. Davies 2000: The effects of bilateral tax treaties on U.S. FDI activity. National Bureau of Economic Research Working Paper Series, No. 7929.

Blonigen, Bruce A. and Ronald B. Davies 2001: The effects of bilateral tax treaties on U.S. FDI activity. Manuscript.

Blonigen, Bruce A., Ronald B. Davies, and Keith Head 2002: Estimating the knowledge-capital model of the multinational enterprise: Comment. Manuscript.

Bond, Eric and Larry Samuelson 1989: Strategic behaviour and the rules for international taxation of capital. *Economic Journal*, 99, 1099–111.

Carr, David, James R. Markusen, and Keith E. Maskus 2001: Estimating the knowledge-capital model of the multinational enterprise. *American Economic Review*, 91, 693–708.

Casson, Mark 1979: *Alternatives to the Multinational Enterprise*. London: Macmillan Press.

Caves, Richard E. 1993: *Multinational Enterprise and Economic Analysis*, 2nd edn. New York: Cambridge Surveys of Economic Literature, Cambridge Press.

Dagan, Tsilly 2000: The tax treaties myth. *New York University Journal of International Law and Politics*, Summer, 939–96.

Devereux, Michael, Rachel Griffith, and Alexander Klemm 2001: Have taxes on mobile capital declined? Manuscript.

Diamond, Walter and Dorothy Diamond 1998: *International Tax Treaties of All Nations*. New York: Ocean Publications.

Doernberg, Richard 1997: *International Taxation in a Nutshell*, 3rd edn. Minnesota: West Publishing.

Graham, Edward and Paul Krugman 1995: *Foreign Direct Investment in the United States*, 3rd edn. Washington, D.C.: Institute for International Economics.

Graham, John R. 1996: Proxies for the corporate marginal tax rate. *Journal of Financial Economics*, 42, 187–221.

Gresik, Thomas A. 2001: The taxing tax of taxing transnationals. *Journal of Economic Literature*, 39, 800–38.

Grubert, Harry and John Mutti 1991: Taxes, tariffs and transfer pricing in multinational corporate decision making. *Review of Economics and Statistics*, 73, 285–93.

Hamada, Koichi 1966: Strategic aspects of taxation on foreign investment income. *Quarterly Journal of Economics*, 80, 361–75.

Hartman, David 1985: Tax policy and foreign direct investment. *Journal of Public Economics*, 26, 107–21.

Hines, James R. Jr. 1988: Taxation and U.S. multinational investment. In Lawrence Summers (ed.), *Tax Policy and the Economy, 2*. Boston: MIT Press, 33–61.

Hines, James R. Jr. 1992: Credit and deferral as international investment incentives. National Bureau of Economic Research Working Paper Series, No. 4191.

Hines, James R. and Kristen L. Willard 1992: Trick or treaty? Bargains and surprises in international tax agreements. Manuscript.

Hufbauer, Gary 1992: *U.S. Taxation of International Income: Blueprint for Reform*. Washington, D.C.: Institute for International Economics.

Janeba, Eckhard 1996: Foreign direct investment under oligopoly: profit shifting or profit capturing? *Journal of Public Economics*, 60, 423–45.

Markusen, James R. and Anthony J. Venables 1997: The role of multinational firms in the wage-gap debate. *Review of International Economics*, 5, 435–51.

Markusen, James R. and Keith E. Maskus 1999: Discriminating among alternative theories of the multinational enterprise. National Bureau of Economic Research Working Paper Series, No. 7164.

Markusen, James R. and Keith E. Maskus 2001: Multinational firms: reconciling theory and evidence. In Magnus Blomstrom and Linda S. Goldberg (eds.), *Topics in Empirical International Economics: A Festschrift in Honor of Robert E. Lipsey*. Chicago: University of Chicago Press for National Bureau of Economic Research, 71–95.

Markusen, James R, Anthony J. Venables, Denise Eby-Konan, and Kevin Honglin Zhang 1996: A unified treatment of horizontal direct investment, vertical direct investment, and the pattern of trade in goods and services. National Bureau of Economic Research Working Paper Series, No. 5696.

Nehru, Vikram and Ashok Dhareshwar 1993: A new database on physical capital stock: sources, methodology and results. *Rivista de Analisis Economico*, 8, 37–59.

Organization for Economic Cooperation and Development 1989: *Explanatory Report on the Convention on Mutual Administrative Assistance in Tax Matters*. Paris: OECD Committee on Fiscal Affairs.

Organization for Economic Cooperation and Development 1997: *Model Tax Convention on Income and on Capital*. Paris: OECD Committee on Fiscal Affairs.

Radaelli, Claudio M. 1997: *The Politics of Corporate Taxation in the European Union*. London: Routledge Research in European Public Policy.

Sinn, Hans-Werner 1993: Taxation and the birth of foreign subsidiaries. In Herberg Heberg and go Van Long (eds.), *Trade, Welfare, and Economic Policies, Essays in Honor of Murray C. Kemp*. Ann Arbor: University of Michigan Press, 325–52.

Summers, Robert and Alan Heston 1991: The Penn-World Table (Mark 5): an expanded set of international comparisons, 1950–1988. *Quarterly Journal of Economics*, 106, 327–68.

White, Fredrick 1991: The United States Perspective. *Tax Treaties and Local Taxes*, International Fiscal Association, 16, 15–24.

Wilson, G. Peter 1993: The role of taxes in location and sourcing decisions. In A. Giovannini, R. G. Hubbard, and Joel Slemrod (eds.), *Studies in International Taxation*. Chicago: University of Chicago Press.

Wilson, John D. 1999: Theories of tax competition. *National Tax Journal*, 52, 269–304.

United Nations 1998: *Bilateral Investment Treaties in the Mid-1990s*. New York: United Nations Conference on Trade and Development.

Index

Abidjan Convention for Co-operation in the Protection and Development of the Marine and Coastal Environment of the West and Central African Region 123, n. 29
Abrego, Lisandro E. 317
agreement and coalitions, concepts of 318
Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPS") 159, 161–2
Agreements
  on Agriculture 322
  Antidumping 84
  on Government Procurement 329, n. 3
  on Sanitary and Phytosanitary Measures (SPS) 6, 12, 22, 24
  on Trade in Civil Aircraft 329, n. 3
  on Trade-Related Aspects of Intellectual Property Rights *see* TRIPS
  Trade-Related Labor and Environment Rights Agreements 107
agriculture 118
aims-and-effects test 196, 198
allocative efficiency test 9
Altshuler, Rosanne 543
Alvarez, Jose E. 157
ambiguity about legal concept of free trade 7
  classical roots of 8–10
American Sugar Refining Company 341
Anders, Dietmar 529
Anderson, James E. 279, 288
  imperfect enforcement 283–4
  model of informal enforcement by Mafia 287
  model of trade and predation under anarchy 287
Anderson, Simon P. 394
Anis, Aslam H. 460
antidumping (AD) 358, 384–92
  activity, worldwide 363
  advantage over AT 384
  business cycle fluctuation 387
  club effect 380
  conditional investment subsidy 468
  and countervailing duty (AD/CVD) actions 74
  duties 73, 276, n. 50
  evolution to standard protectionist instrument 395
  facilitation trade liberalization 364
  game players employing tit-for-tat (TFT) strategies 360
  history and background 360–5
  increased use by developing countries 391
  initiations 388
  injury standard under 380, n. 1
  legislation, short history of 360

antidumping (AD) (*cont'd*)
measures
definitive process of substitution in place of tariffs 389
injury-based 459
national welfare clause 395
policies 384, 393–5
principle of targeting 468
process 341
effect on collusion 342
and regional trade agreements 391–2
schools of thought regarding 364
special protection 366
strategic or protectionist policy 364
strategic use for punishment and/or deterrence 366
TFT 380
theory 365–8
unfair trade, response to 365–6
use
"non-economic" factors behind 359
worldwide upsurge in 379
users
heaviest 364
new 359
traditional 359
versus antitrust, policies trade and competition policy 383
Antidumping Agreement 84
antidumping (AD) and antitrust (AT)
complementarity between 395
as having common target 396
interface between 393
policies, use of 383
antidumping (AD) cases
"big change" hypothesis 369
"big supplier" hypothesis 369
club hypothesis 369
database of all 368
economic incentives 369
preeminence of 359
TFT hypothesis 369
two-tier approach 397
Antidumping Code of the World Trade Organization (WTO) 468
during the Kennedy Round, first 391
see also World Trade Organization (WTO)
antidumping (AD) filing
behavior
four main hypotheses 368
trends in 361–4
estimation the strategic motives for 368–79
antidumping (AD) laws
enactment or drafting as condition for release of loans 369
geography of 386
injury criteria in the material injury decision in 468
and membership of GATT/WTO, countries with 385
for occurrence of predatory dumping 393
role in alleviating the injury 467
user of 364
worldwide implementation of 385
antitrust (AT)
effective instrument in dealing with anticompetitive behavior 396
focus on encouraging competition 396
instruments 384
measures, series of economic tests for justification 396
antitrust policy
in competition policy 396
in open economies
political economy considerations 347
price fixing and international cartels 333
optimal 344–8
Antweiler, Werner 502
ANZCERTA, relationship between Australia and New Zealand 259, 392
see also Australia
Appellate Body (of WTO) 49
in *Asbestos* 221
in Beef Hormones 34, n. 90–4, 226
Decisions
on *EEC–Asbestos* 225
on Hot-Rolled Steel 2001 84
on Lamb Meat, 2001 82
on Line Pipe, 2001 84
on Wheat Gluten, 2000 82

in European Communities – Regime for the Importation, Sale, and Distribution of Bananas 208
in Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef 208
report 269
in US – section 337 in Korean Beef 225
see also World Trade Organization
Argentina 364
antidumping in 363
Armington partial equilibrium models, CADIC and COMPAS as 79
Aruba as tax haven 529
Asia-Pacific Economic Cooperation (APEC) forum 392
Australia 385
countertrade deals 283
proposal on substantially all trade terms 257
Australia and New Zealand (ANZCERTA), relationship between 392
autarky 287–8
Axelrod, Robert 139, 142

Bagwell, Kyle 176, 366
on antidumping measures 364, 366
on impact of trade blocs 310, n. 16
Bagwell-Staiger model
economic effects 375
special protection framework 367
balance-of-payments (BOP) problems 325
Baldwin, R.
PTA expansion domino effects 307, 309, n. 3
Bandiera, Oriana 287
Bank for International Settlements (BIS) 314
Barcelona Convention for Protection of the Mediterranean Sea against Pollution 123
Barratt-Brown, E.P. 36, n. 113, 119
Barrett, Scott 318
Basel Convention on trade or transportation of hazardous waste across international borders 322
Beef/Hormone panel 21
see also Appellate Body of WTO
Berne Convention 414
for the Protection of Literary and Artistic Works (1886) 51, 404
and Universal Copyright Convention in 1992, intellectual property rights, enforcement of 450
Bernhardt, Dan 460
Bernheim, B. Douglas 342
Bernstein, Lisa 285
Bhagwati, Jagdish
Article XXIV of GATT, ambiguity in 305–6
debate with Daly 500
formation of CUs 310, n. 8
Principle of Targeting 505
PTAs 309, n. 3
bilateral integration, stages of 397
bilateral negotiation, conflicts of 529
bilateral tax treaties
costs of treaty formation 529
and foreign direct investment (FDI)
effect on 526, 532–5
empirical framework for statistical analysis 530–2
Blonigen, Bruce A.
bilateral tax treaties and FDI 530, 532, 535
new treaties and FDI 540, 542
indirect effects of AD law 394
retaliatory AD measures 380, 381, n. 14
Boeing and Airbus, example of 150
Boltuck, Richard
AD filing, strategic motives for 368
on CADIC model 101, n. 13
Bond, Eric W.
on multilateral policy cooperation 318
on multi-market collusion 343
Border Tax Adjustments 1970 GATT Working Party Report on 190, 208
Bown, Chad P. 381
retaliatory AD measures 380, 381, n. 14
Brainard, S. Lael 460
Brander, James A.
dumping as international price discrimination 365
endogenous injury in reciprocal dumping model 460

Brander, James A. (*cont'd*)
market failures and trade and environment relationship 502–3
tariff policy and multilateralism 346
Braudel, Fernand 280
social networks and trade 285
Brazil 364
British Commonwealth countries 361
Brittan, Sir L. 326
Brown Weiss, Edith
compliance institutions 135
compliance strategies 145, 147, 175, n. 33
Ozone regime 164, 167, 180, n. 95–8
Brunsdale, A. 79
Burma logging 502
buyback 283
buyer cartels 345
Byrd, Senator 95

CADIC (Comparative Analysis of Domestic Industry's Condition) model 79–80, 101, n. 13
Canada 361, 385
Act to Amend the Customs Tariff, 1897 360
Automotive 219
Certain Measures Affecting the Automotive Industry 210–11
Certain Measures Concerning Periodicals case 197, 203
EU complaint over exemptions/limitations in Canadian patent law 53–4
Excise Tax Act and GATT 1994, Article III 202–3
stockpiling provision 54
Canada-Chile Agreement without AD 392
Canada-US Free Trade Agreement 314
capacity building 146–8, 153, 156, 161–3, 165–6, 167–8, 170
capacity problems of developing countries 170
Carlson, Lisa 139
Carr, David
Markusen and Maskus, general equilibrium model of MNE activity 530, 544, n. 19
Carraro, Carlo
environmental free-riding incentives 329, n. 1
International Environmental Agreements 318
Cartegena Convention for the Protection and Development of the Marine Environment of the Wider Caribbean Region 123, n. 29
cartel
agreements, self-enforcing 338–40
policy
in Cournot oligopoly model 335–8
potential role for international coordination regarding 334
and price-fixing arrangements among firms 335–43
Casella, Alessandra
model of international arbitration supplementing work of courts 285–6
and Rauch, first-generation model of information asymmetry 290
Cass, Ronald 79
Cassing, James H. 460
Casson, Mark 529
causation
for antidumping, hot-rolled steel 84–5
for safeguards
lamb meat 82
line pipe 84
wheat gluten 81–2
certificate of origin 250
Cesar, Herman 316
CFCs *see* chlorofluorocarbons
Chander, Parkash 318
Charnovitz, Steve 238, n. 115, 121
on *Dolphins and Tuna* 29, n. 20, 41, n. 206, n. 211
environmental exceptions 39, n. 70
on free trade 27, n. 6
labour standards 122, n. 24
unilateral trade restrictions 14
Chayes, Abram 142
on compliance systems 142, 147–8
Harmony game 140
Chayes, Antonia
on compliance systems 142, 147–8

Chen, Yongmin
  on parallel imports 417, 456
Chichilnisky, Graciela 502
Chile
  Alcoholic Beverages case, taxes under transitional and new systems 205–6
  Canada, agreement with 392
  Central Preventive Commission and WTO, Article III 205
  Taxes on Alcoholic Beverages 197, 233, n. 75–6, 234, n. 77–81
    and Article III of GATT 203–4, 207, 223
China 169, 364
  joining the Montreal Protocol 166
  response to serious deforestation, and restricted logging 502
chlorofluorocarbons (CFCs) 163–5
Cho, D.S. 280, 289
Choi, P.
  on patent validity 423
chrysotile fibers 214
Cicerone, Ralph 164
Cini, Michelle 350
Claessens, Stijn 478
Coasian bargaining 314, 316
cold-rolled steel AD case 75, 86
  automobiles and other uses 87
  dumping case 98
  econometric study in investigation 99
  factors
    influencing demand 87–8
    influencing the supply of 88
  and hot-dipped galvanized market 87
  and hot-rolled steel 88
  import market 88
  primary inputs 88
  supply of capacity 89
Commercial Policy Analysis System *see* COMPAS model
commercial presence abroad and GATS 492, n. 5
Committee on Regional Trade Agreements (CRTA) 239, 242, 268
common market, goal of 392
comparative advantage
  principle of 504
    trade and 503
  theory of 9
Comparative Analysis of the Domestic Industry's Condition *see* CADIC model
COMPAS (Commercial Policy Analysis System) 79–81, 90, 93, 101, n. 15
compensatory taxation 8
competition policy
  economic analysis 74
  justification for 392–7
  optimal 348
  other forms of 347–8
competitive relationship between imports and domestic products 225
compliance 135–9, 141–2, 144–8, 152, 154–7, 159, 161–2, 164, 168–9
  definition 135
  and game theoretic framework 137–8, 150, 153, 158–9, 163–4
  and game theory 138–9, 144–5, 148–50, 152, 155–7, 163–4
  and GATT 148–9, 152, 154–9, 160, 163, 166–7, 171
    see also GATT dispute settlement system
  GATT expectations for 113
  institutions in international trade law and international environmental law, comparative analysis of 134
  investment and the role of DSU 55–9
  and Montreal Protocol (Ozone Regime) 148, 165–8
    capacity building 165–6
    dispute settlement 165, 167–8
    TEAP 167, 169, 171
  by states, justifications for 45
  strategies
    analysis 145–9
    Types (I, II and III) 156, 163
  systems, approaches to designing 148
  and WTO 136–7, 146, 148, 152–3, 158–9, 160–3, 166, 169–71
    capacity building 156, 161–3
    Dispute settlement 136–7, 148, 159–62, 170–1
    TPRM 137, 161–2
compliance dilemma 136

compliance institutions 136–7, 144–5, 150–1, 154–5, 157, 159, 161–5, 168–9,
  enforcement oriented 146, 148–9, 156–7, 160–1, 164, 166, 168
  management oriented 148–9, 168
    see also institutions
compliance monitoring *see also* international monitoring systems
compliance strategies 139–40, 142, 145, 147, 149, 153, 155–7, 159–62, 167–9
  capacity building 146–8, 153, 156, 160–3, 167–8, 170
  incentive based 136–7, 142, 145–8, 152, 155–6, 161, 163, 167–9,
  modifying commitment levels 136, 144, 146–9, 152, 155–8, 160–1, 163, 167–9
  see also entries under incentive based institutions
compulsory licenses 410, 417
Conconi, Paola 313
conditionality 313, 320
  and their impact on multilateralism 317–21
  linking environmental agreement to multilateral trade cooperation 320
  separation, and open rules in international law 321–6
contract enforcement authority 284
Convention on Biological Diversity 124, n. 32, 418
Convention establishing the World Intellectual Property Organization (WIPO) 121, n. 16
  see also World Intellectual Property Organization
Convention on the International Sale of Goods (CISG) 112
Convention on International Trade in Endangered Species of Flora and Fauna (CITES) 15, 322
Cooter, Robert 164
Copeland, Brian R.
  issue linkage 316
  market failures and linkage between trade and environment 502
Corden, W. Max 364–5
costs associated with mandatory system of adjudication for disputes between sovereign states 59
Council of Trade in Services 487
counterfeit goods/counterfeits
  demand for 429–32, 455
  producers of 434–6
    and imperfect law enforcement 428
  tolerance free for 454
  tolerating 451–3
  two-country model for DVD products 430
counterfeit products 454
  software copying 429
  tolerating 451
counterfeiter
  risk-averse 435
  risk-neutral 436
counterfeiting, socially optimal 453–4
countervailing duty (CVD) 358–9
country of origin 294
Cournot oligopoly, simple two-country model of homogeneous product 335
Court of International Trade 81
Crampes, C.
Crawford, Carol 79
creation of side-agreements 157
cross-issue international cooperation, literature on 316
cross-issue punishment strategies 317
CRTA (Committee on Regional Trade Agreements) of WTO 239, 242–3
  notification prior to PTA to be GATT-compatible 247
  powers of 246, 271
  reports 245
    legal nature of 261–2
  as successor to GATT Article XXIV Working Parties 245
  Track I 244–5
CU *see* Customs Union
Customs Union (CUs) 244, 294, 300
  between the Czech and the Slovak Republic 268, 274, n. 14
  between EC and Turkey 262
  formation of 251
  and an FTA, difference between a 249

Dagan, Tsilly 528
Daly, Herman E. 500
Dam, K. 305
Das, Satya 460
Davey, William J.
economic interdependency, international 149
GATT rounds 157–8
Davidson, Carl
antitrust policy 341, 346, 354, n. 4
Davies, Ronald B.
bilateral tax treaties and FDI 530, 532, 535
new treaties and FDI 540, 542
Deardorff, Alan V.
patent rights 414
PTAs 310, n. 16
trade in goods and trade in services 478
Deneckere, Raymond 354, n. 4
Department of Commerce, US 84
Department of Justice, US 79
successful prosecutions of number of cartels 334
developing countries 50
and enforcement of labor and environmental standards 117
incentive to enforce copyrights or trademarks 450
labor and environmental standards, opposition to incorporation 118
and least-developing countries, compliance with TRIPS Agreement 57
and Montreal Protocol 167
as users of AD 387
Dhareshwar, Ashok 534
Diamond, Jared 503
Dillon Round 157
direct strategic linkages 316
directly competitive or substitutable products 221
discrimination talk
in coercion cases 19–21
in laissez-faire cases 21
dispute settlement 147, 154–6, 158–63, 168
see also WTO dispute settlement
Dispute Settlement Body (DSB) 48, 159
Indian – Measures Affecting the Automotive Sector 218
United States – Tax Treatment for Foreign Sales Corporations – Recourse to Article 21.5 of the DSU by the European Communities 218
in the WTO 266
Dispute Settlement Understanding (DSU) 14, 43, 159
appeal stage 49
Appellate Body's decision 49
Article 3 of the 45, 52, 63–4, n. 22
Article 22 of 160
consultation stage 47
dispute panel rulings and recommendations 49
establishment of a panel 48
informal settlement of disputes 46
international intellectual property regulation 60
legitimate expectations test 49
mandatory adjudicatory team 43
panel report 48
process
as kind of market 45
six progressive stages, formal 47–9
role of compliance in developing countries 58
rule-oriented process 43
signals 56
as supranational adjudicatory process 51
third parties who join complainant 65, n. 36
voluntary compliance and progressive stages of dispute settlement under 45–9
diversification, implication of scale economies 289
Dixit, Avinash K. 31, 309
trade and informal institutions 285
Doernberg, Richard 529
Doha Declaration 418
Development Agenda 161
Ministerial document 397

domestic barriers to international trade 151
  secondary, domestic policies as 317
domestic and imported products, nondiscrimination principle determination of likeness of 193–6
domestic wholesale markets, emergence of 289
double taxation objective 528
Downie, David L.
  Ozone regime 164, 167, 169
Driesen, D.M. 5
dumping
  definition of 381, n. 2
  rules for determining existence of 368
duopolists and the counterfeiter 449
Durling, J.P. 73
duties calculated on dutiable imports 274

Eaton, Jonathan 420 n. 13
EC *see* European Community
econometric study of AD
  case on cold-rolled steel 89–90
  investigation implications 93
  method 103, n. 35
economic study of AD 86–9
  case on cold-rolled steel results 86, 90–3
economic test of like products 221
economics of preferential trade agreements 294
economics in trade remedy cases 79
Ederington, Josh 317
EEC *see* European Economic Community
Elliott, Kimberley 280
endogenous injury 459
  in duopoly model with production and investment stages 460–7
  model, policy implications 467
enforcement
  costs to reduce number of counterfeiters 450
  oriented approach to compliance 148
  strategy for public international standards 113
environment and welfare, long-run effects of trade on the 503
Environmental Conventions 133, n. 108
environmental protection 315
  “competitiveness effect” 116
environmental reform
  under autarky
    long-run effects of 520
    and under trade, long-run effects of 520
  effects on steady states in trade 509
  under free trade different effects of 522
  in open and closed economies 516–17
  under trade, long-run effects 521
environmental stocks, endogenous evolution of 499
environmentalists and economists, debate between 500
Esty, Daniel C. 37, n. 126, 40, n. 193, n. 196
Ethier, Wilfred J. 382
  cyclical dumping model 460
  model of dumping as pricing below production cost 365
  “new protectionism” 359
Europe 361
  asbestos case and GATT, Article XX protection of human health 215
European Common Market and trade in services 486
European Community (EC) 259
  common AD law 399, n. 11
  restriction on beef from cattle injected with hormones 11
  Tariff Treatment on Imports of Citrus Products from Certain Countries in the Mediterranean Region
    panel report 262, 265, 270
    unadopted panel report in 277, n. 60
  WTO obligations 161
European Community Trademark 420, n. 12
European Economic Area (EEA) 392
European Economic Community (EEC) 15, 295
  Appellate Body’s decisions Korean Beef 221
  Asbestos 216, 218
    Appellate Body’s decisions 221
  Bananas 219, 221
    report 265
European Union (EU) 15, 487
  Agreements with Eastern European and Mediterranean Countries 392

ban on growth hormones to both European and foreign beef produces 21
cartel policy 350
complaint against Canada over exemptions/limitations in Canadian patent law 53–4
"crisis cartels" in 350
substitution of AD with AT for intra-union trade 391
treaty, Article 85 of the 350
US complaint, over lack of copyright infringement remedies in Greece 53
exhaustion or first-sale doctrine 410
External Requirement for Free Trade Areas 249

facially neutral tax and regulatory measures, GATT panel case law on 188
Farber, Daniel A. 24
Farrior, Stephanie 127, n. 59
FDI *see* foreign direct investment
Feddersen, Christoph T. 324
Federal Trade Commission (FTC) 79
Fertile Crescent, technological advance and population growth 503
Findlay, R. 311, n. 17
Finger, Michael J.
on antidumping 360, 362, 364, 395
filing behaviour 367, 387
initiations 388–9
legislation 385
firms as Bertrand competitors 341
Fischer, Ronald D. 364
Folmer, Henk 316
Food and Agricultural Organization (FAO) 314
foreign competition 187
foreign direct investment (FDI) 286
activity between treaty partners, amount of 528
data coverage by OECD country 533–4
effect of taxation on 526
non-promotion effect of treaties on 542
foreign governmental compliance with national standards, access to domestic markets upon measures conditioning 25
foreign products, regulatory measure impact on 188
framework–protocol approach *see* ozone regime
free trade
definition 5
three concepts of 7, 26
as trade free of discrimination 26
Free Trade Area of the Americas (FTAA) 392
free trade areas (FTA) 244, 294
and CUs formed under art, XXIV of the GATT 309, n. 1
free-riding incentives 320
Freund, C. 311
Friendship, Commerce, and Navigation (FCN) treaties 51
full employment line (FEL) in model on trade and environmental regulation 515
Fung, K.C. 341

Gallaway, Michael P. 400, n. 22
Gambetta, D. 284
Game Theory 137–8, 144, 146, 150, 152, 157–9
Chicken 141
exogenous shock 143
framework for analyzing international cooperation and institutions 137
framing 143, 146–7, 163
game formation 138, 143, 146, 149, 151–3
grim trigger strategies 143
Harmony 140
incentive-based strategies 144, 146
see also entries under incentive based institutions
and international relations/law 141–2, 147, 150, 152–3, 157–9, 161–2, 164, 167–9
number of players 140–1, 143–4, 152–3, 164
states as rational actors 137–9, 145, 157
penance strategies 143
Prisoners' Dilemma 138, 139, 142, 152, 164 *see also* trade liberalization
reciprocal strategies 142
and role of domestic politics in making of international agreements 85, n. 42

Game Theory (*cont'd*)
Suasion 140, 152
trigger strategies 143
GATS (General Agreement on Trade in Services) 108, 153, 322, 427, 487–92
2000 negotiating round 490
accounting disciplines 489
Article II 240
Article V 242
conditions for PTA to be judged WTO-consistent 261
articles major 487–8
assessment 488–92
commitments toward most-favored-nation (MFN) treatment 487
Council of Trade in Services 489
negotiation of 486
nonbinding guidelines for mutual recognition agreements (MRAs) 489
regime 261
sector negotiations 489
submission of schedules 489
GATT dispute settlement system 154–6, 158–1
GATT *see* General Agreement on Tariffs and Trade
GATT (General Agreement on Trade and Tariffs) 6, 43, 108, 153, 294, 314
adopted panel reports, legal status of 276, n. 48
Agreements
on the Implementation of Article VI 361
on Subsidies and Countervailing Measures (SCM Agreement) 207
Analytical Index, Guide to GATT Law and Practice 275, n. 18
Antidumping Agreement or the Antidumping Code 361
Article I 153, 240
Article II.4 of the 218
Article III *see* GATT Article III
Article VI original formulation of AD 361
Article X 154
Article XI 23, 154
prohibition upon quantitative restriction 19
Sanitary and Phytosanitary Agreement 10–13
Article XII 325, 330, n. 7
Article XIV of 325
Article XVI 154
Article XVIII 330, n. 7
Article XIX, escape clause allowing temporary use of import restraints when imports cause serious injury to domestic industry 154
Article XX *see* GATT Article XX
Article XXI 325
national security 154
Article XXIII 155, 160
Nullification or Impairment 154
Article XXIV *see* GATT Article XXIV
Article XXIV.5 (a) customs union 250
FTAs 249
Article XXIV.6 on customs union 251
Article XXIV.8 258
interpretation 255
PTA and 253
on regulation of commerce and its restriction of trade 261
terms substantially all trade 253–8
Article XXV, wavier authority 154
Balance-of-Payments Committee 330
code providing for the imposition of AD duties 365
contracting parties 13
disciplines, redefining 41, n. 217
evolution of institutional framework 159
exceptions to free trade 22
four principles 153
Generalized System of Preferences 309, n. 1
implicit or origin neutral discrimination 187
most-favored-nation clause (MFN) of 263
principle of nondiscrimination, exceptions to the 309, n. 1
requirement, on external tariffs original 305

and other restrictive regulations of commerce 258–61
rules
consistency of a PTA with the 266
internal barriers to trade 305
shortfalls in compliance 126, n. 48
uniform "multilateralization" of the trade liberalization game 158
and WTO dispute settlement systems
explicit discrimination 187
differences in production or processing methods (PPM's) 222
GATT Article III 35, n. 103, 186, 189, 193–4, 212, 221
"aims-and-effects" approach to determining "likeness" 195
applies to NTBs 158
Canada
Certain Measures Affecting the Automotive Industry 210–11
Certain Measures Concerning Periodicals case 203
Excise Tax Act 202
Chile
Taxes on Alcoholic Beverages 203–4, 207, 223
determination of the "likeness" of the products 200
directly competitive or substitutable products similarly taxed 201
EEC, Bananas 209
Indonesia, Certain Measures Affecting the Automobile Industry 207
internal tax measures and national treatment 197–208
Japan, Film case, panel in 210
Japan, Taxes on Alcoholic Beverages 209
Appellate Body in 204, 207
three-step test 203
Korea, Taxes on Alcoholic Beverages 203
National Treatment principle 19, 208
violation in European Communities – Measures Affecting Asbestos and Asbestos-Containing Products 214–18
prohibition of discriminatory taxes 200
protection to domestic production 202
Shrimp–Turtle case 226
standards for examination of the conformity of an internal tax measures 198
test for like and directly competitive or substitutable products 198
Trade Free from discrimination 10
Tuna–Dolphin case 226
US–Gasoline 209
WTO 198
dispute Settlement system, United States–Standards for Reformulated and Conventional Gasoline 208
GATT Article XX 10–13, 324–6
defenses in decisions about the appropriate scope of national regulation 11
exceptions 17, 22, 222, 226
health and safety regulation 154
limited scope of 227
of trade and environment 323
GATT Article XXIV 242, 256
allowing customs unions and free trade areas 154
complementarity between nondiscriminational and regional trade 256
and its corresponding Article V GATS, role 241
defense conditions 269
FTAs and CUs 244
justiciability of related claims 265
PTAs 244, 304
test 241, 244
understanding on 267, 272
working group 270
Working Parties 271
GATT dispute settlement system 159, 161, 178, n. 61
facially neutral tax or regulatory measures and the principle of national treatment under the 188–96
panels 32, n. 71
process 155
type I strategy aimed at the payoff structure 156

GATT panels 265
decision in US, section 337 225
in the Herring and Salmon Processing case 227
on Japan, Taxes on Alcoholic Beverages case 213
GATT working parties 263
on PTAs
on the consistency of notified 265
pragmatic attitude 263
GATT/WTO regime 148–9, 152–3, 158, 160–1, 163, 166, 169, 171,
and institutions 152–6, 158–61, 163, 166, 169, 171
and game theory 148–9, 152, 157–8, 161
see also GATT, WTO, international trade
gangs acting as contract enforcers in trade 284
see also Mafias
GATT/WTO system 113
case law on national treatment, critique of the 220–8
decisions, Tuna-Dolphin case 221–2
and implementation of AD laws, membership in 388
interpretations of Article III.4
like products 221
jurisprudence, product-process doctrine in 229, n. 37
law 112
multilateral trading regime, nondiscrimination principle 186
provisions of National Treatment under Article III 223
regime 149
as assessing the legitimacy of various trade-inhibiting measures 152
trading system 153–63
sanctioned PTAs 295
trade negotiations in Doha, Qatar 322
General Agreement on Tariffs and Trade, (“GATT”), 146, 148–9, 152–60, 163, 166, 169
and compliance 146, 149, 153, 155–6, 158–9, 160–1, 163, 166–7, 169
General Agreement on Trade in Services *see* GATS
Generalized System of Preferences in GATT 309, n. 1
Germany 385
Gilbert, Richard 407
global climate treaties 321
global trade and environmental institutions 321
global welfare and parallel trade 422
Gould, David M. 290
Grane 179
Grane, Patricio 178, n. 63, 179, n. 78
gravity model 290
Greece, US compliant against 53–4
Green, Edward J.
AD cases 367, 382
and Porter extension 368
Greene, Owen 167
Greening the GATT 27, n. 1, 40, n. 195
Greening International Law 123, n. 31
Greif, Avner 284
Grossman, Gene M.
on carbon steel case 79
and Helpman
analysis of location of outsourcing in a model of imperfect contracting 284, 286
political motivations for protection 367
PTAs 306
and Levinsohn, model of effects of imports on returns to capital 470, n. 8
and Shapiro, trade in counterfeit products model 428–9, 454, 456, n. 6
Guiso, Luigi 470, n. 9
Guzman, Andrew T. 401

Haaland, Jan I. 394
Hague Declaration on the Environment 123, n. 30
Hall, Keith 364
Hamada, Koichi 528
Hansen, Wendy L. 364, 395
“hard” sanctions 148
Harmony game 140, 152
Hartigan, James C. 355, n. 8
AD law 394, 470, n. 12

Hartman, David 529
Hausman, Jerry 96
economic model 97–8, 104, n. 47
Havemeyer, H. 341
health and environmental regulations 10
Helpman, Elhanan 176, n. 43
and Grossman
analysis of location of outsourcing in a model of imperfect contracting 284, 286
political motivations for protection 367
PTAs 306
Henkin, Louis 62
ways that international law is enforced 63, n. 12
Heston, Alan 534
Hexner, Erving 338
Hicks, Laurinda L. 130, n. 85
Higgins, Richard 450
counterfeit products 429
high bargaining power 284
Hillman, Arye L. 460
Hines, James R. Jr. 527–8
Hoekman, Bernard M. 490
Holbein, James R. 130, n. 85
holdup
resolved by institution of social networks 285
resolved by ownership structure 286
in trade 282
Hong Kong, China in Turkey dispute 256
Horn, Elsa 355, n. 16
Hot-Rolled Steel Products from Japan 84
Howse, Robert 318, 325
Hudec, R.E. 24, 126, n. 47, 187–8, 227, 263
GATT/WTO Jurisprudence 40, n. 197
hydro-chlorofluorocarbons (HCFCs) 163

IBRD *see* International Bank for Reconstruction and Development
Ikenson, Daniel J. 470, n. 10
ILO (International Labour Organization) 110, 119, 128, n. 60, 133, n. 108, 314
agreements, ratifications of "core" 122, n. 25
IMF (International Monetary Fund) 314, 325, 389
and OECD guidelines specifying investment as FDI 544, n. 21
income risk associated with predation on trade 290
India 169, 364
in the balance of payments-related dispute 272
joining the Montreal Protocol 166
ODS production capacity 167
Patent Protection dispute 49
Quantitative Restrictions on Imports of Agricultural, Textile and Industrial Products, panel report on 271
TRIPS Article 70.8 implementation 49
Indonesia
antidumping in 363
car program, sales tax benefits 207
Certain Measures affecting the Automobile Industry 197
industrialized country protectionism 118
information, imperfect and asymmetric 290
injury
analyses under US Trade Law, need for economic rigor in 73
from imports standard 360
requirement in antidumping (AD) law 459
to the United States' hot-rolled steel industry 84
innovators concentrated in developed countries 421, n. 29
institutional barriers to trade 281
Institutions
and compliance strategies 136–7, 142, 145, 147–8, 156–7, 160–1, 165, 168–9
and Game theory 137–9, 141, 144, 149–50, 152–3, 158–9, 161–2
and GATT 137, 152–61, 163, 166
Incentive based 136–8, 142–3, 148, 153–4, 157–8, 160–1, 163, 166, 169
negative/sanctions 148, 151–2, 155, 163
positive/subsidy 150–1, 154, 163

Institutions (*cont'd*)
information gathering 145, 146–7, 153, 162, 165
monitoring 146–7, 153, 160, 167, 169 *see also* international monitoring systems
and WTO 146, 148–9, 152–3, 158–9, 161–3, 169–70
intellectual property (IP)
international law 116
protection, equilibrium point for global 61
rights, trade-related field of 108
treaties 50
intellectual property rights (IPR)
developing countries weak systems 414
as established in Article 4 (bis) of the Paris Convention IPRs in Article of the Berne Convention 408
and international or broad exhaustion 411
and labor and environmental rights 112
as private rights under national law 411
and regional exhaustion 411
and subsidized law enforcement 450–1
as territorial rights 404
in the WTO 116
intentional noncompliance 142, 155–6, 161, 163, 168
interdependency problem 170
and need for international cooperation 163–4
international aid policy, conditionality 321
international asset markets 288
International Bank for Reconstruction and Development (IBRD) 314, 325
International Bovine Meat Agreement 329, n. 3
international cartels, globalization effect 334
international cooperation 315
analyzed as Prisoner's Dilemma 138
on cartel policy 347
framing, forming, and playing games 143–5
International Court of Justice (ICJ) 43, 128, n. 62
on Article 32 VCLT 254
ozone regime 168
International Dairy Agreement 329, n. 3
International environmental law 14, 115, 135–6, 142, 164–5, 169, 171
innovations in 137
multilayered legal frameworks 170
and the Ozone regime 163–9
underenforcement of 110
International environmental regimes 149, 152, 164–5, 167, 171
international harmonization of competition policy, potential gains from 361
international labor law 115
underenforcement of 110
International Labour Organization *see* ILO
International Labour Standards Ratifications 133, n. 108
International law
conditionality 321–2
open rules 324–6
separation 322–3
International Law of the Sea Tribunal 62, n. 9
International Law on Trafficking in Women and Children 127, n. 59
International Monetary Fund *see* IMF
international monitoring systems 136–7, 141–2, 144–5, 153, 157–60, 162–3, 166–7, 170 *see also* compliance institutions
international noncoercion principle 7
law and theory supporting 7–18
international obligations under the WTO "Safeguards Agreement" 76–8
international price-fixing agreements 354, n. 5
international standards to rules, evolution 50–2
international trade 135–6, 149–53, 157–9, 162, 169–70
conflicts of interest 281
and formal enforcement 283
in good and services 314
informal enforcement 285–6
and information 290
in listed endangered species and their parts 15
as multiplayer game 152

policy, formal institutions of 280
regime enforcement mechanism 113
and risk diversification 288–90
see also GATT/WTO regime
International Trade Commission (ITC) 74, 99, 428
1999 investigation in Line Pipe case 84
approach to causation and non-attribution 93
formal economic analysis at the 78–81
imports of cameras 101, n. 18
injury investigations under title VII 75–6
US Law 75
litigation 86
non-attribution international obligations on 76
Office of Economics 93
political consequences of embracing economic analysis in cold rolled steel case 95–8
practice
on causation, traditional 81
in injury investigations and recent developments 78–85
international trade law 112
and the GATT/WTO Regime 149–63
underlying interdependency problem and the need for international cooperation 149–53
as public international law 136
international trade in services 472
expansion predictions and the data 479–86
three categories of data on 479
International Union for Conservation of Nature and National Resources 122
involuntary compliance 61
IPR *see* Intellectual Property Rights
Irwin, Douglas A. 274
Ishibashi, Ikuo 428
issue linkage 316
literature of 316–17
ITC *see* International Trade Commission
iterated game(s) 148, 155, 157
iteration(s) 138, 142, 144, 155, 157
ITU 314
Jackson, John H.
on DSU 62, n. 5
economic interdependency, international 149
GATT rounds 157–8
international trade law 187
Jacobs, Brenda A. 132, n. 99
Jacobson, Harold K. 135
Jacques-Laffont, Jean 429
Janeba, Eckhard 528
Jansen, Marion 478
Japan
against AD law 394
Report on Measures Affecting Agricultural Products 34, n. 95
taxes on 197, 204
Taxes on Alcoholic Beverages case under GATT 188–92, 194, 202, 208, 230, n. 45, 231, n. 52–8, 232, n. 59–60
treatment of sound recording in violation of Articles 14.6 and 70.2 of TRIPS Agreement, United States 53
US compliant against 53
Jeddah Regional Convention for Conservation of Red Sea and Gulf of Aden Environment 123, n. 29
Joskow, Paul L. 476

Kaikati, Jack 450
Karp, Larry 502
Kelly, Kenneth H. 470, n. 12
Kemp, M.
Customs Union 300
PTAs 307
Kemp-Wan 310
construction 301–3
and Panagariya-Krishna constructions of welfare improving PTAs 305
theorem demonstrating the possibility of Pareto-improving CUs 241, 301
Kennedy Round (1967) 157, 361
Kilgore Committee 354, n. 1
Klemperer, Paul 407
Klepper, Steven 470, n. 14
Knetter, Michael M. 387
Knox, John H. 136
Koopman, Robert B. 98

Korea
  antidumping in 363
  Measures Affecting Imports of Fresh, Chilled and Frozen Beef 208, 213, 219, 236, n. 90–3
  Taxes on Alcoholic Beverages 197, 204, 233, n. 71
  US Definitive Safeguard Measures on Imports of Circular Welded Carbon Quality Line Pipe from 269
Kranton, Rachel E. 285, 289
Kreinin, Mordechai E.
  discussion of comparative advantage 30, n. 43
Krishna, Pravin
  and Panagariya, FTAs 302–3
  PTAs 306
Krugman, Paul R.
  on dumping 365, 460
  on free trade 150–1
  multilateral policy cooperation 318
Kurdaliev 460
Kuwait Regional Convention for Co-operation on the Protection of the Marine Environment from Pollution 123, n. 29
Kwan Choi, E. 427
Kyoto Protocol 171, n. 4

labor and environment
  problems from formal international scrutiny 115
  protections impact on cost of production 116
  public–private distinction in 112, 125, n. 43
  rights, trade-related fields of 108
  standards 107
    public international law 112
LaGarce, Raymond 450
laissez-faire concept 9–13
  definition of "free trade" 16
  law and theory supporting 7–18
Lamb meat case 82–3, 102, n. 22
Landefeld, Steven J. 484
Langinier, C. 451, 454
legal principles, rules and standards 49
less favorable treatment
  EEC, Bananas case 221
Levine, Ross 478
Levinsohn, James
  antitrust 384
  and Grossman, model of effects of imports on returns to capital 470, n. 8
  optimal competition policy 348, 354, n. 2, 355, n. 16
Levy, P.
  model of trade policy as being determined by majority voting 306
Liebeler, Susan 79
life forms, patenting of 418
Lima Convention for the Protection of the Marine Environment and Coastal Areas of the South-East Pacific 123, n. 29
Limão, Nuno 317
Lindsey, Brink
  AD measures 390, 400, n. 20, 470, n. 10
linkages, importance of 290
Lipsey, R.G. 471
Litan, R.E. 381
Lithuania, AD law 397
Lome Convention 265
London Amendment 1990 *see* ozone regime
Long, Ngo Van 460
Long, Olivier 113

Madrid Pact for trademark protection in 1989, China 450
Mafia 282
  and the State interaction 288
Maggi, Giovanni 317
Malta as tax haven 529
Malueg, David A. 417
management-oriented approach
  to compliance 148
  to international cooperation 137
Mandel, Michael H. 459
Mani, M. 502
Marcouiller, Douglas 291
  and Anderson
    formal model of trade and predation under anarchy 287
    trade flows to Latin America 279
Marin, Dalia 283

Marine Mammal Protection Act (MMPA) 19, 23
market
failures and relationship between trade and the environment 502
thickness principle 286
Markusen, James R.
with Carr and Maskus, general equilibrium theory of MNE activity 530, 532
Marrakesh Agreement 159
establishing the World Trade Organization 28, n. 13
Maruyama, Warren H. 188
Maskus, Keith E.
with Carr and Maskus, general equilibrium theory of MNE activity 530, 532
IPRs and trade 412, 416–17, 422, n. 39
Mastel, Greg 364
Matsumura, Toshihiro 428
Mattoo, Aaditya 187, 189
Maur, Jean-Christophe 368
Mavroidis, Petros C. 259
McCallum, John 279, 293
McDaniel, Christine A. 94
McDonald, Brian 364
McDowell, Mark A. 525
McGowan, Lee 350
McLaren, John 286, 289
analysis of choice between detailed costly contracts and handshakes in markets 284
analysis of the role of sunk costs and trade policy determination 307
McMillan, J. 305
medicines, developing countries and local manufacturing of 417
mercantilist policies 9
MERCOSUR (CU between the Argentine Republic, Brazil, Paraguay, and Uruguay) 295, 487
double AD regime 392
mergers
in Cournot oligopoly 354, n. 4
policy 347
and cartel policies differences between 348
Merrill, William 428
Messerlin, Patrick A.
antidumping measures as retaliation 380
antidumping policy 393
GATS 2000 Round 490
Mexico 15, 364
antidumping in 363
MFN *see* most favored nation
Minehart, Deborah F.
networks in informal enforcement 285, 289
Ministerial Conferences
in Doha 423, n. 44
Singapore Ministerial Declaration 130
see also Doha
mixed market
with counterfeit producers 427
law enforcement as barrier to entry 441
long-run effects of fines and arrests 449
long-run equilibria 441–5
in a two-country framework model 432–9
duopoly and counterfeiters 445–9
effects of counterfeiters 448
law enforcement and short-run Nash equilibrium of 437
long-run Nash equilibria 443–5
long-run Stackelberg equilibrium 445
models
enforcement problems across sovereign jurisdictions resolved by counterpurchase 283
of informal enforcement by a Mafia 287
of informational environment 290
of tacit collusion and multi-market contact 343
modern commercial policy 358
Molina, Mario 164
monopolist, production decision in presence of counterfeiters 433
monopoly
and counterfeit producers 432–9
deterrence of innovation 459
middlemen resolving holdup 283
profit 433

Montreal Noncompliance Procedure (NCP) 111
Montreal Protocol on Substances that Deplete the Ozone Layer 14–15, 111, 125, n. 38, 165, 167
  1990 London Amendment 166
  Article V of 166
  Article 6 of 167
  Environmental Assessment Panel 167
Moore, Mike O. 395
Morkre, Morris E. 470, n. 12
Most-Favored-Nation (MFN) principle 186, 240
  treatment 487
    and IPR under GATT 416
multi-country economic-integration areas 487
multidimensional economic integration 397–8
multidimensional international cooperation, emergence of new issues 326
multilateral cooperation 153, 317
multilateral environmental agreements (MEAs) 322
  free-ride on 319
Multilateral Fund 166
  noncompliance procedure 169
    and implementation committee 168
  Scientific Assessment Panel 167
  Secretariat for the Protocol in Nairobi, Kenya 165
  Technology and Economic Assessment Panel (TEAP) 167, 169, 171
multilateral institutions
  conditionality, separation, and open rules in 313
multilateral negotiation
  additional considerations 326
  emergence of new issues 326
  transaction costs 327
multilateral policy cooperation choices 317
multilateral trading system, evolution of the 158
multilateral, cross-issue conditionality in international cooperation 322
multi-market
  collusion 342–3
  contact model 351
Murray, Tracy 460, 470, n. 12

NAFTA *see* North American Free Trade Agreement
Nairobi Convention for the Protection, Management and Development of the Marine and Coastal Environment of the Eastern African Region 123
Nash-Cournot output decision 460
national cartel laws 348–51
national treatment principle 186, 196
  in Article III of GATT 186
  genesis in international trade law 186
  in international trade law 185
national welfare
  and cartel policy 344–6
  maximization 349
navigation acts 8
Nehru, A. 534
Nelson, Douglas R. 364
Netherlands 15
Netherlands-Antilles as tax haven 529
New Zealand
  Agricultural Implement Manufacture, Importation and Sale Act, 1905 360
  Customs Amendment Act (1921) 360
Newman, Andrew 289
Niels, G.
  antidumping and antitrust 396, 398
No Trade Line (NTL) in model on trade and environmental regulation 515
Noerr-Pennington doctrine 342
Noll, Roger G. 476
non-attribution
  cold-rolled steel 85–94
  ITC statutory language and international obligations requiring 75–8
  legal obligation of 73
noncoercion principle 22, 25
  in international law 15
nondiscrimination principle 153, 294
  even-handed regulation 18
  law and theory supporting 7–18
  national tax 185
  substitutability evaluating in end-uses 205

in trade relations 8
two pillars of 186
nontariff barriers (NTBs) 6, 157–8
Nordhaus, William D. 407
North American Free Trade Agreement (NAFTA) 240, 259, 269, 295, 392, 487
Noumea Convention for Protection of the Natural Resources and Environment of the South Pacific Region 123, n. 29

Oakland, W. H. 164
ODSs *see* ozone-depleting substances
OECD (Organization for Economic Cooperation and Development) countries 110
bilateral tax treaties, effect of FDI, statistical analysis and results 536
Convention on Mutual Assistance in Tax Matters 543, n. 11
estimated treaty effects on outbound FDI stock and flows 538
high cost of patenting 420, n. 13
International Direct Investment Statistics Yearbook 532
members treaty formation 527
and changes in FDI activity 530
model tax treaty 527–8
new treaties from 1983–92 537
Office of Economics 95
and US, FDI activity differences in effects of old and new treaties 539–42
Ohyama, M. 300
Okediji, Ruth L. 42
OPEC *see* Organization of Petroleum Exporting Counties
open rules 313
Organization for Economic Cooperation and Development *see* OECD
Organization of Petroleum Exporting Counties (OPEC) 354, n. 5
case against, for violation of US antitrust policy 345
overlaps between trade and non-trade objectives 315
ownership in trade 282
ozone layer 163–5, 167
Ozone regime 137, 149, 163–71
case study for international environmental cooperation 164
dispute settlement 168
three compliance strategies 168
see also compliance institutions
ozone-depleting substances (ODSs) 36, n. 120, 148, 163–9
production capacity 167

Pakistan, US complaint against 53
Pallemaerts, Marc 124, n. 33
Palmer, G. 122, n. 28
Panagariya, Arvind
and Krishna
design of welfare improving CUs 310, n. 7
FTA 302–3
Principle of Targeting 505
revenue transfer effects 299
welfare-improving PTAs 305
parallel imports 410–11
Parigi, Giuseppe 470, n. 9
Paris Convention for the Protection of Industrial Property (1883) 404, 414
patentability requirements and priority date 408
patents
breadth 407
economic role of 405–6
effects on trade 404–12
height 407
holder renunciation solution 451
infringed 420, n. 11
and inventive step 407
length 407
in medicines, issues in 417
origin of 405
policy role in the open economy 403
and principles of territoriality and independence 408
protection, market failure argument for granting 405
rights, trade and globalization 403
strategies 407–10

patents (*cont'd*)
strength 403–4, 407
choice of 412–14
minimum international agreements on 414–16
system, effects of trade on choice of 412–18
path dependence in institutional models 281
Paul, K. 470
payment of sanctions 61
Penn World Tables 534–5
Penubarti, Mohan 412
People's Republic of China *see* China
Perroni, Carlo 314–15
Pindyck, R.S. 79, 470, n. 12
Piretz 349
Poland, antidumping in 363
political and economic costs of dispute resolution 47
political economy of cartel policy 351
political justification for antidumping 394
Pomfret, John 502
Porter 367
Porter, Robert H. 367–8
Portes, Richard 290
Portugal
tariffs of 251
US compliant against 53
pragmatism versus legalism 112
Preamble to the Protocol on Acceleration of Free Trade in Goods 400, n. 17
preferential trade agreements *see* PTAs
PTAs (preferential trade agreements)
compatibility with WTO contract 244–61
consistency of interplay between the two tracks 263–4
control of consistency under both Article XXIV GATT and Article V GATS 243
covering both goods and services in the WTO 242–3
CRTA forum to review consistency of notified 244
multilateral review 269
economics of 294
implications and income-redistribution effects 300
external requirement 249–53
for Customs Unions 250–2
formation, legal test under WTO 240–1
and Free Trade Areas (FTAs) 302–4
in the GATS-context
external requirement for 252–3
notifying under 248
impact on the world trading system 307
internal requirement 253–61
limited judicial review by panels 270
model of influence of powerful producers in decision making over entry into a 306
and the multilateral trade system 306–7
necessarily welfare improving 300–6
Pareto-improving 300
related issues
panels dealing with 272
WTO adjudicating bodies 242, 268
revenue transfer effects 299–300
and rules of origin 276, n. 39
substantive requirements 249
and trade barriers 304
two forms of 244
and welfare
analysis 295–300
implications of 239
preferential trading 391
price regulations 417
priority, principle of 408
Prisoners Dilemma game 140, 152, 173, n. 13–16, 367
basic games 139–43
private international law 112
private market activity and public international law concerns 114
process opportunism and TRIPS dispute settlement 42
production decisions of a monopolist 429
protection
of public health and nutrition 423, n. 41
through duties or through quantitative restrictions (QRs) 258
protectionist government policies 204
between the two world wars 186
Protocol of Fortaleza, proposed 392

Prusa, Thomas J.
antidumping 364
analysis of process 341–2, 355, n. 8
and antitrust 382, 387, 390, 394–5
rising activity 357–9, 364, 367
cold rolled steel case study 87, 96–8, 102, n. 28, 31, 32, 103, n. 34, 40
PTA *see* preferential trade agreement
public international law 326
public policy intervention to protect intellectual property rights (IPRs) 403
purgatorial trade agreement 329, n. 3
Putnam, Robert D. 65, n. 42

Radaelli, Claudio M.
tax treaties and FDI 527–30
Raiffa, Howard 316
Rauch, James 285, 290
and Casella, first-generation model of information asymmetry 290
Raustiala, Kal 60
commendation of nonbonding instruments and systems implementation review (SIRs) 60
Harmony game 140
Montreal Protocol Noncompliance Procedure 168, 171, n. 1–2
Razin, Assaf 288
Reciprocal Trade Agreements Act of 1934 186
red-herring distinctions on labor and environmental laws
pragmatism versus legalism 112
public-private 112
Reed, Geoff
AD measures 380
Reformulated Gasoline case 25
Regan, D. 227
regional agreements and regional integration agreements (RAs and RIAs) 264
regional rules of origin 250
regional trade agreements (RTAs) 383–4, 398, 411
for regional AD legislation 399
see also European Union (EU)
regionalism, two paradoxes of 240
regulatory measures 188
and national treatment 208–20
Reichman, J. 51
remedies available under GATT and WTO 179, n. 78
renegotiation market, value of the 52
resource extraction and welfare, short-run effect of trade on 515–16
Rey, Helen 290
Ricardo, David 7–9
Richardson, Martin 348, 355, n. 16
Riezman, Raymond 318
Riley, John G. 288
Rio Declaration on Environment and Development
Principle 7 of 142
risk
associated with trade forms of 288
aversion in counterfeiters 435
to consumption 288
diversification
absence of 287
achieved by networks 289
to income 288
neutrality in counterfeiters 436
Rockfeller, Senator 95–6
Roessler, Frieder
Article XXIV of GATT 267
institutional balance of WTO 270
RTAs 242, 263
Rorschact Test 5
Rose, Nancy L. 476
Rotemberg, Julio 79, 470, n. 12
effects of tariffs and quotas 341
Rousslang, D.J. 470, n. 12
Rowland, Sherwood 164
RTA *see* regional trade agreements
Rubin, Paul 429, 450
rules
definition 50
at the international level, unique costs of negotiating 50
versus standards
debates 56
distinction between 60
Russia, AD law 399, n. 10

Sacheti, S. 502
safeguard measures *see* WTO

Salant, S. 354, n. 4
sales tax benefits 207
Saloner, Garth 34
Sapir, Andre 240–1
Scherer, Frederic M. 350, 407
Schneider, Norman 428
Schnitzer, Monika 283
Schott, Jeffrey 264–5
  unanimous decisions of working parties 274, n. 14
Schwartz, Marius 417
Schwartz, Warren F. 160
Seattle Round of GATT/WTO trade negotiations 120, n. 7, 327
  see also WTO
Sebenius, J.K. 316
second best, trade and theory of 503–5
self-enforcing agreements 328
seller cartels 345
separation regime 313, 320
Serra, J. 278
services, international trade in 474
  commercial presence 474
    (physical) presence of a branch, subsidiary, or affiliate 484
  consumption abroad 474
  cross-border 474
  and goods, dissimilarities and similarities in trade in 473–8
  international trade in temporary presence (presence of natural persons) 475
  international trade in US affiliates data 484–5
  liberalization of trade in 488
  major categories of trade and ten leading country exporters importers 482–3
  multilateral trade negotiations 486–92
  professional 477
  reduced barriers to 485
  ten leading country exporters and importers of goods and 482
  in Uruguay Round 480
  world trade data 480–2
Shapiro, Carl
  markets with counterfeit producers 407, 428–9, 454
Sharp, T.J. 102
Sherman Antitrust Act *see* US
Shrimp/Turtle case 7, 11, 26, 227
  Appellate Body 19, 20
  decision 6, 21
    concern about international coercion 18
    trade free from international coercion 13–16
signals
  communicated at each stage of the DSU process 61
  effectiveness during the dispute resolution 53
  efficacy of 49–55
Simons, Kenneth L. 470, n. 14
1996 Singapore ministerial meeting 324
Siniscalco, D. 329, n. 1
Sinn, H.-W. 529
Skeath, Susan
  antidumping 359, 380, 390
Smith, Adam 7–9
Smith, Pamela J. 412
“soft” negative sanctions 148
software copying 429
South Africa 364
  antidumping in 363
South East Asian nations 240
Spagnolo, Giancarlo 317, 322
Spain
  patent system 420, n. 10
  tariffs of 251
Spencer 346
SPS *see* Agreements
Srinivasan, T.N. 312, 505
Stackelberg Leader Equilibrium 439–40
Staiger, Robert W. 460
  antidumping 364
    model 460
    strategic use of 366, 368
  trade and informal institutions 280
  see also Bagwell–Staiger special protection framework
standards
  definition 50
  in the domestic legislation 56
  negotiated at the international level 56
Stegeman, Klaus 415

Stockholm Declaration of United Nations Conference on the Human Environment 111
Stolarski, Richard 164
Suasion game 140
Subramanian, Arvind 187, 189
substantially all trade, ordinary meaning of the terms 254
Summers, Robert 534
supranational powers, issue of 384
Sykes, Alan O., Jr. 9, 149, 157–8, 160
Syropoulos, C.
  antitrust policy 343
  multilateral cooperation 318
  tariffs 346, 354, n. 7
System-Wide Medium-Term Environment Programme 123, n. 31
systems of implementation review (SIRs) 171, n. 1–2

Targeting, Principle of 503, 505
Tariff Commission, ITC 78
tariff policy
  and multilaterialism 346–7
  optimal 346
  setting as Prisoner's Dilemma 346
tariffs 153, 156–7, 166
  and quotas in a model of Bertrand competition, effects of 341
  reduction 153
  and trust activity link between 350
tax haven countries 529
tax treaties, effect on FDI 527
Taylor, M.S. 502
TEAP *see* Technology and Economic Assessment Panel
Technology and Economic Assessment Panel ("TEAP") 167, 169, 171
Technology and Economic Assessment Panel (TEAP) *see* Montreal Protocol
ten Kate, Adriaan
  on antidumping and antitrust 396, 398
territoriality, principle of 408
textiles sector 118
Thomas, C. 119, n. 4
Tiefer, C. 158
TIPs agreement 116
tit-for-tat (TFT) 358, 366–7
Title VII of the Tariff Act of 1930 74
Tokyo Round of GATT 157, 329, n. 2, 361, 364
  "improvements" in the AD Code 361
  Understanding 155
Töpfer, K. 117
TPRM *see* Trade Policy Review Mechanism
trade 137, 150–4, 156, 158–9, 161–3, 170
  conflict of interest channels 281
  contract enforcement 282–6
  creation 295–300
  diversion 295–300
  effects and environment reform, on the steady states 508
  in goods and trade in services, dissimilarities, and similarities, between 473–8
  and labor
    environment, conflicts 117
    rights 324
  long-run effects environmental reform 519
  measures and global environmental objectives principle of separation between 323
  and monetary policy 325
  negotiations recently launched in Doha, Qatar 327
  restriction 17
    alternative form of 289
    and barriers 16–18
  and security policy 325
  versus nontrade distinction 108–9
Trade Act of 1974 *see* US
Trade Agreements Act of 1979 79
trade barrier 17
  effect on cartel members, single market collusion 343
  global changes with multi-market collusion 343
trade and environment 325
  cooperation, linkages between 319
trade and environment debate 5
  as discussion of proposals for exceptions to free trade principles 24
  refocusing 22–6

trade and environmental reform
  existence of multiple steady states 506
  general model long-run effects of 506–10
  long-run effects of 517–23
  short-run effects of 514–17
  special model 510–14
trade in goods and trade in services
  dissimilarities, intangibility 473–6
trade and informal institutions 279
  predation and enforcement 287–8
trade liberalization
  and AD, strong correlation between 388
  and competition policy 397
    relationship between 398
  and entry 355, n. 16
  and environmental policy harmonization
    dynamic effect of 499
    ideas from economy theory 503–5
  and environmental reform
    discussion 523–4
    general equilibrium model 523
    general model on 506–10
  multilateral approaches to 294
  protectionism paradigm theories 175, n. 40
trade liberalization game 150, 152, 154, 158
trade negotiation 156
trade policies
  effect on collusion 341, 351
  and environmental policies cooperation incentives 319
  formal institutions, national and international 280
  and labor rights, interaction between 324
  level effects on optimal number of firms 355
  and the sustainability of agreements in single markets 341–2
Trade Policy Review Body, ("TPRB") 162
Trade Policy Review Mechanism (TPRM) 137, 161–2
Trade-Related Labor and Environment Rights Agreements 107
trademark holders and counterfeiters in foreign countries 450
transaction costs 147
  modeling 280
treaty proliferation 14
Treaty of Rome (1958) 399–400, n. 15
Trebilcock, Michael J. 318, 325
Trindade, Vitor 290
TRIPS Agreement 42–3, 107, 322, 404
  in age of law 59–60
  Article 30 and Canada 54
  Council 418
  in developing countries, voluntary compliance with 58
  disputes 52
  enforcement provisions of the 56
  in light of International Relations theories 44
  minimum standards for protection of IPRs 415
  and parallel imports 416–17
  patents the 416
  property law 419
  provisions, interpretations of 52
  rules and standards 51
  United States and DSU process against India 59–60
Tulkens, Henry 318
tuna imports, ban on 6
Tuna/Dolphin I 26, 40, n. 198
Tuna/Dolphin II 25–6
Tuna/Dolphin case 7
  decisions 6, 21
    concern about international coercion 18
    trade free from international coercion 13–16
  panels 23
    dispute between Mexico and Canada over extraterritorial application of US Marine Mammal Protection Act 323
Turkey
  Restrictions on Imports of Textile and Clothing Products case 270–1
    panel report on 256, 268, 273

Ulen, Thomas 164
UN World Investment Directory 535
UNCTAD 314
UNDP 314

UNIDROIT Principles (Principles of International Commercial Contracts) 126, n. 44
unilateral trade restrictions 14
unintentional noncompliance 142, 147
United Kingdom
 no antitrust legislation 349
 policy on cartels 350
 Restrictive Trade Practices Act of 1956 350
United Nations
 study of effects of bilateral treaties for promotion and protection of FDI 543–4, n. 16
United Nations Conference on Environment and Development (UNCED) 111, 123, n. 31
 Rio Declaration on Environment and Development, Principle 7 174, n. 21
United Nations Convention on the Law of the Sea (LOS Convention) 62, n. 9
United Nations Environment Program (UNEP) 111, 123, n. 29, 128, n. 60, 133, n. 108, 165
 Ad Hoc Working Group of Legal and Technical Experts 165
 Töpfer, Klaus 117
 see also ozone regime
United Nations Human Rights Commission 128, n. 60
United States 15, 364
 antidumping 392
  and countervailing duty (AD/CVD) law 79, 400, n. 22
  disputes 359
  law 394
 Antidumping Act (1921) 360
 antitrust case against Microsoft 459
 ban on tuna imports 15
 bilateral tax treaty activity 540
 Bureau of Economic Activity (BEA) 540
 and Canadian Merger Enforcement Guidelines 224
 Clayton Antitrust Act 349
 complaints
  Denmark and Sweden over failure to enact provisional measures as required by TRIPS Article 50, 63, and 65, 53
  Greece and European Union over lack of copyright infringement remedies in Greece 53
  Pakistan, failure to comply with TRIPS Articles 27, 65, and 70, 53
  Portugal, failure to implement TRIPS provisions with regard to patent terms 53
 Copyright Act, section 110(5) of the 47–8
 cotton trade 285
 Cuban Liberty and Democratic Solidarity (LIBERTAD) Act of 1996 (Helms-Burton Act) 325
 Cyprus treaty 529
 Definitive Safeguard Measures on Imports of Circular Welded Carbon Quality Line Pipe from Korea 269
 Department of Commerce (DOC) 75
 Department of Justice (DOJ) 333
 deregulation or regulatory reform efforts in trade in services 476
 discriminating against Mexican tuna 19
 dumping law 93
 embargo against Nicaragua 325
 and European Union over section 110(5) of US Copyright Act, dispute 46
 Food, Drug, and Cosmetic Act of 1938 25
 German treaty 529
 goods imports 482
 hegemony, role of 159
 Import Milk Act of 1927 25
 international obligations under WTO agreements 76
 International Trade Commission (USITC) 73, 280, 473
 law
  injury investigations under section 201
  international obligations 77–8
 Malt Beverages case GATT case 194–5
 Omnibus Trade Act (1988) 361

United States (*cont'd*)
Omnibus Trade and Competitiveness Act (1998) 361, 415
policy
on cartels 349
consistent with the national welfare model 349
Poultry Products Inspection Act 25
reliance on unilateral trade restrictions 14
Restrictions on Imports of Tuna, case 193
Revenue Act of 1916 360
Safeguarding of Industries Act of 1921 360
services
exports 482
net exporter of 485
trade data 482–4
Sherman Antitrust Act of 1890 348–9, 351, 360, 399, n. 3
Tariff Act of 1930 361
section 337 of 192
title VII of the 75
and the Trade Agreements Acts of 1974 and 1979 361
tax definitions, 1986 Tax Reform Act revised 528
Tax Treatment for "Foreign Sales Corporations", Recourse to Article 21.5 of the DSU by the European Communities 237, n. 101–4
Taxes and Automobiles, GATT panel decision in 195–6
Textile Quota Program 132, n. 99
Trade Act of 1974
US Law section 201 of the 77
US section 301 14, 415
trade deficit
with Asian countries 428; with China 456, n. 2
in homogeneous products like agriculture 289
Trade Representative (USTR) investigation 415
Trade and Tariff Act 1984 361, 415
treaty with Pakistan 543, n. 8
and United Kingdom, policies toward price-fixing in 348
versus American Tobacco 343, 345
versus Trenton Potteries 356
wheat gluten industry 81
Uruguay Round of GATT 14, 50, 108, 113, 157, 159, 254, 327, 364
Agreements
on AD 361
on PTA 273
textiles and agriculture under trade-liberalizing rules in the 118
IP protection, multilateral instruments for 57, 109, 415
negotiations
freer trade in agriculture and textiles 57
on relationship of Article XIX GATT with Article XXIV GATT 260, 267, 271
in trade in services 486
trade negotiations 7, 315
Understanding of the Balance-of-Payments Provisions of the GATT 330
USF (upstream firms) and DSF (downstream firms) 286

van Dijk, Theon W. P. 407
van Wincoop, Eric 279, 290
Vanek, J. 300
Verdier, Thierry 460
Verhoosel, G. 223
integrated necessity test under Article III and Article XX 223
proposals 227
Victor, David 135, 168
Vienna Convention 49, 55–6
for the Protection of the Ozone Layer 165
Article 11 of 168
see also ozone regime
Vienna Convention of the Law of Treaties (VCLT) 66, n. 55, 250–1, 253–4, 260
Viner, Jacob 402
antidumping 381, n. 2
as remedy for predatory behavior 385–6, 399, n. 2

PTAs, analysis of 295, 306–7
work on trade creation/trade diversion 240
Vousden, Neil 460

Walker, Vern R.
on scientific uncertainty and WTO fact-finding 33, n. 77
Wan, H. 307, 310
see also Kemp-Wan model
Washington Convention on International Trade in Endangered Species of Wild Fauna and Flora 122
Watal, Jayashree 423, n. 41
Wathen, Tom
definition of free trade 27, n. 6
WCO 314
Webb Pomerene Act 355, n. 13
welfare implications of PTAs 241
Whalley, John 314
Wheat Gluten case 81
Wheeler, D. 502
Whinston, Michael D. 342
White, Lawrence J.
examination of international trade in accounting services 475, 478, 489
Wilson, G.P. 527
Winston, C. 476
Winters, A. 312
WIPO *see* World Intellectual Property Organization
Wolak, Frank A. 364
Wolak, Frank S. 460
Wooton, Ian 394
workers rights 315
Working Party report on
EC, Agreements with Portugal 257
EEC, Association with African and Malgasy States, 1970 259
report on EFTA 257
World Bank 389, 534
World Health Organization Model List of Essential Drugs 423, n. 42
World Intellectual Property Organization (WIPO) 109, 118, 314, 414–15
World Trade Organization (WTO) 6, 25, 113, 136, 146, 148–9, 152, 158–62, 169–71, 280, 294, 314
agreements
Agreement on Agriculture 132, n. 104
Agreement on Trade-Related Property Rights *see* TRIPS
Antidumping (AD) Agreement 75–6, 78, 394
breach as a resort to protectionism 59
Safeguards Agreement 77–8, 82
Annex 1
GATT 1994, the General Agreement on Trade in Services (“GATS”), and the Agreement on Trade-Related Aspects of Intellectual Property Rights (“TRIPS”) 159
Annex 2
dispute settlement rules 159
Annex 4
“plurilateral agreements” 159
Beef/Hormone Appellate Body 12
Canada
Certain Measures Affecting the Automotive Industry 235, n. 88–9
Certain Measures Concerning Periodicals 232–3, n. 60–70
capacity problems 160–2, 170
case on National Treatment principles issues of facially neutral nontax regulatory measures 209
case on National Treatment principles, Japan, Measures Affecting Consumer Photographic Film and Paper 209
Committee on Trade and Environment 323
Committee on Trade and Environment study 116
Community/Forums Page 129, n. 68
and compliance 136–7, 146, 148, 153, 158–9, 161–3, 166, 169, 171
contract (Article XXIV GATT and Article V GATS) 240
Council Trade in Services 473
creation of the 153
definition of like products in Article III:2 199

World Trade Organization (WTO) (*cont'd*)
Dispute Settlement Body (DSB) of the 218
dispute settlement system 99, 127, n. 53, 148, 160–1, 208
as compliance institution 136
Dispute Settlement Understanding (DSU) 42
shortcomings 81
Type II and Type III strategies 161
dispute settlement system, facially neutral tax or regulatory measures and the principle of national treatment under the 196–220
economic approach to the definition of like products proposal for 224
European Communities, Measures Affecting Asbestos and Asbestos-Containing Products 236, n. 94–7
European Communities, Regimes for the Importation, Sale and Distribution of Bananas 235, n. 86
expanding versus controlling international trade 115–17
facially neutral regulatory measure 214
Fourth Ministerial Conference in Doha, Qatar 161
and game theory 148–9, 152, 158–9
institutional balance in 270–2
institutional competence" justification 107
and institutions 146, 148, 152–3, 158–60, 166
internal requirement issue 271
and IP rights and labor and environmental rights red-herring distinctions 112–15
Japan, Taxes on Alcoholic Beverages 230, n. 45
judicial and scholarly efforts to justify the 18–26
Korea
Measures Affecting Imports of Fresh, Chilled and Frozen Beef 236, n. 90–3
Taxes on Alcoholic Beverages 233, n. 71
Law, Advisory Centre for 69, n. 110
legal protection 358
and other legal regimes, tension between the 6
legitimacy 26
legitimate" protectionism and AD 359
likeness" of domestic and imported products 212
MEAs currently in force outside the 323
Members deciding to enter into PTA, obligation to notify 245–7
ministerial conferences
considering labor and environmental issues 119–20, n. 5
in Doha on AD law 395
in Seattle, disruptions, and GATS 490
National Treatment principle 212
package of agreements 159
panel
aim-and-effect test in regard to Article III 231
India, Automotive case 219
internal tax and regulatory measures 196
issue of less favorable treatment 219
recent 2003 Technical Assistance/Capacity Building Plan 161
red herring distinctions
market versus nonmarket or economic versus social 114
open versus closed forum 114
rigidity versus flexibility 114
regime, trade measures 14
science policy 33, n. 77
Seattle meeting 115
in the Shrimp/Turtle case 17
Singapore meeting in 115
Secretariat 162
separation regime between trade measures and labor rights issues 324
TBT and SPS Agreements 227
tests
to determine the validity of 212

for Japan, Taxes on Alcoholic Beverages 214
for likeness of products, US – Gasoline 214
Trade Policy Review Mechanism 159
trading system 159
and TRIPS, exclusion of labor and environment standards 111–19
see also TRIPS
underenforcement problem common to IP, Labor and Environmental Regimes 109–11
understanding on Article XXIV GATT 251, 259
WTO Appellate Body
decisions in wheat gluten, lamb meat, and hot-rolled steel 99
Japan
Report on Measures Affecting Agricultural Products 34, n. 95
Taxes on Alcoholic Beverages 202, 208
rulings in wheat gluten, lamb meat, line pipe, and hot-rolled steel cases 85–94
in Korea – Measures on Beef 212–13
and panel rejection in the Japanese Alcohol case of the aims-and-effects test 221
WTO dispute settlement 148, 158–62, 169–71
Dispute Settlement Understanding, (“DSU”) 159–60
WTO regime 137, 152–3, 158, 160, 169–71
see also GATT/WTO regime, WTO
WTO trade assistance 153, 158–9, 161-2

Yarbrough, Beth V. 149, 153
Yarbrough, Robert M. 149, 153
Yi, S. 307, 318
Young, Kan H. 484
Young, Leslie
formal enforcement 283–4

Zanardi, Maurizio 402
AD laws and WTO membership, countries with 385–6
antitrust policy 355, n. 8
Zeeuw, Aart de 316
Zhao, Jinhua 499, 502