Studies in Choice and Welfare

Prasanta K. Pattanaik
Koichi Tadenuma
Yongsheng Xu
Naoki Yoshihara
Editors

Rational Choice and Social Welfare

Theory and Applications

Springer

Studies in Choice and Welfare

Kotaro Suzumura

Prasanta K. Pattanaik • Koichi Tadenuma
Yongsheng Xu • Naoki Yoshihara
Editors

Rational Choice and Social Welfare

Theory and Applications

Essays in Honor of Kotaro Suzumura

Springer

Editors

Professor Prasanta K. Pattanaik
Department of Economics
University of California
Riverside
CA, 92521
U.S.A
prasanta.pattanaik@ucr.edu

Professor Koichi Tadenuma
Department of Economics
Hitotsubashi University
2-1 Naka, Kunitachi
Tokyo 186-8601
Japan
tadenuma@econ.hit-u.ac.jp

Professor Yongsheng Xu
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992
Atlanta, GA 30302-3992
U.S.A.
yxu3@gsu.edu

Professor Naoki Yoshihara
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi
Tokyo 186-8603
Japan
yosihara@ier.hit-u.ac.jp

ISBN 978-3-540-79831-6 ISBN 978-3-540-79832-3 (eBook)

Studies in Choice and Welfare ISSN 1614-0311

Library of Congress Control Number: 2008931049

Cover design: WMX Design GmbH, Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

This volume brings together papers, which were first presented at the International Conference on Rational Choice, Individual Rights and Non-Welfaristic Normative Economics, held in honour of Kotaro Suzumura at Hitotsubashi University, Tokyo, on 11–13 March 2006, and which have subsequently gone through the usual process of review by referees. We have been helped by many individuals and institutions in organizing the conference and putting this volume together. We are grateful to the authors of this volume for contributing their papers and to the referees who reviewed the papers. We gratefully acknowledge the very generous fundings by the Ministry of Education, Culture, Sports, Science and Technology, Japan, through the grant for the 21st Century Center of Excellence (COE) Program on the Normative Evaluation and Social Choice of Contemporary Economic Systems, and by the Japan Society for the Promotion of Science, through the grant for International Scientific Meetings in Japan, and the unstinted effort of the staff of the COE Program at Hitotsubashi University, without which the conference in 2006 would not have been possible. We thank Dr. Martina Bihn, the Editorial Director of Springer-Verlag for economics and business, for her advice and help.

Finally, we would like to mention that it has been a great pleasure and privilege for us to edit this volume, which is intended to be a tribute to Kotaro Suzumura's immense intellectual contributions, especially in the theory of rational choice, welfare economics, and the theory of social choice.

Riverside	*Prasanta K. Pattanaik*
Tokyo	*Koichi Tadenuma*
Atlanta	*Yongsheng Xu*
Tokyo	*Naoki Yoshihara*
2007	

Professor Kotaro Suzumura Bibliographical Data

Books in English

Rational Choice, Collective Decisions and Social Welfare, Cambridge University Press, 1983.
Industrial Policies for Pacific Economic Growth, George Allen and Unwin, 1986. Joint editor with H. Mutoh, S. Sekiguchi, and I. Yamazawa.
Industrial Policy of Japan, Academic Press, 1988. Joint editor with R. Komiya and M. Okuno.
The Economic Theory of Industrial Policy, Academic Press, 1991. Joint author with M. Itoh, K. Kiyono, and M. Okuno-Fujiwara.
Choice, Welfare, and Development: A Festschrift in Honor of Amartya K. Sen, Oxford University Press, 1995. Joint editor with K. Basu and P. K. Pattanaik.
Competition, Commitment, and Welfare, Oxford University Press, 1995.
Social Choice Re-examined, Macmillan, 2 vols., 1996/1997. Joint editor with K. J. Arrow and A. K. Sen.
Development Strategy and Management of the Market Economy, Oxford University Press, 1997. Joint author with E. Malinvaud, J.-C. Milleron, M. Nabli, A. K. Sen, A. Sengupta, N. Stern, and J. E. Stiglitz.
Handbook of Social Choice and Welfare, Elsevier, 2 vols., 1st vol., 2002; 2nd vol., forthcoming in 2008. Joint editor with K. J. Arrow and A. K. Sen.
Intergenerational Equity and Sustainability, Elsevier, 2007. Joint editor with J. Roemer.
Choice, Opportunities, and Procedures: Selected Papers on Social Choice and Welfare, Basil Blackwell, forthcoming in 2008.

Articles in English

"Balanced Growth Properties of Generalized Dynamic Leontief Model," *Economic Studies Quarterly*, Vol. 20, 1969, pp. 64–74. Joint paper with S. Nabeya.
"Price Divergence and Gains from Trade in Leontief Model with Variable Coefficients," *Economic Studies Quarterly*, Vol. 21, 1970, pp. 60–66.
"International Equalization of Factor Prices in Leontief Model with Variable Coefficients," *Metroeconomica*, Vol. 23, 1971, pp. 1–15.
"Uniqueness of the Cournot Oligopoly Equilibrium," *Economic Studies Quarterly*, Vol. 22, 1971, pp. 81–83. Joint paper with K. Okuguchi.
"The Economic Theory of Organization and Planning: A Review Article," *Economic Studies Quarterly*, Vol. 24, 1973, pp. 33–51.
"Professor Uzawa's Equivalence Theorem," *Economic Studies Quarterly*, Vol. 24, 1973, pp. 60–70.

"Boundedness of the Closed Economy with Samuelson-Leontief Technology," *Hitotsubashi Journal of Economics*, Vol. 13, 1973, pp. 43–46.
"On Acyclic Preferences," *Economic Review*, Vol. 26, 1975, pp. 273–275.
"Rational Choice and Revealed Preference," *Review of Economic Studies*, Vol. 43, 1976, pp. 149–158.
"Remarks on the Theory of Collective Choice," *Economica*, Vol. 43, 1976, pp. 381–390.
"Impossibility Theorems without Collective Rationality," *Journal of Economic Theory*, Vol. 13, 1976, pp. 361–379. Joint paper with D. H. Blair, G. Bordes, and J. S. Kelly. Reprinted in K. J. Arrow and G. Debreu, eds., *The Foundations of 20th Century Economics*, Vol. 3, *Landmark Papers in General Equilibrium Theory, Social Choice, and Welfare*, Cheltenham, Glos.: Edward Elgar, 2001, pp. 660–678.
"Houthakker's Axiom in the Theory of Rational Choice," *Journal of Economic Theory*, Vol. 14, 1977, pp. 284–290.
"On the Consistency of Libertarian Claims," *Review of Economic Studies*, Vol. 45, 1978, pp. 329–342. "A Correction," *Review of Economic Studies*, Vol. 46, 1979, p. 743.
"On Distributional Value Judgments and Piecemeal Welfare Criteria," *Economica*, Vol. 47, 1980, pp. 125–139.
"Liberal Paradox and the Voluntary Exchange of Rights-Exercising," *Journal of Economic Theory*, Vol. 22, 1980, pp. 407–422. Reprinted in C. K. Rowley, ed., *The International Library of Critical Writings in Economics*, Vol. 27, *Social Justice and Classical Liberal Goals*, Cheltenham, Glos.: Edward Elgar, 1993, pp. 483–498.
"On the Possibility of 'Fair' Collective Choice Rule," *International Economic Review*, Vol. 22, 1981, pp. 307–320.
"On Pareto-Efficiency and the No-Envy Concept of Equity," *Journal of Economic Theory*, Vol. 25, 1981, pp. 367–379.
"Collective Choice Rules without the Pareto Principle," *International Economic Review*, Vol. 23, 1982, pp. 299–308. Joint paper with J. Fountain.
"Equity, Efficiency and Rights in Social Choice," *Mathematical Social Sciences*, Vol. 3, 1982, pp. 131–155.
"Resolving Conflicting Views of Justice in Social Choice," in P. K. Pattanaik and M. Salles, eds., *Social Choice and Welfare*, Amsterdam: North-Holland, 1983, pp. 125–149.
"Perron-Frobenius Theorem on Non-Negative Square Matrices: An Elementary Proof," *Hitotsubashi Journal of Economics*, Vol. 24, 1983, pp. 137–141.
"Equity and Efficiency in the Public Goods Economy: Some Counterexamples," *Hitotsubashi Journal of Economics*, Vol. 26, 1985, pp. 59–82. Joint paper with K. Sato.
"Gibbardian Libertarian Claims Revisited," *Social Choice and Welfare*, Vol. 3, 1986, pp. 61–74. Joint paper with K. Suga.
"An Economic Analysis of Industrial Policy: A Conceptual Framework Through the Japanese Experience," in H. Mutoh, S. Sekiguchi, K. Suzumura, and I. Yamazawa, eds., *Industrial Policies for Pacific Economic Growth*, George Allen and Unwin, 1986, pp. 23–41. Joint paper with M. Okuno.
"Entry Barriers and Economic Welfare," *Review of Economic Studies*, Vol. 54, 1987, pp. 157–167. Joint paper with K. Kiyono.
"Arrow and the Problem of Social Choice," in G. Feiwel, ed., *Arrow and the Foundations of the Theory of Economic Policy*, Macmillan, 1987, pp. 255–271. Joint paper with K. Suga.
"Industrial Policy in Japan: Overview and Evaluation," in R. Sato and P. Wachtel, eds., *Trade Friction and Economic Policy*, Cambridge University Press, 1987, pp. 50–79. Joint paper with M. Okuno-Fujiwara.
"Social Welfare Function," in J. Eatwell, M. Milgate, and P. Newman, eds., *The New Palgrave: A Dictionary of Economics*, Vol. 4, London: Macmillan, 1987, pp. 418–420.
"Reduction of Social Choice Problems: A Simple Proof of Arrow's General Possibility Theorem," *Hitotsubashi Journal of Economics*, Vol. 29, 1988, pp. 219–221.
"Strategic Information Revelation," *Review of Economic Studies*, Vol. 57, 1990, pp. 25–47. Joint paper with M. Okuno-Fujiwara and A. Postlewaite.

"Oligopolistic Competition and Economic Welfare: A General Equilibrium Analysis of Entry Regulation and Tax-Subsidy Schemes," *Journal of Public Economics*, Vol. 42, 1990, pp. 67–88. Joint paper with H. Konishi and M. Okuno-Fujiwara.

"Dictator Lemma: Detailed Proof," *Hitotsubashi Journal of Economics*, Vol. 31, 1990, pp. 35–36.

"Endogenous Preferences: Through the Looking-Glass of Social Choice Theory," in G. Brennan and C. Walsh, eds., *Rationality, Individualism and Public Policy*, Center for Research on Federal Financial Relations, 1990, pp. 34–39.

"Alternative Approaches to Libertarian Rights in the Theory of Social Choice," in K. J. Arrow, ed., *Issues in Contemporary Economics*, Vol. 1, *Markets and Welfare*, London: Macmillan, 1990, pp. 215–242.

"On the Voluntary Exchange of Libertarian Rights," *Social Choice and Welfare*, Vol. 8, 1991, pp. 199–206.

"Alternative Firm Objectives and the Welfare Effects of Entry," *Osaka Economic Papers: Special Issue in Honor of Professor H. Shibata*, Vol. 40, 1991, pp. 194–198. Reprinted in Ali M. El-Agraa, ed., *Public and International Economics*, London: Macmillan, 1993.

"Individual Rights Revisited," *Economica*, Vol. 59, 1992, pp. 161–177. Joint paper with W. Gaertner and P. K. Pattanaik. Reprinted in C. K. Rowley, ed., *The International Library of Critical Writings in Economics*, Vol. 27, *Social Justice and Classical Liberal Goals*, Cheltenham, Glos.: Edward Elgar, 1993, pp. 592–608.

"Taxation and Welfare in Oligopoly with Strategic Commitment," *International Economic Review*, Vol. 33, 1992, pp. 413–431. Joint paper with T. Besley.

"Cooperative and Noncooperative R&D in Oligopoly with Spillovers," *American Economic Review*, Vol. 82. 1992, pp. 1307–1320.

"Symmetric Cournot Oligopoly and Economic Welfare: A Synthesis," *Economic Theory*, Vol. 3, 1993, pp. 43–59. Joint paper with M. Okuno-Fujiwara.

"Cooperative and Non-Cooperative R&D in an Oligopoly with Spillovers: Strategic Substitutes Versus Strategic Complements," *Hitotsubashi Journal of Economics*, Vol. 34, 1993, pp. 1–11. Joint paper with N. Yanagawa.

"Rights, Welfarism and Social Choice," *American Economic Review: Papers and Proceedings*, Vol. 84, 1994, pp. 435–439. Joint paper with P. K. Pattanaik.

"Industrial Policy of Japan: Whither Now?" in C. Filippini, ed., *The Italian and Japanese Economies in the '80s*, Milano: SGEA, 1994, pp. 64–95.

"On Symmetric Agents: Comments on Kemp and Shimomura," *Japanese Economic Review*, Vol. 46, 1995, pp. 295–298. Joint comment with M. Kaneko.

"Further Comments," *Japanese Economic Review*, Vol. 46, 1995, p. 300. Joint comment with M. Kaneko.

"Entry and Cost Reduction: Comment," *Japan and the World Economy*, Vol. 7, 1995, pp. 411–418.

"Individual Rights and Social Evaluation: A Conceptual Framework," *Oxford Economic Papers*, Vol. 48, 1996, pp. 194–212. Joint paper with P. K. Pattanaik.

"Welfare, Rights and Social Choice Procedure: A Perspective," *Analyse & Kritik*, Vol. 18, 1996, pp. 20–37.

"On the Social Responsibility of Corporations under Competition," in B. F. Neil, ed., *Ethical Universals and International Business*, Berlin: Springer Verlag, 1996, pp. 40–57.

"Oligopolistic Competition and Economic Welfare: The Effects of Ownership Structure," in R. Sato, R. Ramachandran, and H. Hori, eds., *Organization, Performance and Equity: Perspectives in the Japanese Economy*, Hingham: Martinus Nijhoff Publishing, 1996, pp. 93–107.

"Voluntary Export Restraints and Economic Welfare," *Japanese Economic Review*, Vol. 48, 1997, pp. 176–186. Joint paper with J. Ishikawa.

"Japan's Industrial Policy and Accession to the GATT: A Teacher by Positive or Negative Examples?" *Hitotsubashi Journal of Economics*, Vol. 38, 1997, pp. 101–123.

"Collaborative R&D and Competition Policy: Economic Analysis in the Light of Japanese Experience," in W. Comanor, A. Goto and L. Waverman, eds., *Competition Policy in a Global Economy: Modalities for Co-operation*, London: Routledge & Kegan Paul, 1997, pp. 197–223. Joint paper with A. Goto.

"Keiretsu – Inter-Firm Relationship in Japan" in W. Comanor, A. Goto and L. Waverman, eds., *Competition Policy in a Global Economy: Modalities for Co-operation*, London: Routledge & Kegan Paul, 1997, pp. 361–378. Joint paper with A. Goto.

"Interpersonal Comparisons of the Extended Sympathy Type and the Possibility of Social Choice," in K. J. Arrow, A. K. Sen and K. Suzumura, eds., *Social Choice Re-examined*, Vol. 2, London: Macmillan, 1997, pp. 200–227.

"Japan's Industrial Policy for Telecommunications: The 1985 Institutional Reform and its Aftermath," in K. Odaka and J. Teranishi, eds., *Markets and Government: In Search of Better Coordination*, Tokyo: Kinokuniya Publishing Co., 1998, pp. 263–303.

"Consequences, Opportunities and Procedures," *Social Choice and Welfare*, Vol. 16, 1999, pp. 17–40.

"Paretian Welfare Judgments and Bergsonian Social Choice." *Economic Journal*, Vol. 109, 1999, pp. 204–220. Reprinted in G. C. Wood and M. McLure, eds., *Paul Samuelson: Critical Assessment of Contemporary Economists*, Vol. 2, 2nd Series, London: Routledge, 2004, pp. 378–396.

"Welfare Economics and the Welfare State," *Review of Population and Social Policy*, No. 8, 1999, pp. 119–138.

"Consequentialism and Procedural Evaluations in Social Welfare Judgments," in R. Sato, R. V. Ramachandran, and K. Mino, eds., *Global Competition and Integration*, Boston: Kluwer Academic Publishers, 1999, pp. 65–81.

"Comments on 'Regulatory Priorities for Infrastructure Reform in Developing Countries,' by P. L. Joskow," in B. Pleskovic and J. E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998*, Washington, D.C.: The World Bank, 1999, pp. 228–232.

"Welfare Economics Beyond Welfarist-Consequentialism," *Japanese Economic Review*, Vol. 51, 2000, pp. 1–32.

"Competition and Regulation in Telecommunications: Japan's Reform Experience Since 1985," *Hitotsubashi Journal of Economics*, Vol. 40, 2000, pp. 1–28. Joint paper with Y. Yasaki.

"On the Rule-Oriented Approach to Fair Trade," *Journal of Japanese Trade and Industry*, July/August 2000, pp. 26–29.

"Pareto Principles from Inch to Ell," *Economics Letters*, Vol. 70, 2001, pp. 95–98.

"An Interview with Miyohei Shinohara: Nonconformism in Japanese Economic Thought," *Journal of the Japanese and International Economies*, Vol. 15, 2001, pp. 341–360. Joint paper with A. Amsden.

"Characterizations of Consequentialism and Non-Consequentialism," *Journal of Economic Theory*, Vol. 101, 2001, pp. 423–436. Joint paper with Y. Xu.

"Upper Semicontinuous Extensions of Binary Relations," *Journal of Mathematical Economics*, Vol. 37, 2002, pp. 231–246. Joint paper with W. Bossert and Y. Sprumont.

"Recoverability of Choice Functions and Binary Relations: Some Duality Results," *Social Choice and Welfare*, Vol. 21, 2003, pp. 21–37. Joint paper with Y. Xu.

"On Constrained Dual Recoverability Theorems," *Mathematical Social Sciences*, Vol. 45, 2003, pp. 143–154. Joint paper with Y. Xu.

"Consequences, Opportunities, and Generalized Consequentialism and Non-Consequentialism," *Journal of Economic Theory*, Vol. 111, 2003, pp. 293–304. Joint paper with Y. Xu.

"Welfarist-Consequentialism, Similarity of Attitudes, and Arrow's General Impossibility Theorem," *Social Choice and Welfare*, Vol. 22, 2004, pp. 237–251. Joint paper with Y. Xu.

"Extended Social Ordering Functions for Rationalizing Fair Allocation Rules as Game Forms in the Sense of Rawls and Sen," *International Journal of Economic Theory*, Vol. 1, 2005, pp. 21–41. Joint paper with R. Gotoh and N. Yoshihara.

"Consistent Rationalizability," *Economica*, Vol. 72, 2005, pp. 185–200. Joint paper with W. Bossert and Y. Sprumont.

"Arrovian Aggregation in Economic Environments: How Much Should We Know About Indifference Surfaces?" *Journal of Economic Theory*, Vol. 124, 2005, pp. 22–44. Joint paper with M. Fleurbaey and K. Tadenuma.

"The Informational Basis of the Theory of Fair Allocation," *Social Choice and Welfare*, Vol. 24, 2005, pp. 311–341. Joint paper with M. Fleurbaey and K. Tadenuma.

"Maximal-Element Rationalizability," *Theory and Decision*, Vol. 58, 2005, pp. 325–350. Joint paper with W. Bossert and Y. Sprumont.

"An Interview with Paul Samuelson: Welfare Economics, 'Old' and 'New', and Social Choice Theory," *Social Choice and Welfare*, Vol. 25, 2005, pp. 327–356.

"Competition, Welfare, and Competition Policy," in U. Schmidt and S. Traub, eds., *Advances in Public Economics: Utility, Choice and Welfare*, Netherlands: Springer, 2005, pp. 1–15.

"Rationalizability of Choice Functions on General Domains Without Full Transitivity," *Social Choice and Welfare*, Vol. 27, 2006, pp. 435–458. Joint paper with W. Bossert and Y. Sprumont.

"Shigeto Tsuru (1912–2006): Life, Work and Legacy," *European Journal of the History of Economic Thought*, Vol. 13, 2006, pp. 613–620.

"Competition, Regulation and Welfare in the Presence of Network Externalities," *Japanese Economic Review*, Vol. 58, 2007, pp. 47–70. Joint paper with A. Soubeyran and S. Weber.

"Ordering Infinite Utility Streams," *Journal of Economic Theory*, Vol. 135, 2007, pp. 579–589. Joint paper with W. Bossert and Y. Sprumont.

"Domain Closedness Conditions and Rational Choice," *Order*, Vol. 24, 2007, pp. 75–88. Joint paper with W. Bossert.

"Non-Deteriorating Choice Without Full Transitivity," Analyse & Kritik, Vol. 29, 2007, pp. 163–187. Joint paper with W. Bossert.

"A Characterization of Consistent Collective Choice Rules," *Journal of Economic Theory*, Vol. 138, 2008, pp. 311–320. Joint paper with W. Bossert.

"Continuity and Egalitarianism in the Evaluation of Infinite Utility Streams," *Social Choice and Welfare*, Vol. 31, 2008, 179–191. Joint paper with C. Hara, T. Shinotsuka, and Y. Xu.

"Rational Choice on General Domains", in K. Basu and R. Kanbur, eds., *Arguments for a Better World: Essays in Honor of Amartya Sen, Vol. I: Ethics, Welfare and Measurement*, Oxford: Oxford University Press, forthcoming. Joint paper with W. Bossert.

"Consequentialism and Non-Consequentialism: Axiomatic Approach," in C. Puppe, P. Anand and P. K. Pattanaik, eds., *Handbook of Decision Theory and Social Choice*, Oxford: Oxford University Press, forthcoming. Joint paper with Y. Xu.

Professor Kotaro Suzumura Biographical Data

Education

Hitotsubashi University, Tokyo (B.A. in Economics 1966, M.A. in Economics 1968)
Hitotsubashi University (Doctoral Course in Economics, 1968–1971)
Hitotsubashi University (Doctor in Economics, 1980)
Doctor Thesis: *Rational Choice, Collective Decisions and Social Welfare*, New York: Cambridge University Press, 1983

Professional Appointments

Lecturer, Department of Economics, Hitotsubashi University, 1971–1973
Associate Professor, Institute of Economic Research, Kyoto University, 1973–1982
British Council Visiting Scholar, Department of Economics, Cambridge University, 1973–1974
Lecturer, Department of Economics, London School of Economics and Political Science, 1974–1976
Visiting Associate Professor, Department of Economics, Stanford University, 1979–1980
Associate Professor, Institute of Economic Research, Hitotsubashi University, 1982–1984
Professor, Institute of Economic Research, Hitotsubashi University, 1984–2008
Visiting Fellow, Faculty of Economics and Commerce and Research School of Social Sciences, Australian National University, 1986
Visiting Fellow, Department of Economics, University of Pennsylvania, 1987
Visiting Fellow, All Souls College, Oxford University, 1988
Director General, Tokyo Center for Economic Research, 1990–1992
Visiting Professor, Contemporary Japan Centre and Department of Economics, University of Essex, 1990–1991
Fulbright Senior Research Fellow, Department of Economics, Harvard University, 1993
Visiting Professor, Department of Economics, University of British Columbia, 1994
Visiting Professor, Institute of Public Economics, University of Graz, 1994
Associate Editor, *Economic Design*, 1994–1998
Associate Editor, *Economics and Philosophy*, 1995–
Nissan Visiting Fellow, Nissan Institute of Japanese Studies, St. Antony's College, Oxford University, 1996
Visiting Professor, Ecole des Hautes Etudes en Sciences Sociales, 1998
Visiting Professor, Department of Economics, Southern Methodist University, 1999

Visiting Professor, Centre for Philosophy of Natural and Social Sciences, London School of Economics and Political Science, 1999
Visiting Fellow Commoner, Trinity College, Cambridge University, 2001
Visiting Professor, Department of Economics, University of Montreal, 2001
Director, Competition Policy Research Center, Fair Trade Commission of Japan, 2003–2008
Visiting Professor (Full-time), School of Political Science and Economics, Waseda University, 2008–

Professional Elections

Fellow, Econometric Society, 1990–
Secretary General, Japan Association of Economics and Econometrics, 1993–1996
Council Member, Econometric Society, 1995–2000
Chair, Far East Standing Committee of the Econometric Society, 1995–2000
Chair, Sub-Committee on Unfair Trade Policies and Measures, WTO Committee, Industrial Structure Council, Japan, 1996–2001
President-elect, Society for Social Choice and Welfare, 1998–1999
President, Japanese Economic Association, 1999–2000
Executive Committee Member, International Economic Association, 1999–2005
President, Society for Social Choice and Welfare, 2000–2001
Member, Science Council of Japan, 2000–2003; 2005–
Vice-President, Science Council of Japan, 2006–

Professional Awards

Nikkei Economics Book Prize (*Rational Choice, Collective Decisions and Social Welfare*, Cambridge University Press, 1983), 1984
Nikkei Economics Book Prize (*The Economic Analysis of Industrial Policy*, The University of Tokyo Press, in Japanese, 1988; joint author with M. Itoh, K. Kiyono, and M. Okuno-Fujiwara), 1988
Medal with Purple Ribbon, 2004
Japan Academy Prize, 2006

Editorial Responsibilities

Editor, *Economic Studies Quarterly* (Journal of the Japan Association of Economics and Econometrics), 1984–1987
Editor, *Social Choice and Welfare*, 1984–
Book Review Editor, *Journal of the Japanese and International Economies*, 1986–1991
Associate Editor, *Journal of Industrial Economics*, 1986–1995
Editor, *Hitotsubashi Journal of Economics*, 1989–1993
Board of Editors, *Economic Record*, 1990–
Editor, *Journal of the Japanese and International Economies*, 1992–1994
Editor-in-Chief, *Economic Review*, Institute of Economic Research, Hitotsubashi University, 1994–2008

Conference Organizations

Program Chair and Local Organization Chair, The First Far Eastern Meeting of the Econometric Society, Tokyo, October 1987.

Program Chair, Nara Conference on Normative Economics, Nara, October 1987.

Standing Committee, The Second Far Eastern Meeting of the Econometric Society, Kyoto, June 1989.

Program Chair, The First TCER Summer Conference on Economic Theory, Hakone, June 1989.

Program Committee, The Sixth World Congress of the Econometric Society, Barcelona, August 1990.

Program Committee, The Third Far Eastern Meeting of the Econometric Society, Seoul, June 1991.

Program Committee, The First World Meeting of the Society for Social Choice and Welfare, Caen, July 1992.

Program Committee, The First India and South East Asian Meeting of the Econometric Society, Bombay, December 1992.

Program Committee, The Fourth Far Eastern Meeting of the Econometric Society, Taipei, June 1993.

Program Co-Chair, The International Economic Association Round Table Meeting on Social Choice Theory, Vienna, May 1994.

Program Chair, The Fourth TCER Summer Conference on Economic Theory, Tateshina, July 1994.

Program Committee, The Second World Meeting of the Society for Social Choice and Welfare, Rochester, July 1994.

Program Committee, The Seventh World Congress of the Econometric Society, Tokyo, August 1995.

Program Committee, The Twelfth World Congress of the International Economic Association, Buenos Aires, August 1999.

Program Chair, The International Symposium on Social Choice and Welfare, Tokyo, July 2004.

Program Co-Chair, The International Economic Association Round Table Meeting on Intergenerational Equity and Sustainability, Hakone, March 2005.

Program Chair, The International Conference on Sustainable Well-Being: Science and Technology for Sustainability, Science Council of Japan, Tokyo, September 2008.

Contents

Introduction 1

Part I Arrovian Social Choice Theory and its Developments

Limited Rights as Partial Veto and Sen's Impossibility Theorem 11
Maurice Salles

Harmless Homotopic Dictators 25
Nicholas Baigent

Remarks on Population Ethics 35
Tomoichi Shinotsuka

On Non-Welfarist Social Ordering Functions 43
Naoki Yoshihara

Part II Social Choice and Fair Allocations

Monotonicity and Solidarity Axioms in Economics and Game Theory . . . 71
Yves Sprumont

To Envy or to be Envied? Refinements of the Envy Test for the Compensation Problem 95
Marc Fleurbaey

Choice-Consistent Resolutions of the Efficiency-Equity Trade-Off 119
Koichi Tadenuma

Characterization of the Maximin Choice Function in a Simple Dynamic Economy 139
Koichi Suga and Daisuke Udagawa

Part III Rational Choice, Individual Welfare, and Games

Suzumura Consistency 159
Walter Bossert

On the Microtheoretic Foundations of Cagan's Demand for Money Function 181
Rajat Deb, Kaushal Kishore, and Tae Kun Seo

Hicksian Surplus Measures of Individual Welfare Change When There is Price and Income Uncertainty 195
Charles Blackorby, David Donaldson, and John A. Weymark

Beyond Normal Form Invariance: First Mover Advantage in Two-Stage Games with or without Predictable Cheap Talk 215
Peter J. Hammond

Part IV Social Welfare and the Measurement of Unemployment and Diversity

Unemployment and Vulnerability: A Class of Distribution Sensitive Measures, its Axiomatic Properties, and Applications 237
Kaushik Basu and Patrick Nolen

Ordinal Distance, Dominance, and the Measurement of Diversity 259
Prasanta K. Pattanaik and Yongsheng Xu

Contributors

Nicholas Baigent
Institute of Public Economics, Graz University, A-8010 Graz, Austria
nickbaigent@waitrose.com

Kaushik Basu
Department of Economics, Cornell University, Ithaca NY 14853, USA
kb40@cornell.edu

Charles Blackorby
Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
c.blackorby@warwick.ac.uk

Walter Bossert
Department of Economics, University of Montréal, Montreal, Canada
walter.bossert@umontreal.ca

Rajat Deb
Department of Economics, Southern Methodist University, Dallas, TX 75275, USA
deb@smu.edu

David Donaldson
Department of Economics, University of British Columbia, Vancouver, B.C, V6T 1Z1, Canada
dvdd@telus.net

Marc Fleurbaey
CERSES-University Paris 5, 45 rue des Sts Peres, 75270 Paris Cedex, France
marc.fleurbaey@univ-paris5.fr

Peter J. Hammond
Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
p.j.hammond@warwick.ac.uk

Kaushal Kishore
Department of Economics, Southern Methodist University, Dallas, TX 75275
USA

Patrick Nolen
Department of Economics, University of Essex, Colchester CO4 3SQ, Essex,
UK
pjnolen@essex.ac.uk

Prasanta K. Pattanaik
Department of Economics, University of California, Riverside, CA 92521,
USA
ppat@ucrac1.ucr.edu

Maurice Salles
CREM, Université de Caen, 14032 Caen cedex, France
maurice.salles@unicaen.fr

Tae Kun Seo
Department of Economics, Southern Methodist University, Dallas, TX 75275, USA

Tomoichi Shinotsuka
Graduate School of Humanities and Social Sciences, University of Tsukuba, Japan
shino@social.tsukuba.ac.jp

Yves Sprumont
Département de Sciences, Économiques and CIREQ, Université de Montréal,
Montréal, Canada
yves.sprumont@umontreal.ca

Koichi Suga
School of Political Science and Economics, Waseda University, Shinjuku-ku,
Tokyo 169-8050, Japan
ksuga@waseda.jp

Koichi Tadenuma
Department of Economics, Hitotsubashi University, Kunitachi, Tokyo 186-8601,
Japan
koichi.t@econ.hit-u.ac.jp

Daisuke Udagawa
Graduate School of Economics, Waseda University, Shinjuku-ku, Tokyo 169-8050,
Japan
udagawa_daisuke@suou.waseda.jp

John A. Weymark
Department of Economics, Vanderbilt University, VU Station B #35189, Nashville,
TN 37235-1819, USA
john.weymark@vanderbilt.edu

Yongsheng Xu
Department of Economics, Georgia State University, Atlanta, GA 30303, USA
yxu3@gsu.edu

Naoki Yoshihara
Institute of Economic Research, Hitotsubashi University, Kunitachi,
Tokyo 186-8603, Japan
yosihara@ier.hit-u.ac.jp

Introduction

Kotaro Suzumura received his doctoral degree in economics from Hitotsubashi University in 1980. He has taught at various institutions, including Hitotsubashi University (1971–1973, 1982–2008), Kyoto University (1973–1982), London School of Economics (1974–1976), Stanford University (1979–1980), University of Pennsylvania (1987), University of Essex (1990–1991), and University of British Columbia (1994). He has been a British Council Visiting Scholar at Cambridge University (1973–1974), a Visiting Fellow at Australian National University (1986) and at All Souls College, Oxford University (1988), a Fulbright Senior Research Fellow at Harvard University (1993), a Nissan Visiting Fellow at St. Antony's College, Oxford University (1996), and a Visiting Fellow Commoner at Trinity College, Cambridge University (2001).

Though Kotaro is well-known mainly as the author of many seminal contributions to social choice theory, welfare economics, and theoretical industrial organization, he started his academic career by working in the fields of economic growth theory, general equilibrium in linear multisector models, and international trade. His papers in these fields were published in *Economic Studies Quarterly* (*Japanese Economic Review*) and *Metroeconomica*.

After completing his doctoral course work, Kotaro earnestly began to work on the theory of rational choice, social choice, and welfare economics. Between 1976 and 1982, he published many influential papers in leading economic journals, including *Review of Economic Studies*, *Journal of Economic Theory*, *Economica*, and *International Economic Review*. Those papers constituted the basis of his first seminal research monograph, *Rational Choice, Collective Decisions, and Social Welfare* (Cambridge University Press, 1983), which has been referred to by numerous researchers since its publication. He also launched his research project on competition and social welfare. His first major research work on competition and welfare was published in *Review of Economic Studies*. This was followed by many influential articles on this subject in *Journal of Public Economics*, *American Economic Review*, *International Economic Review*, and *Economic Theory*. These articles were finally developed into his second important monograph, *Competition, Commitment, and Welfare* (Oxford University Press, 1995).

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

Kotaro's recent research has spanned several important areas, including consistent preferences and rationality of choice; welfare, rights, and social choice procedures; economics of well-being and freedom; non-welfaristic foundations of welfare economics; intergenerational equity and global warming. He has published articles on these subjects in major professional journals such as *American Economic Review*, *Journal of Economic Theory*, *Economic Journal*, *Journal of Mathematical Economics*, *Social Choice and Welfare*, and *Economica*. Jointly with Kenneth J. Arrow and Amartya K. Sen, he edited *Handbook of Social Choice and Welfare*, Volumes 1 and 2 (North-Holland, Armsterdam).

Kotaro has also been interested in and worked on issues relating to economic policies in Japan. In particular, he has taken keen interest in the competition and industrial policy in Japan, competition and regulation in telecommunications and Japan's Reform experience, and welfare policies in Japan. In these areas, he has published many articles and has edited several books, including Yongsheng, Prasanta, and Naoki; *Industrial Policy of Japan* (Academic Press, 1988), jointly with R. Komiya and M. Okuno; *The Economic Theory of Industrial Policy* (Academic Press, 1991), jointly with M. Itoh, K. Kiyono, and M. Okuno-Fujiwara; and *Development Strategy and Management of the Market Economy* (Oxford University Press, 1997), jointly with E. Malinvaud, J.-C. Milleron, M. Nabli, A.K. Sen, A. Sengupta, N. Stern, and J.E. Stiglitz.

Kotaro has made a vast contribution to the profession through his involvement in various academic associations and conferences and also as an editor of several journals. He served as the editor of the *Journal of the Japanese and International Economies* in 1992–1994, and as an editor of the *Japanese Economic Review* in 1995–1998. In 1990, he was elected a Fellow of the Econometric Society. He was the President of the Japanese Economic Association (1999–2000) as well as of the Society for Social Choice and Welfare (2000–2001). Since October 2006, he has served as a Vice-President of the Science Council of Japan.

He received the Nikkei Prize twice, first in 1984 for *Rational Choice, Collective Decisions, and Social Welfare* (Cambridge University Press, 1983) and second in 1988 for *Economic Analysis of Industrial Policy* (Academic Press, 1988). In 2004, he was awarded the Medal with Purple Ribbon for his contributions to theoretical economics. In 2006, he received the Japan Academy Award for his contributions to welfare economics and social choice theory.

It is not possible to represent in one volume all the varied interests that mark Kotaro Suzumura's work. We have therefore decided to restrict attention to the intersection of our research interests with his.

Part I Arrovian Social Choice Theory and its Developments

The first part of the volume consists of four papers on Arrovian social choice theory and its developments.

Maurice Salles' "Limited rights as partial veto and Sen's impossibility theorem," deals with individual rights and social choice, a subject to which Kotaro has made important contributions. Salles considers a weakening of Sen's well-known condition of *minimal liberalism*. Sen's original condition of minimal liberalism required the existence of at least two individuals, each of whom is *locally decisive*. In contrast, Salles requires the existence of at least two individuals each of whom is only *locally semi-decisive*. Salles shows that even this weaker condition does not provide an escape route from Sen's famous paradox of a Paretian liberal if we replace Sen's *social decision function* by a *social welfare function* or any other aggregation function that may lie "between" these two in terms of the "rationality" of social preferences.

Nick Baigent's paper on "Harmless homotopic dictators" considers the possibility of continuous Paretian social welfare functions. It was Chichilnisky (1982) who first showed that, for all continuous Paretian social welfare functions, there must be a *homotopic dictator*. Baigent presents a reappraisal of the intuitive content of Chichilnisky's theorem insofar as he shows that the existence of the homotopic dictator does not entail an undesirable concentration of decisive powers in the hands of such a person. Baigent demonstrates that, even if there is a homotopic dictator, one can construct social preferences arbitrarily close to the preferences of the other agents whenever their preferences are not "opposite" to that of the dictator.

One of Kotaro's recent research interests is the intergenerational social choice problems, in which the size and composition of populations may naturally differ between social states, which arise at different points of time. Derek Parfit (1976, 1982, 1984) was the first to address the issue of "population ethics," and to criticize classical utilitarianism for its *repugnant conclusion*. In a series of papers and a monograph, Blackorby, Bossert, and Donaldson have studied the possibility of generalized utilitarian principles that avoid this repugnant conclusion. However, Arrhenius (2003) recently proposed a more serious problem, called the *very repugnant conclusion.* In his paper, "Remarks on population ethics," Tomoichi Shinotsuka investigates what happens to Blackorby, Bossert, and Donaldson's results on generalized utilitarianism in population ethics if the axiom of *avoidance of the repugnant conclusion* is replaced by the axiom of *avoidance of the very repugnant conclusion*.

Naoki Yoshihara's paper, "On non-welfarist social ordering functions," discusses *extended social ordering functions* (ESOFs), each of which yields a social ordering over alternative combinations of a resource allocation and an allocation rule (visualized as a game form). Yoshihara shows the possibility of reasonable non-welfarist ESOFs, which meet the condition of individual autonomy, a non-welfarist principle of distributive justice, and the welfarist Pareto principle, using a weaker lexicographic application method. It may be recalled that Kotaro, together with Prasanta Pattanaik, initiated the study of the framework of extended social ordering functions in the context of individual rights and social welfare, and, together with Reiko Gotoh and Naoki Yoshihara, applied this framework to resource allocation problems. Yoshihara's paper is an attempt to treat appropriately the values of procedural fairness and non-consequentialism in the context of social choice of rules and social institutions.

Part II Social Choice and Fair Allocations

The second part of the volume consists of four papers on fair allocations. Among various principles or concepts of fairness, these papers basically focus on fairness as no-envy, monotonicity, solidarity, and the maximin criterion.

In his paper, "Monotonicity and solidarity axioms in economics and game theory," Yves Sprumont provides an excellent survey of contributions that use solidarity and monotonicity principles in fair division problems and transferable utility cooperative games. In each of the fair division problems and cooperative games, Sprumont introduces *resource monotonicity* and its variants, and surveys the works on the compatibility of these axioms with other ethical principles such as efficiency and its stronger variant, core principle, as well as no-envy and its weaker variant, equal split lower bound. Then, he introduces *population monotonicity* and its variants, and surveys the works on the compatibility of these axioms with other ethical principles mentioned above.

Fairness as no-envy is one of the prominent notions of fair allocations, but it is well-known that envy-free allocations are hard to achieve in the context of the "compensation problem." It was Suzumura (1981a, b, 1983) who first systematically studied the ranking of social states on the basis of fairness as no-envy. In his paper, "To envy or to be envied? Refinements of the envy test for the compensation problem," Marc Fleurbaey explores further the possibility of a systematic use of rankings based on the notion of no-envy. In addition to the rankings based on the *number of envy relations* proposed by Suzumura (1983), Fleurbaey introduces two criteria for rankings based on the idea of *undominated diversity* (van Parijs, 1990, 1995), as well as three criteria for rankings based on the notion of *envy intensity*. Then, he examines whether allocation rules, which are respectively derived from the rankings based on the above mentioned criteria, satisfy some variants of the basic principles of *responsibility* and *compensation* for the two types of compensation problems.

Koichi Tadenuma's "Choice-consistent resolution of the efficiency-equity trade-off" deals with the social choice of equitable and efficient allocations. Tadenuma adopts a choice-theoretic approach to the issue of the efficiency-equity trade-off, and formulates the two contrasting principles, the *equity-first and efficiency-second principle* and the *efficiency-first and equity-second principle*, in the form of axioms on social choice correspondences. As equity notions, he considers both equity as *no-envy* and equity as *egalitarian-equivalence*. Then, he examines whether various social choice correspondences derived from either the equity-first principle or the efficiency-first principle can satisfy certain consistency properties of choice, such as *path independence* and *contraction consistency*. Tadenuma also discusses the relationships of his paper to Kotaro's seminal work (Suzumura, 1981a,b).

In contrast to the above three papers, which are devoted to the study of *intra*generational equity, Koichi Suga and Daisuke Udagawa contribute to the issue of intergenerational equity. They consider problems of social choice over infinite consumption paths in a simple dynamic economy à la Arrow (1973) and Dasgupta (1974a,b). They focus on the *Rawlsian social choice function* in this context, and provide a characterization of it. In an earlier contribution to the choice-theoretic approach to intergenerational equity, Asheim, Bossert, Sprumont, and Suzumura (2006) provided characterizations of all infinite-horizon choice functions by means of efficiency and time-consistency. Since the Rawlsian choice function does not satisfy time-consistency, the analysis of Suga and Udagawa may be seen as independent of, but complementary to, Asheim, Bossert, Sprumont, and Suzumura (2006).

Part III Rational Choice, Individual Welfare, and Games

The third part of the volume consists of four papers on the rationality of individual choice in single-person and/or multi-person decision problems and the welfare of individuals.

Rational choice theory constitutes the foundation for economic theory in general, and is applied to problems of individual choices as well as problems of social choice and welfare economics. In the field of rational choice theory, Kotaro Suzumura has made important contributions by providing and analyzing the notion of *Suzumura Consistency* (S-Consistency). S-Consistency is an axiom imposed on binary relations, and it is a necessary and sufficient condition for the existence of an *ordering extension* of a binary relation. Although this notion was first introduced by Kotaro more than 30 years ago, it still provides us with a variety of interesting research agendas. Walter Bossert's paper, "Suzumura consistency," provides a survey of recent works on S-Consistency. Bossert reviews how this notion can be used in a variety of applications, and provides some new observations to emphasize the importance of this axiom.

In market economies, money is doubtlessly irreplaceable by any other commodities. It was Cagan (1956) who introduced the "demand for money function," to explain the demand for money in inflationary environments. An extensive literature, both theoretical and empirical, has used his functional form in analyzing hyperinflation and the associated problem of "inflation tax." However, the use of Cagan's demand for money function has been "ad-hoc" and no attempt has been made to rationalize it in terms of "utility maximizing" behavior. Rajat Deb, Kaushal Kishore, and Tae Kun Seo's "On the microtheoretic foundations of Cagan's demand for money function," studies this unexplored but conceptually important issue of *rationalizability*. Deb, Kishore, and Seo assume that individuals are rational, that money is both a medium of exchange and a store of value, and that the demand for money is a result of intertemporal consumption smoothing. Then, they ask the question whether Cagan's demand function for money can be generated from some underlying process of utility maximization.

As exact measures of individual welfare change, the *Hicksian compensating and equivalent variations* are well-established, but the validity of these measures depends on the absence of uncertainty. Under uncertainty, the validity of expected versions of the Hicksian compensating and equivalent variations is restricted. Some papers, such as Helms (1984, 1985), characterized the restrictions on preferences, for which the expected compensating variation is a valid measure of individual welfare change when only one price is uncertain. However, Helms' framework is quite restrictive, since it is often the case that the incomes of consumers, as well as one or more prices, are also uncertain. The paper, "Hicksian surplus measures of individual welfare change when there is price and income uncertainty," by Charles Blackorby, David Donaldson, and John A. Weymark extend Helms' framework by considering the case where the consumer's income and some or all of the prices are uncertain, and identifies the circumstances in which the Hicksian compensating variation is a valid measure of individual welfare change.

The three papers discussed above basically study the problems of single-person decision-making and individual welfare. The last paper in this section, "Beyond normal form invariance: first mover advantage in two-stage games with or without predictable cheap talk," by Peter Hammond, however, considers multiperson decision problems in the context of non-cooperative games. It will be recalled that von Neumann (1928) was a strong believer in *normal form invariance*, which implies that the reduction of an extensive form game to the corresponding normal form game involves no loss of generality, and which has been a key assumption of the standard paradigm in the theory of non-cooperative games. In contrast, it has been recognized in experimental economics that there is a *first mover advantage* in Battle of the Sexes and similar games, which seems to indicate that normal form invariance is invalid. Hammond explores this critical view against the position of von Neumann, and introduces a "sophisticated" refinement of Nash equilibrium, which is capable of explaining the first mover advantage. This refinement depends on the extensive form of the game, and so it violates normal form invariance.

Part IV Social Welfare and the Measurement of Unemployment and Diversity

The fourth part of the volume consists of two papers, one on the measurement of unemployment and the other on the measurement of diversity. While the papers do not deal with social choice and welfare directly, they deal with phenomena which have indirect links with the notion of social welfare.

Kaushik Basu and Patrick Nolen's "Unemployment and vulnerability: A class of distribution sensitive measures, its axiomatic properties and applications," develops a way of measuring *effective* unemployment. Aggregate measures of unemployment have been recently criticized for ignoring *vulnerability* (the existence of people under the risk of *becoming* unemployed in the near future). In contrast, Basu and Nolen argue that the issue of vulnerability is not relevant for the underestimation of the pain of unemployment, but rather for the *inequality of the pain of unemployment*. They develop a class of *distribution-sensitive unemployment measures*, which take account of their normative stance regarding this issue. Basu and Nolen present a class of unemployment measures that satisfy several attractive axioms. They also provide a full characterization of the class of such measures, and apply them to data for the USA and for South Africa.

Prasanta K. Pattanaik and Yongsheng Xu's, "Ordinal distance, dominance, and the measurement of diversity," characterizes a class of rules for comparing sets of objects in terms of the *degrees of diversity*. Some previous works, such as Weitzman (1992, 1993, 1998) and Weikard (2002), developed cardinal measures of diversity. Pattanaik and Xu, however, use an ordinal notion of distance between objects and develop a notion of *dominance* between sets of objects. The class of rules for ranking sets of objects that they define and characterize is the class of rules that constitute extensions of this dominance relation.

Prasanta K. Pattanaik
Koichi Tadenuma
Yongsheng Xu
Naoki Yoshihara

Part I
Arrovian Social Choice Theory and its Developments

Limited Rights as Partial Veto and Sen's Impossibility Theorem

Maurice Salles

1 Introduction

The origin of the tremendous development of studies on rights and freedom within social choice theory and normative economics can be traced back to the famous short paper of Amartya Sen published in 1970 (Sen, 1970b); see also his book published the same year (Sen, 1970). In this paper, it is shown in the framework of aggregation procedures that there is a conflict between collective rationality (in terms of properties of choice functions or in terms of a transitivity-type of the social preference property – in fact, acyclicity of the asymmetric part of the social preference), Paretianism (a unanimity property) and some slight violation of neutrality (neutrality meaning that the names of options or social states are not to be taken into account) possibly combined with some slightly unequal distribution of power among individuals interpreted as an individual liberty property. Although, since then, rights have been considered within another paradigm, viz. game forms (see for instance Gärdenfors (1981, 2005), Gaertner, Pattanaik, and Suzumura (1992), Peleg (1998a,b) and Suzumura (2008)), and freedom has been mainly analyzed in the context of opportunity sets following the pioneering paper of Pattanaik and Xu (1990) (see also the survey by Barberà, Bossert, and Pattanaik (2004)), some authors (for instance Igersheim (2006) and Saari and Pétron (2006)) have recently revisited the foundational framework of Sen and Gibbard (1974) either by studying the informational structure of the aggregation procedure or by examining the consequences of taking a Cartesian structure to define the set of social states, consequences that take the form of a restriction of individual preferences. The purpose of this paper is different. I wish to formally study a weakening of the conditions associated with the notion of individual liberty. I have always considered that this condition was rather

M. Salles
CREM, Université de Caen, 14032 Caen cedex, France, and Institute for SCW, BP 1, 14111 Louvigny, France
e-mail: maurice.salles@unicaen.fr

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

strong in Sen's original paper. In fact, the condition is quite strong in the mathematical framework and only the interpretation, to my view, makes it not only acceptable but obvious. In his comments to a paper by Brunel (now Pétron) and Salles (1998), Hammond (1998) writes:

> In the social choice rule approach ..., local dictatorship becomes a desideratum, provided that the 'localities' are appropriate. Our feelings of revulsion should be reserved for non-local dictatorships, or local dictatorships affecting issues that should not be treated as personal.

I entirely share this opinion, but there is nothing in the basic mathematical framework that guarantees this personal aspect (in contrast with a suitable Cartesian product structure). In this basic framework, it is, however, possible to weaken local dictatorships. Unfortunately, I will show that this weakening does not offer a very interesting escape route from Sen's negative result. From a formal point of view, I believe that there is a sort of analogy that can be made between a family of Sen's impossibility theorems and Arrovian impossibility theorems.

After introducing general definitions and recalling Sen's theorems, I will present new Sen-type impossibility theorems, then will make a comparison with Arrovian impossibility theorems, commenting on similarities and obvious differences.

2 Basic Definitions and Sen's Theorem

Let X be the set of social states. Nothing specific is assumed for this set. A binary relation, a *preference*, over X is a subset of $X \times X$. It will be denoted by $\succeq$. I will write $x \succeq y$ rather than $(x,y) \in \succeq$. All binary relations considered in this chapter are supposed to be *complete* (for all x and $y \in X$, $x \succeq y$ or $y \succeq x$) and, consequently, *reflexive* (for all $x \in X$, $x \succeq x$). The *asymmetric* part of $\succeq$, denoted $\succ$ is defined (since $\succeq$ is complete) by $x \succ y$ if $\neg y \succeq x$. The *symmetric* part of $\succeq$ is defined by $x \sim y$ if $x \succeq y$ and $y \succeq x$. Intuitively, $x \succeq y$ will mean 'x is at least as good as y', $x \succ y$ will mean 'x is preferred to y' and $x \sim y$ will mean 'there is an indifference between x and y'. A preference $\succeq$ is *transitive* if for all x, y and $z \in X$, $x \succeq y$ and $y \succeq z \Rightarrow x \succeq z$. The asymmetric part of $\succeq$, $\succ$, is transitive if for all x, y and $z \in X$, $x \succ y$ and $y \succ z \Rightarrow\ x \succ z$. The symmetric part of $\succeq$, $\sim$, is transitive if for all x, y and $z \in X$, $x \sim y$ and $y \sim z \Rightarrow x \sim z$. If $\succeq$ is transitive, $\succ$ and $\sim$ are transitive too. We will say that $\succeq$ is *quasi-transitive* if $\succ$ is transitive (then $\sim$ is not necessarily transitive), and that $\succ$ is *acyclic* if there is no finite subset of X, $\{x_1, \ldots, x_k\}$, for which $x_1 \succ x_2, x_2 \succ x_3, \ldots, x_{k-1} \succ x_k$ and $x_k \succ x_1$. A complete and transitive binary relation is a *complete preorder* (sometimes called 'weak ordering'). Let $\mathbb{B}$ denote the set of complete binary relations over X, $\mathbb{P}$ denote the set of complete preorders over X, $\mathbb{Q}$ denote the set of complete and quasi-transitive binary relations over X, and $\mathbb{A}$ denote the set of complete binary relations over X whose asymmetric part is acyclic.

Let N be the set of individuals. Nothing specific will be assumed for this set unless it is clearly indicated that it is finite. Individual $i \in N$ has her preference given by a complete preorder $\succeq_i$ over X. A *profile* π is a function from N to $\mathbb{P}'$, $\pi : i \mapsto \succeq_i$, where $\mathbb{P}' \subseteq \mathbb{P}$ with $\mathbb{P}' \neq \emptyset$. Let Π' be the set of profiles when the $\succeq_i$'s are in $\mathbb{P}'$ and Π be the set of all profiles (when the $\succeq_i$'s are in $\mathbb{P}$). When N is finite and $\#N = n$, a profile is a n-list $(\succeq_1, \ldots, \succeq_n)$ with each $\succeq_i$ in $\mathbb{P}'$. Then $\Pi' = \mathbb{P}'^n$ and $\Pi = \mathbb{P}^n$ ($\mathbb{P}'^n$ and $\mathbb{P}^n$ are n-times Cartesian products of $\mathbb{P}'$ and $\mathbb{P}$).

Definition 1. An *aggregation function* is a function $f : \Pi' \to \mathbb{B}$.

An aggregation function associates a unique complete binary relation, a social preference, denoted by $\succeq_S$, to individual preferences (one preference for each individual).

Given an aggregation function f, and two (distinct) social states x and $y \in X$,[1] we will say that individual $i \in N$ is (x,y)-decisive if for all $\pi \in \Pi'$, $x \succ_i y \Rightarrow x \succ_S y$, where $\succ_S$ is the asymmetric part of $\succeq_S = f(\pi)$.

Definition 2. An individual who is (x,y)-decisive and (y,x)-decisive will be said to be $\{x,y\}$*-decisive* or a $\{x,y\}$*-dictator.*

I can now define Sen's two liberalism conditions. Let f be an aggregation function.

Definition 3. (Liberalism, general 2-D$^+$) For all $i \in N$, *there exist* a_i *and* $b_i \in X$ such that i is a $\{a_i, b_i\}$-dictator.

It should be noticed that $\mathbb{P}'$ must be large enough to have a non-trivial satisfaction of general 2-D$^+$: for each individual i it must be possible to have both $a_i \succ_i b_i$ and $b_i \succ_i a_i$. Also, it should be outlined that the condition is rather fair since each individual is endowed with the same kind of power. The theorem can be proved by using a weaker form of the foregoing condition.

Definition 4. (Minimal liberalism, minimal 2-D$^+$) There exist *two individuals i and* $j \in N$, and *a, b, c, d* $\in X$ such that i is a $\{a,b\}$-dictator and j is a $\{c,d\}$-dictator.

Of course, the fairness property disappeared. The options are to be 'interpreted' as being specific to the concerned individual, i.e., a and b are specific to individual i; a and b can even be 'interpreted' as perfectly identical social states except for some features that are personal to individual i. Clearly general 2-D$^+$ implies minimal 2-D$^+$.

As mentioned earlier, the domain of the aggregation function f must be rich enough. This will be taken care of (with some excess) by the following condition U.

Definition 5. (Universality, U) Let f be an aggregation function. *Universality* requires that $\mathbb{P}' = \mathbb{P}$.

This means that an individual preference can be any complete preorder. There is no restriction imposed by some kind of upper rationality or the existence of inter-individual constraints. The last condition (condition P) is a weak form of unanimity (Pareto principle).

[1] They have to be distinct so that saying it is superfluous since we consider that $x \succ_i y$ and $\succ_i$ is asymmetric.

Definition 6. (Pareto principle, P) Let f be an aggregation function, $\pi \in \Pi'$ and $x, y \in X$.[2] If for all $i \in N$, $x \succ_i y$, then $x \succ_S y$ where $\succ_S$ is the asymmetric part of $\succeq_S = f(\pi)$.

Sen's theorem is obtained within a large class of aggregation functions (Sen called them *social decision functions*).

Definition 7. A $\mathbb{A}$-*valued aggregation function* (or *social decision function*) is a function $f : \Pi' \to \mathbb{A}$.

The collective rationality imposed in this case is rather weak. It has an interesting consequence on the non-emptiness of the set of maximal elements in any finite subset of X (or since we are considering complete binary relations on the non-emptiness of maximum elements or choices).

Theorem 1. *If there are at least two individuals and if $\#X \geq 2$, there is no $\mathbb{A}$-valued aggregation function satisfying minimal 2-D^+, U and P.*

An immediate corollary is:

Corollary 1. *If there are at least two individuals and if $\#X \geq 2$, there is no $\mathbb{A}$-valued aggregation function satisfying general 2-D^+, U and P.*

3 Partial Veto and Sen-Type Theorems

It is in reading Pattanaik's paper (Pattanaik (1996)) that I got the impetus to work on this topic. In particular in this paper, Pattanaik discusses Sen's possible views regarding a distinction between a conception of rights as the ability to prevent something and a conception of rights as the obligation to prevent something which seems to be endowed in the liberalism conditions. Although I wished to devote some time to introduce modal theoretic techniques to deal with this distinction, I will be in this chapter more modest and will consider a weakening of the liberalism conditions. It is however obvious that this weakening is not a real response to the ability-obligation problem. Nevertheless, at least from a semantical point of view, having a (partial) veto corresponds rather well to the idea of an ability to prevent something. I will then introduce the notion of partial veto and will show how robust Sen's theorem is.

Given an aggregation function f, and two (distinct) social states x and $y \in X$, we will say that individual $i \in N$ is (x,y)-semi-decisive if for all $\pi \in \Pi'$, $x \succ_i y \Rightarrow x \succeq_S y$, where $\succeq_S = f(\pi)$.

Definition 8. An individual who is (x,y)-semi-decisive and (y,x)-semi-decisive will be said to be $\{x,y\}$-*semi-decisive* or a $\{x,y\}$-*vetoer*.

As can be seen, the difference between a $\{x,y\}$-vetoer and a $\{x,y\}$-dictator is the difference between $x \succeq_S y$ and $x \succ_S y$. A $\{x,y\}$-vetoer's power amounts to the

[2] Again, x and y are necessarily distinct.

assurance that y will not be 'ranked' before x in the social preference.[3] I can now define weak versions of liberalism.

Definition 9. (Weak liberalism, general 2-V$^+$) For all $i \in N$, there exist a_i and $b_i \in X$ such that i is a $\{a_i, b_i\}$-vetoer.

Definition 10. (Minimal weak liberalism, minimal 2-V$^+$). There exist two individuals i and $j \in N$, and a, b, c, $d \in X$ such that i is a $\{a,b\}$-vetoer and j is a $\{c,d\}$-vetoer.[4]

One of the possible Pareto extension functions (based on the weak form of the Pareto principle of condition P) that will be discussed later, indicates that having weak liberalism will not make $\succeq_S$ non-quasi-transitive. Consequently, problems can be met only for the transitivity of social preference, or, as will be seen, for binary relations which are, in some sense, between preorders and quasi-transitive binary relations. We will define two of these 'intermediate' relations, interval orders and semiorders. These definitions can be stated as properties of the asymmetric part of the complete binary relation $\succeq$ over X.

Definition 11. A binary relation $\succeq$ on X is an *interval order* if for all w, x, y, and $z \in X$, $w \succ y$ and $x \succ z \Rightarrow w \succ z$ or $x \succ y$.

The set of interval orders over X will be denoted by $\mathbb{I}$.

Definition 12. *A binary relation $\succeq$ on X is a semiorder* if it is an interval order and if for all w, x, y, and $z \in X$, $w \succ x$ and $x \succ y \Rightarrow w \succ z$ or $z \succ y$.

The set of semiorders will be denoted by $\mathbb{S}$. These two concepts have mainly been introduced in measurement theory to deal with possible intransitive indifference. Although indifference is not necessarily transitive contrary to what is the case with preorders, it should be noted that, for both concepts, $\succ$ is transitive (see Fishburn (1985) and Suppes, Krantz, Luce, and Tversky (1989)).

My first result is for social welfare functions.

Definition 13. A $\mathbb{P}$-*valued aggregation function* (or *social welfare function*) is a function $f : \Pi' \to \mathbb{P}$.

Theorem 2. *If there are at least two individuals, there is no $\mathbb{P}$-valued aggregation function satisfying U, P and minimal 2-V$^+$, provided that, in the definition of minimal 2-V$^+$, $\{a,b\} \neq \{c,d\}$.*

---

[3] I use quotation marks for 'ranked' since a ranking can only be meaningful for X being finite with a (some) social state(s) ranked first, etc.

[4] Although I suspected that the notion of weak liberalism was not new, it is in reading (Sen, 1982) that I discovered that it had been first proposed in a different form by Karni (1974) in a working paper that is still unpublished. In Karni's paper the condition applies to subsets of alternatives. This working paper also includes important comments about the Cartesian product structure and unconditional preferences.

Proof. Let f be an aggregation function. i is a $\{a,b\}$-vetoer and j is a $\{c,d\}$-vetoer. $\{a,b\} \neq \{c,d\}$. Note that $\#X \geq 3$. Let us assume first that $\{a,b\} \cap \{c,d\} \neq \emptyset$. Without loss of generality, assume that $b = c$. Let π be a profile such that $a \succ_i b$, $b \succ_j d$, and for all $k \in N$, $d \succ_k a$. Then, we have for i, $d \succ_i a \succ_i b$ and for j, $b \succ_j d \succ_j a$. Since i is a $\{a,b\}$-vetoer, we have $a \succeq_S b$, and since j is a $\{b,d\}$-vetoer, we have $b \succeq_S d$. If f were $\mathbb{P}$-valued, by transitivity, we should have $a \succeq_S d$, but by condition P, we have $d \succ_S a$, a contradiction.

Consider now the case where $\{a,b\} \cap \{c,d\} = \emptyset$. Let π be a profile such that $a \succ_i b$, $c \succ_j d$ and for all $k \in N$, $b \succ_k c$ and $d \succ_k a$. Then, we have $d \succ_i a \succ_i b \succ_i c$ and $b \succ_j c \succ_j d \succ_j a$. By condition P, we have $b \succ_S c$ and $d \succ_S a$. Since j is a $\{c,d\}$-vetoer, we have $c \succeq_S d$. If f were $\mathbb{P}$-valued, $b \succ_S c$ and $c \succeq_S d$ and $d \succ_S a$ would imply $b \succ_S a$. But $a \succeq_S b$ since i is a $\{a,b\}$-vetoer, a contradiction. □

Although one can obtain a cycle with two options ($a \succ_S b$ and $b \succ_S a$), three options are necessary for an intransitivity. Now, I will consider the case of interval orders.

Theorem 3. *If there are at least two individuals, there is no $\mathbb{I}$-valued aggregation function satisfying U, P and minimal 2-V$^+$, provided that, in the definition of minimal 2-V$^+$, $\{a,b\} \cap \{c,d\} = \emptyset$.*

Proof. Let f be an aggregation function. Obviously, $\#X \geq 4$. Let π be a profile such that $a \succ_i b$, $c \succ_j d$, and for all $k \in N$, $b \succ_k c$ and $d \succ_k a$. Observe that $d \succ_i a \succ_i b \succ_i c$ and $b \succ_j c \succ_j d \succ_j a$. Since $b \succ_S c$ and $d \succ_S a$ by condition P, we should have if $\succeq_S$ were an interval order, i.e., if f were $\mathbb{I}$-valued, $b \succ_S a$ or $d \succ_S c$. But we have ($a \succeq_S b$ and $c \succeq_S d$) since i is a $\{a,b\}$-vetoer and j is a $\{c,d\}$-vetoer, a conjunction that is the negation of the disjunction ($b \succ_S a$ or $d \succ_S c$) (given completeness of $\succeq_S$). □

If $\#X = 3$, the condition given in Definition 11 is reduced to the transitivity of $\succ_S$. Then, the Pareto extension function based on the weak form of the Pareto principle given in condition P is a counter-example. Let me define this Pareto extension function.

Definition 14. Let $\pi \in \Pi'$ and $x, y \in X$. f is the *weak Pareto extension function* if $x \succ_S y \Leftrightarrow \forall k \in N\ x \succ_k y$, and $y \succeq_S x$ otherwise.

One can easily see that $\succ_S$ is transitive and that each individual i is a $\{x,y\}$-vetoer for all $\{x,y\} \subseteq X$.

Semiorders are ‘between’ preorders and interval orders. Can we expect to have some progress? In fact, one obtains, as could be expected, a theorem ‘between’ Theorem 2 and 3, although the refinement is quite modest.

Theorem 4. *If there are at least two individuals, there is no $\mathbb{S}$-valued aggregation function satisfying U, P and minimal 2-V$^+$, provided that, in the definition of minimal 2-V$^+$, $\{a,b\} \neq \{c,d\}$ and provided that $\#X \geq 4$.*

Proof. Let f be an aggregation function. Suppose first that $\{a,b\} \neq \{c,d\}$, but $\{a,b\} \cap \{c,d\} \neq \emptyset$. Without loss of generality, assume that $a = d$. Consider a profile

π such that $a \succ_i b$, $c \succ_j a$, and for all $k \in N$, $b \succ_k e$ and $e \succ_k c$, where e is a fourth social state. Note that $a \succ_i b \succ_i e \succ_i c$ and $b \succ_j e \succ_j c \succ_j a$. Then, by condition P, $b \succ_S e$ and $e \succ_S c$. If $\succeq_S$ were a semiorder, i.e., if f were $\mathbb{S}$-valued, we should have ($b \succ_S a$ or $a \succ_S c$). But we have ($a \succeq_S b$ and $c \succeq_S a$), since i is a $\{a,b\}$-vetoer and j is a $\{a,c\}$-vetoer, a conjunction that is the negation of the disjunction ($b \succ_S a$ or $a \succ_S c$), a contradiction.

If $\{a,b\} \cap \{c,d\} = \emptyset$, the proof is of course similar to the proof of Theorem 3. □

Theorems 2–4 have obvious corollaries (omitted) when minimal weak liberalism is replaced by weak liberalism.

This hierarchy of results is reminiscent of the family of Arrovian impossibility theorems. In Sect. 4, I will present a parallel between these two families of impossibility results.

4 Comparing Sen-Type Impossibilities with Arrovian Impossibilities

I will very briefly state Arrovian theorems with their necessary supplementary definitions. For all these theorems N is supposed to be finite with $\#N = n$. In all the definitions of this section, we suppose that f is an aggregation function.

Definition 15. (Independence-binary form, I) Let π and $\pi' \in \Pi'$ with $\pi : i \mapsto \succeq_i$ and $\pi' : i \mapsto \succeq'_i$. Consider any x, $y \in X$. If $\succeq_i |\{x,y\} = \succeq'_i |\{x,y\}$ for all $i \in N$, then $\succeq_S |\{x,y\} = \succeq'_S |\{x,y\}$ where $\succeq_S = f(\pi)$ and $\succeq'_S = f(\pi')$. ($\succeq_S |\{x,y\}$ is the restriction of $\succeq_S$ to $\{x,y\}$.)

Definition 16. A *dictator* is an individual who is a $\{x,y\}$-dictator for all $\{x,y\} \subseteq X$.

Definition 17. (Condition D$^-$, non-dictatorship) There is no dictator.

Theorem 5. (Arrow, 1950, 1951, 1963) If $n \geq 2$ and $\#X \geq 3$, there is no $\mathbb{P}$-valued aggregation function (social welfare function) satisfying U, P, I and D$^-$.

The Pareto extension function is a counter-example to a theorem which would be similar to Arrow's theorem except that $\mathbb{P}$-valuedness would be replaced by $\mathbb{Q}$-valuedness. However, if non-dictatorship is replaced by a no-vetoer condition, the result is restaured.

Definition 18. A *vetoer* is an individual who is a $\{x,y\}$-vetoer for all $\{x,y\} \subseteq X$.

Definition 19. (Condition V$^-$, no-vetoer) There is no vetoer.

Theorem 6. (Gibbard, 1969) *If $n \geq 2$ and $\#X \geq 3$, there is no $\mathbb{Q}$-valued aggregation function satisfying U, P, I and V$^-$.*

There is more in the original Gibbard's paper, since Gibbard shows that if f is a $\mathbb{Q}$-valued aggregation function satisfying U, P, I, there exists an oligarchy, a group of individuals having full power if they act unanimously and whose members are all vetoers. For $n = 2$, majority rule gives a quasi-transitive social preference, but, in this case, each of the two individuals is a vetoer.

If one considers $\mathbb{A}$-valued aggregation function (social decision function), one can still get an impossibility provided that the aggregation function is increasing (this property is often called strict monotonicity or positive responsiveness).

Definition 20. (Increasing aggregation function, IF) An aggregation function f is an increasing aggregation function if for all π, $\pi' \in \Pi'$, and all x, $y \in X$, if for all $i \in N$, ($x \succ_i y \Rightarrow x \succ'_i y$ and $x \sim_i y \Rightarrow x \succeq'_i y$), and there exists $j \in N$ such that ($y \succ_j x$ and $x \succeq'_j y$) or ($x \sim_j y$ and $x \succ'_j y$), then $x \succeq_S y \Rightarrow x \succ'_S y$, where $\succeq_S = f(\pi)$ and $\succ'_S$ is the asymmetric part of $\succeq'_S = f(\pi')$.

Intuitively this condition means that if option x does not decrease vis-à-vis option y in all individual preferences, and if x increases vis-à-vis y in at least one individual preference, then, this increase must be reflected at the social level, when possible.

Theorem 7. (Mas-Colell and Sonnenschein, 1972) *If $n \geq 4$ and #$X \geq 3$, there is no $\mathbb{A}$-valued increasing aggregation function (increasing social decision function) satisfying U, P, I,and V^-.*

A rather confidential result offers a refinement of this theorem.

Definition 21. A *quasi-dictator* is an individual i who is a vetoer such that for all $\pi \in \Pi'$, and all x, $y \in X$, $x \succ_i y$ and $x \sim_S y \Rightarrow$ for all $j \neq i$, $y \succ_j x$, where $\sim_S$ is the symmetric part of $\succeq_S = f(\pi)$.

A quasi-dictator is then nearly exactly similar to the Arrovian dictator except in the case where all other individuals have a strict preference that is the inverse of his strict preference.[5]

Definition 22. (Condition Q-D$^-$, non-quasi-dictatorship) There is no quasi-dictator.

Theorem 8. (Bordes & Salles, 1978) *If $n \geq 4$ and #$X \geq 3$, there is no $\mathbb{A}$-valued increasing aggregation function (increasing social decision function) satisfying U, P, I, and Q-D$^-$.*

Surprisingly, Arrovian theorems regarding semiorder-valued (or interval order-valued) aggregation functions appeared later. It was, however, quite important to know that Arrow's theorem could be obtained without postulating that the social preference be a complete preorder. These important and somewhat neglected results are due to Blair and Pollack (1979) and Blau (1979) (in fact their papers appeared in the same issue of the *Journal of Economic Theory*).

---

[5] I prefer to use 'his' rather than 'her' for this sort of people! I hope that this is not politically incorrect.

Table 1 Comparing Arrovian and Sen-type impossibility theorems

	Impossibility theorems	
Aggregation function	Arrovian (with U+I+P)	Sen-Type (with U+P)
$\mathbb{P}$-valued	D$^-$	Minimal 2-V$^+$
$\mathbb{S}$-valued	D$^-$	Minimal 2-V$^+$
$\mathbb{I}$-valued	D$^-$	Minimal 2-V$^+$
$\mathbb{Q}$-valued	V$^-$	Minimal 2-D$^+$
$\mathbb{A}$-valued	(V$^-$ or Q-D$^-$)+IF	minimal 2-D$^+$

Theorem 9. (Blair and Pollack, 1979; Blau, 1979) *If $n \geq 4$ and $\#X \geq 4$, there is no $\mathbb{S}$-valued aggregation function or no $\mathbb{I}$-valued aggregation function satisfying U, P, I and D$^-$.*

To the best of my knowledge and contrary to the case of Sen-type impossibility theorems, there is no way to make a distinction between $\mathbb{S}$-valued and $\mathbb{I}$-valued aggregation function. Of course, for three options, the properties of interval orders and of semiorders are both reduced to quasi-transitivity and then the Pareto extension function gives an appropriate counter-example.

Table 1 provides a useful summary of the preceding results and establishes a sort of parallel.

What this table shows is the connection between D$^-$ and 2-V$^+$, and between V$^-$ and 2-D$^+$. However, as shown in the next section, this parallel should not be taken too seriously.

5 Discussion

In this section, I will outline the major differences between the two categories of impossibility theorems. The first one concerns the set of individuals N. As I mentioned at the beginning of the preceding section, for Arrovian impossibility theorems we assume that N is finite. From a historical point of view, after the publication of Arrow's papers and book, a question was whether this assumption was only there to make the proof of the theorem easier (and, after all, N finite is a rather easily justifiable property) or whether this was a necessity. Fishburn (1970) was the first to show that it was a necessary assumption. With N infinite, Fishburn provided a counter-example. Incidentally, Fishburn's short paper was the starting point of an active research with possibly metaphysical implications.

As clearly stated in the Table 1, independence (of irrelevant alternatives–different from the Chernoff variety) is a very important feature of Arrovian theorems. This property is not used at all in Sen-type theorems. The interplay between the various conditions of social rationality, the different Pareto principles and the properties of decisiveness and semi-decisiveness in the light of Sen's Paretian epidemic deserves to be scrutinized in a future work.[6]

Both categories of impossibility theorems are stated with condition U. It is, however, possible to define smaller domains so that we could obtain impossibilities (see Kalai and Muller (1977) for Arrovian social welfare functions). For Sen-type theorems, one only need a domain rich enough to include the profiles leading to the impossibilities.

Sen's theorem is often considered with social choice functions rather than $\mathbb{A}$-valued functions (social decision functions). A *social choice function* is a function $f^{\star} : 2^X - \emptyset \times \Pi' \to 2^X - \emptyset$ such that for all $S \in 2^X - \emptyset$ and all $\pi \in \Pi'$, $f^{\star}(S,\pi) \subseteq S$. This function selects social states in each non-empty subset of the set of social states. To define *liberalism*, one can say that for all $i \in N$, there exist two social states $\{a_i, b_i\}$ such that for all non-empty $S \subseteq X$ and all $\pi \in \Pi'$, if $a_i \in S$ and $a_i \succ_i b_i$, then $b_i \notin f^{\star}(S,\pi)$ and if $b_i \in S$ and $b_i \succ_i a_i$, then $a_i \notin f^{\star}(S,\pi)$. A similar definition can be given for *minimal liberalism* by restricting the definition to only two individuals in a way similar to what was done previously. Using this framework, the proof of the corresponding theorem consists in emptying $f^{\star}(S,\pi)$ for specific S and π. This proof is as easy as the proof for $\mathbb{A}$-valued functions, and it is not surprising given the strong relations between $\succ$-cycles and the absence of maximal elements (or between acyclicity and the existence of maximal elements, as previously mentioned). Things are less simple with weak liberalism. With liberalism, if $a_i \succ_i b_i$, then b_i is rejected from all choice sets $f^{\star}(S,\pi)$ such that $a_i \in S$. This means that $f^{\star}(\{a_i, b_i\},\pi) = \{a_i\}$. This corresponds intuitively well to $a_i \succ_S b_i$. Weak liberalism only tells us that $a_i \succeq_S b_i$. Intuitively but also in the standard choice literature, this corresponds to $a_i \in f^{\star}(\{a_i, b_i\},\pi)$. This means that a_i must be selected but it does not say that b_i is rejected, and, furthermore, we cannot say anything about the selection from the other sets to which a_i belongs. Of course this difficulty can be probably taken care of by imposing to $f^{\star}$ properties borrowed from the revealed preference and rationalizability literature (this will be the subject of another paper).

Finally and this is the main difference, difference which is at the origin of recent major developments on non-welfaristic issues in normative economics, Sen-type theorems are non-welfaristic.[7] The word *welfarism* is associated with the idea that the goodness of social states are evaluated only on the basis of individual utilities attached to these social states. This leads to the following observation. If we have four social states w, x, y and z and if each individual i attributes the same utility to w and to x, and the same utility to y and to z, then, the social ranking of w and y must be the same as the social ranking of x and z. This can lead to various properties of neutrality for functions defined on profiles of utility functions and this can be extended to profiles of individual preferences in which case one obtains intra or inter profiles neutrality (for an introduction to the non-welfaristic literature, I rec-

[6] Paretian epidemic was first described and analyzed in Sen (1976, 1982). This remark was prompted by a comment from the referee, a comment that is very gratefully acknowledged.

[7] The following remarks owe much to Kotaro Suzumura. I am very grateful to him for calling my attention to this crucial aspect.

ommend the remarkable article by Pattanaik published in a too confidential book, see Pattanaik (1994)). Intuitively, neutrality means that names of social states do not matter. The liberalism conditions obviously violate neutrality since specific social states are attached to specific individuals.

6 Conclusion

In this chapter, Sen's liberalism conditions have been weakened. Partial dictatorship has been replaced by partial veto. This weakening could be justified to some extent by a wish to consider rights as the ability to prevent something to happen rather than the obligation to prevent something to happen. Unfortunately, this weakening does not take us very far since impossibilities will occur if we replace *social decision functions* by *social welfare functions* or other aggregation functions 'between' social welfare functions and social decision functions. Considering a kind of hierarchy of aggregation functions on the basis of the collective rationality of the associated social preference, it was natural to compare Sen-type impossibility theorems with Arrovian impossibility theorems. Although this comparison shows that there is some interesting relations from a mathematical point of view, the discrepancies are probably more important from an interpretative point of view. In particular, the discrepancy between welfaristic aspects of the Arrovian theorems and the non-welfaristic aspects of Sen-type theorems is a very important one that has been outlined. Sen's theorem has been justly considered as the foundational result of non-welfaristic normative economics. In, forthcoming papers, I will consider the extension of the results of the present paper to a choice-theoretic framework. I will also come back to the debate between ability (possibility) and obligation by using modal logic.

Appendix 1

Blau (1979) and to some extent Blair and Pollack (1979) (and probably others) define interval orders and semiorders differently from Definitions 11 and 12. Given that $\succeq_S$ is complete, the following propositions show the equivalence between the definitions of the present chapter which are borrowed from Fishburn (1985) and Suppes, Krantz, Luce, and Tversky (1989) and the definitions used by Blau (these results are probably already known, but I have been unable to find where it could be).

Proposition 1. *Let $\succeq$ be complete. Then the following two statements are equivalent.*

(i) *For all w, x, y, and $z \in X$, $w \succ y$ and $x \succ z \Rightarrow w \succ z$ or $x \succ y$.*
(ii) *For all w, x, y, and $z \in X$, $w \succ x$ and $x \sim y$ and $y \succ z \Rightarrow w \succ z$.*

Proof. (i) $\Rightarrow$ (ii). Suppose (i) and $w \succ x$ and $x \sim y$ and $y \succ z$. Since $w \succ x$ and $y \succ z$, then by (i), $w \succ z$ or $y \succ x$. But $\neg y \succ x$ since $x \sim y$, and then $w \succ z$.
(ii) $\Rightarrow$ (i). Suppose (ii) but not (i). Then there exist a, b, c and $d \in X$ such that $a \succ c$ and $b \succ d$ and $\neg(a \succ d$ or $b \succ c)$. But $\neg(a \succ d$ or $b \succ c)$ is equivalent to $(d \succeq a$ and $c \succeq b)$ since $\succeq$ is complete. $c \succeq b$ is either $c \succ b$ or $c \sim b$. If $c \succ b$, $a \succ c$ and $c \succ b$ and $b \succ d$ imply $a \succ d$ since $\succ$ is transitive by (ii) (take $x = y$ and note that $\sim$ is reflexive). This contradicts $d \succeq a$. If $c \sim b$, $a \succ c$ and $c \sim b$ and $b \succ d$ imply $a \succ d$ by (ii), contradicting $d \succeq a$ again. $\square$

The next result concerns Definition 12.

Proposition 2. *Let $\succeq$ be complete. Then the following two statements are equivalent.*

(i) For all w, x, y, and $z \in X$, $w \succ x$ and $x \succ y \Rightarrow w \succ z$ or $z \succ y$.
(ii) For all w, x, y, and $z \in X$, $w \succ x$ and $x \succ y$ and $y \sim z \Rightarrow w \succ z$.

Proof. (i) $\Rightarrow$ (ii). Suppose we have (i) and that $w \succ x$ and $x \succ y$ and $y \sim z$. But $w \succ x$ and $x \succ y$ imply by (i) ($w \succ z$ or $z \succ y$). Since $y \sim z$, then $\neg z \succ y$, and $w \succ z$.
(ii) $\Rightarrow$ (i). Suppose we have (ii) and not (i). Then there exist a, b, c and $d \in X$ such that $a \succ b$ and $b \succ c$ and $\neg(a \succ d$ or $d \succ c)$, i.e., ($d \succeq a$ and $c \succeq d$) by completeness of $\succeq$. If $c \succ d$, $a \succ b$ and $b \succ c$ and $c \succ d$ imply $a \succ d$ since $\succ$ is transitive by (ii) (take $y = z$ and note that $\sim$ is reflexive). But this contradicts $d \succeq a$. If $c \sim d$, $a \succ b$ and $b \succ c$ and $c \sim d$ imply $a \succ d$ by (ii), which contradicts $d \succeq a$. $\square$

Acknowledgments This paper is dedicated to Kotaro Suzumura. It is a variation on and a tribute to Amartya Sen's 1970 landmark paper. I am grateful to Feng Zhang for the discussions on this topic and to Nick Baigent and other participants at the International Conference on Rational Choice, Individual Rights and Non-Welfaristic Normative Economics, Hitotsubashi University, Tokyo, 11–13 March 2006, for their remarks and comments. The conversation I had with Kotaro Suzumura after my talk there was crucial for the comparison section. The comments on previous drafts of the paper from an anonymous referee, Edi Karni and Prasanta Pattanaik were very helpful for the preparation of the final version. Financial support from the French ANR through contract NT05-1_42582 (3LB) is gratefully acknowledged.

References

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, *58*, 328–346
Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York (1st ed. 1951): Wiley
Barberà, S., Bossert, W., & Pattanaik, P. K. (2004). Ranking sets of objects. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of utility theory, volume 2* (pp. 893–977). Dordrecht: Kluwer
Blair, D. H., & Pollack, R. A. (1979). Collective rationality and dictatorship: The scope of the Arrow theorem. *Journal of Economic Theory*, *21*, 186–194
Blau, J. H. (1979). Semiorders and collective choice. *Journal of Economic Theory*, *21*, 195–206
Bordes, G., & Salles, M. (1978). Sur l'impossibilité des fonctions de décision collective: Un commentaire et un résultat. *Revue d'Economie Politique*, *88*, 442–448

Brunel, A., & Salles, M. (1998). Interpretative, semantic and formal difficulties of the social choice approach to rights. In J.-F. Laslier, M. Fleurbaey, & A. Trannoy (Eds.), *Freedom in economics: New perspectives in normative economics* (pp. 101–111). London: Routledge

Fishburn, P. C. (1970). Arrow's impossibility theorem: Concise proof and infinite voters. *Journal of Economic Theory*, *2*, 103–106

Fishburn, P. C. (1985). *Interval orders and interval graphs*. New York: Wiley

Gaertner, W., Pattanaik, P. K., & Suzumura, K. (1992). Individual rights revisited. *Economica*, *59*, 161–177

Gärdenfors, P. (1981). Rights, games and social choice. *Nous*, *15*, 341–356

Gärdenfors, P. (2005). *The dynamics of thought*. Heidelberg: Springer

Gibbard, A. (1969). *Social choice and Arrow's conditions* (mimeo). University of Michigan

Gibbard, A. (1974). A Pareto consistent libertarian claim. *Journal of Economic Theory*, *7*, 388–410

Hammond, P. J. (1998). Some comments on Brunel and Salles. In J.-F. Laslier, M. Fleurbaey, & A. Trannoy (Eds.), *Freedom in economics: New perspectives in normative economics* (pp. 112–115). London: Routledge

Igersheim, H. (2006). *Invoking the Cartesian product structure on social states: New resolutions of Sen's and Gibbard's impossibility theorems* (mimeo). Universitat Autonoma de Barcelona

Kalai, E., & Muller, E. (1977). Characterizations of domains admitting nondictatorial social welfare functions and nonmanipulable voting procedures. *Journal of Economic Theory*, *16*, 457–469

Karni, E. (1974). *Individual liberty, the Pareto principle and the possibility of social decision functions* (mimeo). The Foerder Institute for Economic Research, Tel-Aviv University

Mas-Colell, A., & Sonnenschein, H. (1972). General possibility theorems for group decision functions. *Review of Economic Studies*, *39*, 185–192

Pattanaik, P. K. (1994). Some non-welfaristic issues in welfare economics. In B. Dutta (Ed.), *Welfare economics* (pp. 197–248). Delhi: Oxford University Press

Pattanaik, P. K. (1996). On modelling individual rights: Some conceptual issues. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Social choice re-examined, volume 2* (pp. 100–128). London: Macmillan

Pattanaik, P. K., & Xu, Y. (1990). On ranking opportunity sets in terms of freedom of choice. *Recherches Economiques de Louvain*, *56*, 383–390

Peleg, B. (1998a). Effectivity functions, game forms, games, and rights. *Social Choice and Welfare*, *15*, 67–80

Peleg, B. (1998b). Effectivity functions, game forms, games, and rights. In J.-F. Laslier, M. Fleurbaey, & A. Trannoy (Eds.), *Freedom in economics: New perspectives in normative economics* (pp. 116–132). London: Routledge

Saari, D. G. , & Pétron, A. (2006). Negative externalities and Sen's liberalism theorem. *Economic Theory*, *28*, 265–281

Sen, A. K. (1970a). *Collective choice and social welfare*. San Francisco: Holden-Day

Sen, A. K. (1970b). The impossibility of a Paretian liberal. *Journal of Political Economy*, *78*, 152–157

Sen, A. K. (1976). Liberty, unanimity, and rights. *Economics*, *49*, 217–245

Sen, A. K. (1982). *Choice, welfare and measurement*. Oxford: Blackwell

Suppes, P., Krantz, D. H., Luce, R. D., & Tversky, A. (1989). *Foundations of measurement, volume 2*. New York: Academic Press

Suzumura, K. (2008). Rights, opportunities, and social choice procedures. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare, volume 2.* Amsterdam: Elsevier

Harmless Homotopic Dictators

Nicholas Baigent

1 Introduction

This paper constructs continuous Paretian social welfare functions for which one agent is a homotopic dictator but another is, in a precise sense, almost all powerful. The significance of this arises from the widely differing views[1] that have been expressed about a theorem in Chichilnisky (1982) showing that, for all continuous Paretian social welfare functions there must be a homotopic dictator. What the analysis in this paper therefore shows is that Chichilnisky's theorem is not a genuine Arrow-type impossibility theorem in the sense that desirable properties are not shown to entail some undesirable concentration of power.

While this does not necessarily mean that Chichilnisky's theorem is not significant, at least it calls for a reappraisal. One possible argument for the significance of this theorem starts from the fact that a homotopic dictator is also a strategic manipulator. However, as argued below, this argument does not establish the independent significance of Chichilnisky's theorem. At best, its significance seems to be derivative.

Section 2 introduces the main concepts and definitions. Section 3 provides an informal overview drawing heavily on diagrams. Section 4 presents results and a final Section 5 concludes with a summary and a few remarks towards a reappraisal of Chichilnisky's theorem.

N. Baigent
Institute of Public Economics, Graz University A-8010, Graz, Austria
e-mail: nickbaigent@waitrose.com

[1] See Chichilnisky (1982), Heal (1997), Sen (1986) on the one hand and Baigent (2003), Lauwers (2000), MacIntyre (1998), Saari (1997) on the other hand.

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,
Studies in Choice and Welfare.

2 Concepts and Definitions

Consider parallel linear indifference curves on a two dimensional space of alternatives, and call the underlying preferences *linear preferences*. Figure 1 shows two indifference curves for each of two linear preferences. For a given linear preference, draw a vector of length 1 perpendicular to an indifference curve at an arbitrary alternative. Such vectors are called unit normals. Since they are independent of the arbitrary alternative, a linear preference may be represented by such a *unit normal*. Also, since each unit normal takes a point in the Euclidean plane to another point on a circle of radius 1, the set of all preferences may be taken as the set of points on a unit circle. For convenience, re-centre this circle at the origin as in Fig. 2. Thus, the set of all linear preferences will be taken as: $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} = 1\}$.

For all vectors, $x = (x_1, x_2) \in S^1$, its polar coordinates are $(1, \rho_x)$ where ρ_x is the distance around the circle S^1 from the vector $(1,0)$ to x in the positive (anticlockwise) direction as shown by a bold arc in Fig. 2.

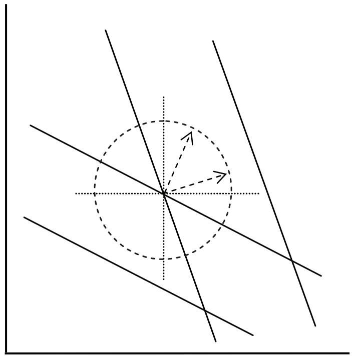


Fig. 1

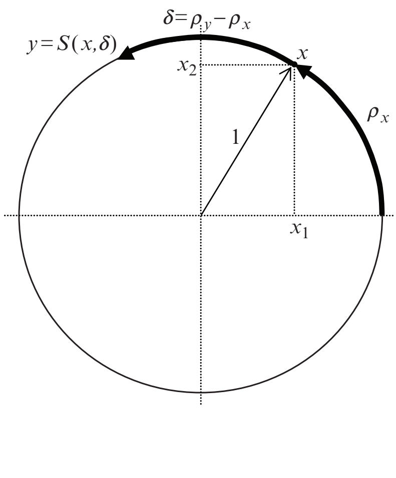


Fig. 2

Let $[0,2\pi]$ denote the closed interval of real numbers from 0 to 2π, and let $(0,2\pi)$ denote the open interval from 0 to 2π.[2] For all $x \in S^1$ and all $\delta \in [0,2\pi]$, let $s(x,\delta) \in S^1$ denote the point in S^1 that is a distance of δ around S^1 from x in an anticlockwise direction. Thus, for all $x,y \in S^1$, $s(x,\delta) = y$ if and only if $\rho_y - \rho_x = \delta$, see Fig. 2. That is, $s(x,\delta)$ determines an anticlockwise rotation from $x \in S^1$. Since the circumference of the unit circle is equal to 2π, it follows immediately that:

$$s(x,0) = s(x,2\pi) = x \tag{1}$$

For simplicity, consider the case of only two agents. A *social welfare function* is then a function $f : S^1 \times S^1 \to S^1$ that assigns a group linear preference $f(x,y) \in S^1$ to all pairs of individual linear preferences $(x,y) \in S^1 \times S^1$. Since the domain and range of a social welfare function are subsets of Euclidean spaces, when continuity is required it is taken in the usual sense for functions between subsets of Euclidean spaces.[3]

3 Overview

Continuous Paretian social welfare functions on a two-dimensional space of alternatives may be illustrated in a simple diagram. This diagram is used in this section to offer an informal presentation of the main point of the paper that is presented more precisely in Section 4.

The Weak Pareto property of social welfare functions requires that the group preference rank one alternative strictly above another whenever every individual does. In Fig. 3, an indifference curve for each agent is given in bold for which a is ranked above b. This is also the case for the indifference curve of the group preference, shown by the dotted line. Indeed, for the social preferences illustrated, any

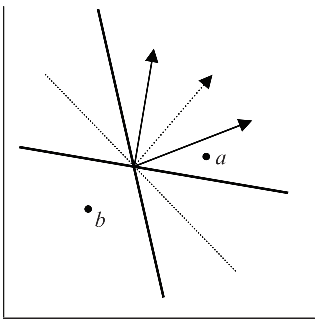
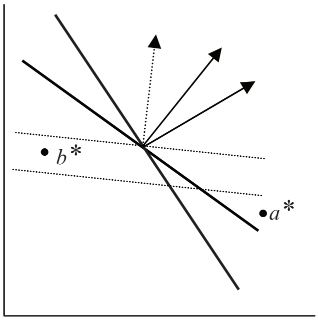


Fig. 3

[2] Though the same sort of parenthesis is used for both intervals of real numbers and vectors in $\mathbb{R}^2$, confusion is avoided by explicitly designating vectors, for example by writing, "the vector $(0,1)$".

[3] That is, with respect to the relative topologies given the Euclidean topologies on $\mathbb{R}^4$ which contains the domain and $\mathbb{R}^2$ which contains the range.

Fig. 4



alternative ranked by both individual agents above another is also ranked above it by the group preference. In fact, this must be the case for all group preferences whose unit normal is contained in the cone spanned by the agents' unit normals. This is shown by the arrows in Fig. 3. Now, consider the case shown in Fig. 4. Both agents rank a^* above b^*, but the group ranks these alternatives in the opposite way. In this case, the unit normal for the group is not in the cone spanned by the agents' unit normals.

Alternatively but equivalently, for agents' preferences $x, y \in S^1$, the group preference must be on the shortest arc in S^1 from x to y. For example, in the case shown in Fig. 2, the group preference must be on the bold arc going anti-clockwise from x to y, and its distance δ' from x along this arc must satisfy $0 \leq \delta' \leq \delta$.

To illustrate a continuous weakly Paretian social welfare function, consider an arbitrary $x \in S^1$, and $f(x, s(x, \delta))$ as δ varies from 0 to 2π. This is shown in Fig. 5 in which values of δ are shown on the horizontal axis and the anticlockwise distance, $\rho_{f(x,s(x,\delta))} - \rho_x$, of the social preference from x is shown on the vertical axis.

The relevant details are all shown in the square with sides of length 2π, which is sub-divided into four sub-squares each with sides of length π. As δ goes from 0 to 2π on the horizontal axis, the height of the S-shaped curve, shown by a continuous line from the point $(0,0)$ to the point $(2\pi, 2\pi)$, shows the anticlockwise distance of the social preference around S^1 from x. At the point $(0,0)$ agents 1 and 2 both have preferences given by $x \in S^1$, and this is also the case at the point $(2\pi, 2\pi)$. At $\delta = \pi$, agent 2's preference is exactly opposite 1's preference in S^1 and exactly the same as the social preference since the S-curve goes through the point (π, π).

Now consider values of δ between 0 and π. In this case, the height of the S-curve is less than the height of the diagonal. This implies that in S^1, the anticlockwise distance from x to the social preference is less that that to 2's preference, thus satisfying the requirement of the cone restriction. This is also true for values of δ between π and 2π. In this case, the height of the S-curve is greater than the height of the diagonal. This implies that in S^1, the anticlockwise distance from x to the social preference is greater than that to 2's preference, and again the requirement of the cone restriction is satisfied.

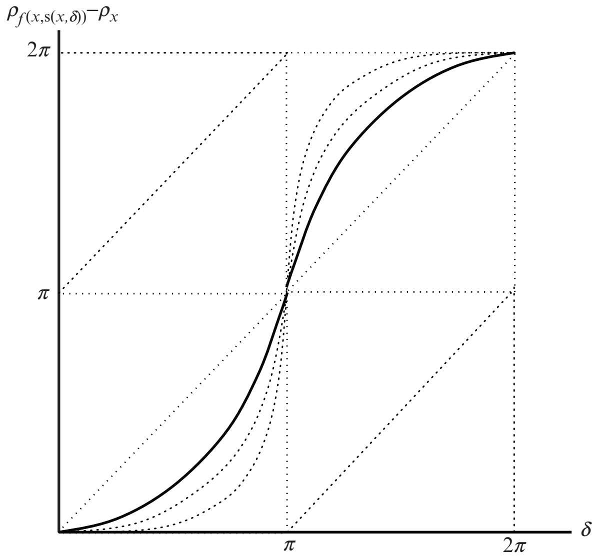


Fig. 5

Another crucial feature of the social welfare function illustrated by the S-curve is that the social preference is never the exact opposite of 2's preference. That is, the point in S^1 that gives the social preference is never exactly opposite the point that gives agent 2's preference. If it were, it would intersect the diagonals of the northwest or southeast sub-squares in Fig. 5, shown by dotted lines.

Now consider the case of a social welfare function that is illustrated by the diagonal of the square. In this case, as the preference of agent 2 rotates anticlockwise from x, the social preference also goes through exactly the same rotation. That is, the preferences of society and agent 2 are always identical. If this is the case for all possible preferences $x \in S^1$ that agent 1 could have, then this social welfare function is *dictatorial* and agent 2 is the *dictator*.

Finally, a crucial role is played by two continuous deformations of the S-curve. In one of these, the continuous S-curve is continuously deformed into the diagonal. Just continuously raise the S-curve for all $\delta \in (0,\pi)$ and lower it for all $\delta \in (\pi,2\pi)$. Such pairs of functions that can be continuously deformed into each other are called *homotopic* functions. Thus, the social welfare function illustrated by the S-curve in Fig. 5 and the social welfare function illustrated by the diagonal are homotopic. Furthermore, agent 2 is then called a *homotopic dictator* for the social welfare function illustrated by the S-curve. Indeed, there must be a homotopic dictator by Chichilnisky's theorem.

The other important observation is that the social welfare function illustrated by the S-curve may also be continuously deformed as shown by the broken lines in Fig. 5. For this continuous deformation, for all $\delta \in (0,\pi)$, the heights of the curves shown by the broken lines decrease towards 0 and for all $\delta \in (\pi,2\pi)$, the heights of the curves shown by the broken lines increase towards 2π. For this class of deformations of the S-curve, apart from its end points, only the point (π,π) remains constant. In other words, it may be concluded that, *if agents do not have opposite preferences, the group preference may be made arbitrarily close to the preference of agent* 1*, even though agent* 2 *remains a homotopic dictator*. It is only if agents have opposite preferences that agent 2 is necessarily asymmetrically powerful.

4 Results

This section makes precise concepts that are used informally in Sect. 3 and the results given in this section justify the conclusion of the Sect. 3.

Projection functions on $S^1 \times S^1$ are continuous social welfare functions that have a special role. They are functions $p_i : S^1 \times S^1 \to S^1, i = 1,2$, such that, for all $x, y \in S^1$, $p_1(x,y) = x$ and $p_2(x,y) = y$. For the social welfare function p_i, $i = 1,2$, i is called the *dictator*. Note that if agent 2 is a *dictator* then, for all $x \in S^1$ and all $\delta \in [0,2\pi]$, $f(x,s(x,\delta)) = s(x,\delta)$.[4]

The concept of homotopic dictatorship first requires the concept of homotopic functions. For arbitrary continuous functions F, G from A to B, F and G are *homotopic* if and only if there is a continuous function $h : A \times [0,1] \to B$ such that, for all $a \in A$, $h(a,0) = F(a)$ and $h(a,1) = G(a)$. Thus, homotopic functions F and G may be continuously deformed into each other. For a social welfare function $f : S^1 \times S^1 \to S^1$, agent $i, i = 1,2$, is a *homotopic dictator* if and only if f and p_i are homotopic. A dictator is a homotopic dictator but not necessarily vice versa.

Next, the cone restriction is made precise. For two agents the satisfaction of the cone restriction is equivalent to the Weak Pareto property, though for more than two agents it is strictly weaker though still sufficient for Chichilnisky's theorem.

For all $x \in S^1$ and $\delta \in [0,2\pi]$, the *closed circular cone* spanned by x and $s(x,\delta)$ is defined as follows:

$$C(x,s(x,\delta)) = \begin{cases} \{y \in S^1 : y = s(x,\delta'), 0 \le \delta' \le \delta\} \text{ if } 0 \le \delta < \pi, \\ \{y \in S^1 : y = s(x,\delta'), \delta \le \delta' \le 2\pi\} \text{ if } \pi < \delta \le 2\pi. \end{cases} \tag{2}$$

A social welfare function $f : S^1 \times S^1 \to S^1$ satisfies the *cone restriction* if and only if, for all $x \in S^1$ and $\delta \in [0,2\pi] \setminus \{\pi\}$, $f(x,s(x,\delta)) \in C(x,s(x,\delta))$. That is, as long as the agents do not have opposite preferences, the social preference is on the shortest arc between them. Note that if agents have opposite preferences so that $\delta = \pi$, the cone restriction does not restrict the social preference. Finally, as noted already, a social welfare function has the *Weak Pareto* property if and only if it satisfies the cone restriction.

The class of social welfare functions that are illustrated in Fig. 5 may now be defined as follows. For all real numbers t, $t \ge 1$:

[4] Dictatorship of agent 1 would require $f(x,s(x,\delta)) = x$.

$$f_t(x,s(x,\delta)) = \begin{cases} s(x,\pi^{1-t}\delta^t) \text{ if } \delta \in [0,\pi], \\ s(x,2\pi - \pi^{1-t}(2\pi-\delta)^t) \text{ if } \delta \in [\pi,2\pi], \end{cases} \tag{3}$$

$f_t : S^1 \times S^1 \to S^1$ are easily shown to be continuous, and their properties are established by the following results.

Proposition 1. *For all $t \geq 1$ and all $x \in S^1$: (i) $f_t(x,s(x,0)) = x$; (ii) $f_t(x,s(x,2\pi)) = x$ and (iii) $f_t(x,s(x,\pi)) = s(x,\pi)$.*

Proof. (i) Substituting $\delta = 0$ into the first part of (3) and then using (1) gives $f_t(x,s(x,0)) = s(x,0) = x$. A similar argument substituting $\delta = 2\pi$ into (3) and again using (1) proves (ii). For (iii), substitute $\delta = \pi$ into both parts of (3) gives what is required. For example, substituting into the first part gives $f_t(x,s(x,\pi)) = s(x,\pi^{1-t}\pi^t) = s(x,\pi)$. □

Proposition 2. *For all $t \geq 1$, $f_t : S^1 \times S^1 \to S^1$ satisfies the cone restriction.*

Proof. There are four cases to consider.

(i) $\delta = 0$: Substituting $\delta = 0$ into (2) gives $C(x,s(x,0)) = \{x\}$. Using (1) and (3) now gives $f_t(x,s(x,0)) = x$, so that $f_t(x,s(x,0)) \in C(x,s(x,0))$.

(ii) $\delta = 2\pi$: A similar argument as used in (i) but beginning by substituting $\delta = 2\pi$ into (2) leads to $f_t(x,s(x,2\pi)) \in C(x,s(x,2\pi))$.

(iii) $\delta \in (0,\pi)$: (3) implies that $f_t(x,s(x,\delta)) = s(x,\pi^{1-t}\delta^t)$. Therefore satisfying the cone restriction in this case requires that $0 \leq \pi^{1-t}\delta^t \leq \delta$ from (2). Since π and δ are both positive, $0 < \pi^{1-t}\delta^t$. Since $\delta < \pi$, it follows that $\pi^{1-t}\delta^t < \delta^{1-t}\delta^t = \delta$.

(iv) $\delta \in (\pi,2\pi)$: (3) implies that $f_t(x,s(x,\delta)) = s(x,2\pi - \pi^{1-t}(2\pi-\delta)^t)$. Therefore satisfying the cone restriction in this case requires that $\delta \leq 2\pi - \pi^{1-t}(2-\delta)^t \leq 2\pi$ from (2). Since $\delta \in (\pi,2\pi)$, it follows that $0 < 2\pi - \delta < \pi$. Using the argument in (iii) with $\delta' = 2\pi - \delta$ instead of δ, it follows that $\pi^{1-t}(2\pi-\delta)^t < 2\pi - \delta$ or, rearranging, $\delta < 2\pi - \pi^{1-t}(2-\delta)^t$ which is part of what is required. For the other part, note that $\pi^{1-t}(2\pi-\delta)^t > 0$ since both π and $2\pi - \delta$ are positive. Therefore, $2\pi - \pi^{1-t}(2\pi-\delta)^t < 2\pi$ and this completes the proof. □

Corollary of Propositions 1 and 2: For all $\delta \in [0,2\pi]$, $f_t(x,s(x,\delta)) \neq -s(x,\delta+\pi)$.

That is, the social preference is never the exact opposite of 2's preference.

Proposition 3. *For all $\delta \in (0,\pi) \cup (\pi,2\pi)$, $\lim_{t\to\infty} f_t(x,s(x,\delta)) = x$.*

Proof. There are two cases to consider.

(i) $\delta \in (0,\pi)$. In this case (3) implies $f_t(x,s(x,\delta)) = s(x,\pi^{1-t}\delta^t)$, so that $\lim_{t\to\infty} f_t(x,s(x,\delta)) = \lim_{t\to\infty} s(x,\pi^{1-t}\delta^t)$. From continuity, $\lim_{t\to\infty} s(x,\pi^{1-t}\delta^t) = s(x,\lim_{t\to\infty} \pi^{1-t}\delta^t)$. Since $\pi^{1-t}\delta^t = \pi(\delta/\pi)^t$, $\lim_{t\to\infty} \pi^{1-t}\delta^t = 0$ since $\lim_{t\to\infty} \pi(\delta/\pi)^t = \pi \lim_{t\to\infty}(\delta/\pi)^t$ and $|(\delta/\pi)| < 1$. Therefore, using (1), $\lim_{t\to\infty} f_t(x,s(x,\delta)) = \lim_{t\to\infty} s(x,\pi^{1-t}\delta^t) = s(x,\lim_{t\to\infty} \pi^{1-t}\delta^t) = s(x,0) = x$.

(ii) $\delta \in (\pi,2\pi)$. In this case, (3) implies $f_t(x,s(x,\delta)) = s(x,2\pi - \pi^{1-t}(2\pi-\delta)^t)$. $\lim_{t\to\infty} s(x,2\pi - \pi^{1-t}(2\pi-\delta)^t) = s(x,\lim_{t\to\infty}(2\pi - \pi^{1-t}(2\pi-\delta)^t))$ from continuity, and furthermore, $\lim_{t\to\infty}(2\pi - \pi^{1-t}(2\pi-\delta)^t) = 2\pi - \lim_{t\to\infty}(\pi^{1-t}(2\pi-\delta)^t)$.

Also $\pi^{1-t}(2\pi-\delta)^t = \pi\left(\frac{2\pi-\delta}{\pi}\right)^t$ and $\lim_{t\to\infty}\pi\left(\frac{2\pi-\delta}{\pi}\right)^t = \pi\lim_{t\to\infty}\left(\frac{2\pi-\delta}{\pi}\right)^t$. Therefore, since $|(\frac{2\pi-\delta}{\pi})^t| < 1$, it follows that $\lim_{t\to\infty}\left(\frac{2\pi-\delta}{\pi}\right)^t = 0$, and this implies that $s(x, \lim_{t\to\infty}(2\pi - \pi^{1-t}(2\pi-\delta)^t)) = s(x, 2\pi)$. Therefore, $f_t(x, s(x,\delta)) = s(x, 2\pi) = x$ which completes the proof. □

Since, for all $t, t \geq 1$, $f_t : S^1 \times S^1 \to S^1$ is continuous and satisfies the cone restriction, it follows from Chichilnisky's theorem that either agent 1 or agent 2 must be a *homotopic dictator*. The final result shows that the homotopic dictator is agent 2.

Proposition 4. *For all $t, t \geq 1$, f_t and p_t are homotopic.*

Proof. First, it will be shown that $f_t(x, s(x,\delta)) \neq s(x,\delta)$. If $\delta = \pi$ then this follows from part (iii) of Proposition 1. If $\delta \neq \pi$ and $f_t(x, s(x,\delta)) = -s(x,\delta)$ then the cone restriction would not be satisfied, contrary to Proposition 3. Therefore, for all $x \in S^1$ and all $\delta \in [0, 2\pi]$, $f_t(x, s(x,\delta)) \neq -s(x,\delta)$. Since $s(x,\delta) = p_2(x, s(x,\delta))$, it follows that $f_t(x, s(x,\delta)) \neq -p_2(x, s(x,\delta))$. Given this, the following homotopy between f_t and p_2 is well defined. For all $x \in S^1$, all $\delta \in [0, 2\pi]$ and all $\lambda \in [0,1]$:

$$h_t(x, s(x,\delta), \lambda) = \frac{\lambda f_t(x, s(x,\delta)) + (1-\lambda)p_2(x, s(x,\delta))}{||\lambda f_t(x, s(x,\delta)) + (1-\lambda)p_2(x, s(x,\delta))||}.$$

Recall from the definitions of f_t and p_2 that all values of these functions lie in the unit circle, S^1, and thus unit norms. It is then straightforward to check that, for all $x \in S^1$ and $\delta \in [0, 2\pi]$, $h_t(x, s(x,\delta), 1) = f_t(x, s(x,\delta))$ and $h_t(x, s(x,\delta), 0) = p_2(x, s(x,\delta))$, and also that h_t is continuous as required. □

Propositions 3 and 4 justify and make precise the claim that concludes Sect. 2. Namely, *if agents do not have opposite preferences, the group preference may be made arbitrarily close to the preference of agent 1, even though agent 2 is a homotopic dictator.*

5 Conclusion

One possible reservation about the analysis in this paper is that it is limited to two agents. However, given the nature of the issue, it is only necessary to establish the conclusion for a simple case, and this has been accomplished. Indeed, Chichilnisky's theorem is not an Arrow-type impossibility result in the sense that it shows that desirable properties entail an undesirable concentration of power.

It may be argued that a homotopic dictator is also a strategic manipulator in the sense of being able to get any particular social preference, for all preferences of other agents. This is indeed the case. It can be seen from Fig. 5 and easily checked from (3), that, for all $t, t \geq 1$, and all $x \in S^1$, $f_t(x, s(x, [0, 2\pi])) = S^1$. Thus, for any possible preference, agent 2 can choose a possibly different preference so that the former is the social preference. This does concentrate a certain sort of power in agent 2. However, if strategic manipulation is of concern, then conditions for its existence can be given directly, and there seems to be no purpose served by tying it to an analysis of homotopic dictatorship.

Acknowledgements This paper was presented at a conference in honor of Kotaro Suzumura at Hitotsubashi University in March, 2006. I am grateful to all participants, especially Peter Hammond and John Weymark. Yongsheng Xu's help with the final manuscript was generous and very much appreciated. I also wish to express my gratitude to Kotaro Suzumura for his general encouragement of my interest in social choice theory in general and in topological social choice theory in particular.

References

Baigent, N. (2003). Topological theories of social choice. In K. Arrow, A. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare: Vol. 2*. Amsterdam: Elsevier

Chichilnisky, G. (1982). The topological equivalence of the Pareto condition and the existence of a dictator. *Journal of Mathematical Economics*, *9*, 223–234

Heal, G. (1997). Social choice and resource allocation: A topological perspective. *Social Choice and Welfare*, *14*, 147–160

Lauwers, L. (2000). Topological social choice. *Mathematical Social Sciences*, *40*, 1–39

MacIntyre, I. D. A. (1998). Two-person and majority continuous aggregation in 2-good space in social choice: A note. *Theory and Decision*, *44*, 199–209

Saari, D. G. (1997). Informational geometry of social choice. *Social Choice and Welfare*, *14*, 211–232

Sen, A. K. (1986). Social choice theory. In K. J. Arrow & M. D. Intriligator (Eds.), *Handbook of mathematical economics, vol. iii* (pp. 1073–1181). Amsterdam: North-Holland

Remarks on Population Ethics

Tomoichi Shinotsuka

1 Introduction

Population ethics is about principles for social evaluation of alternatives with different population sizes. Different environmental policies lead to different population sizes as well as different quality of lives involved. Therefore, as a necessary step towards laying foundations for such policy recommendations, discussing relevant issues on population principles is of critical importance.

One of the most important issues in population ethics has been the repugnant conclusion introduced by Derek Parfit (1976, 1982, 1984). He criticized classical utilitarianism as it implies the following conclusion:

> *The Repugnant Conclusion*: For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are being equal, would be better, even though its members have lives that are barely worth living (see Parfit (1984, p. 388)).

Since then, avoiding the repugnant conclusion has been one of the most important axioms in population ethics. And, this is well-documented by two facts. First, Blackorby, Bossert and Donaldson, leading figures in population ethics, survey the literature concerning the repugnant conclusion in a handbook chapter on social choice and welfare (Blackorby, Bossert, and Donaldson, 2002). Second, there is a book that is entirely devoted to the issues of the repugnant conclusion (Ryberg and Tännsjö, 2004). Despite these, a number of theorists have argued that the repugnant conclusion may not be so repugnant and thus avoiding the conclusion is not that compelling (see Arrhenius (2003, p. 168)). This motivates Arrhenius (2003) to modify the concept of the repugnant conclusion in such a way that even those theorists critical of the original version would find the modified version very hard to accept:

T. Shinotsuka
Graduate School of Humanities and Social Sciences, University of Tsukuba, Japan
e-mail: shino@social.tsukuba.ac.jp

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,
Studies in Choice and Welfare.

The Very Repugnant Conclusion: For any perfectly equal population A with very high positive welfare, and for any number of lives with very negative welfare, there is a population B consisting of the lives with negative welfare and lives with very low positive welfare which is better than population A, other things being equal (Arrhenius (2003, p. 167)).

Arrhenius (2003), then, proceeds to formalize this idea and shows that a version of the mere addition paradox (Parfit, 1984) still holds even if one replaces avoidance of the repugnant conclusion with avoidance of the very repugnant conclusion. In this paper, we investigate what happens to the results on generalized utilitarianism in population ethics established by Blackorby, Bossert, and Donaldson (2004, 2006) when we replace avoidance of the repugnant conclusion with avoidance of the very repugnant conclusion. Arrhenius' own version of the very repugnant conclusion is stated in a model that has considerably different structures than ours. Therefore, we reformulate the very repugnant conclusion in our framework. Consequently, Arrhenius' own version of the very repugnant conclusion and ours are non-comparable.

Arrhenius (2000) introduces two versions of the sadistic conclusion and argues that it should be avoided, too. If a population principle implies that adding people with negative utilities can make a society better off, the conclusion is sadistic. Blackorby, Bossert, and Donaldson (2004) explore logical relations between avoidance of sadistic conclusion and critical-level generalized utilitarian principles. In this paper, we reexamine one of their results.

In Section 2, we introduce the model and state avoidance of the repugnant conclusion and avoidance of the very repugnant conclusion. We show that the incompatibility between Pareto plus and avoidance of the repugnant conclusion is rather robust in the sense that replacing the latter with avoidance of the very repugnant conclusion does not upset the result. In Section 3, we state avoidance of sadistic conclusion. The last section concludes with some remarks.

2 The Model

We work with the model set up by Blackorby, Bossert, Donaldson, and Fleurbaey (1998).

Let $\mathbb{N}$ be the set of natural numbers and let $\mathbb{R}(\mathbb{R}_{++}, \mathbb{R}_{--})$ be the set of all (positive, negative) real numbers. $\mathbb{R}^{\mathbb{N}}$ be the set of all maps from $\mathbb{N}$ into $\mathbb{R}$. Let $\mathcal{N}$ be the set of all non-empty and finite subsets of $\mathbb{N}$. Typical elements of $\mathcal{N}$ are denoted by L, M, N and so on. For each $N \in \mathcal{N}$, $\mathbb{R}^N(\mathbb{R}^N_+)$ is the set of all maps from N into $\mathbb{R}(\mathbb{R}_+)$. Typical elements of $\mathbb{R}^N(\mathbb{R}^N_+)$ are denoted by $u = (u_i)_{i\in N}$, $v = (v_i)_{i\in N}$, $w = (w_i)_{i\in N}$ and so on. For each $N \in \mathcal{N}$, 1_N is the element in $\mathbb{R}^N$ defined by $(1_N)_i = 1$ for each $i \in N$. For all disjoint sets $N, M \in \mathcal{N}$, for all $u = (u_i)_{i\in N}$, $v = (v_i)_{i\in N}$, (u, v) is the element of $\mathbb{R}^{N\cup M}$ defined by $(u, v)_i = u_i$ for $i \in N$ and $(u, v)_j = v_j$ for $j \in M$.

We take a welfarist approach to population ethics: To discuss evaluations of social states, all we need to know is information about population and utility allocations.[1] We employ a comprehensive notion of utilities as indicators of lifetime well-being to avoid counter-intuitive results on the termination of lives.[2] Let $\mathcal{D} = \{(N;u) | N \in \mathcal{N} \text{ and } u \in \mathbb{R}^N\}$. A typical element $(N;u) \in \mathcal{D}$ consists of population N and utility allocation u for N. A social-evaluation ordering is a complete and transitive binary relation R on $\mathcal{D}$.[3] For $(N;u)$, $(M;v) \in \mathcal{D}$, $(N;u)R(M;v)$ means $(N;u)$ is socially at least as good as $(M;v)$. The asymmetric part of R is denoted by P and the symmetric part by I.

An individual considers her or his life neutral if it is as good as the one without any experiences. We employ the convention in population ethics that utilities are normalized so that the zero level of utility represents neutrality.[4]

Repugnant conclusion: For all $N \in \mathcal{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $\varepsilon \in (0, \xi)$, there exists $M \in \mathcal{N}$ such that $M \supset N$ and $(M; \varepsilon 1_M)P(N; \xi 1_N)$.

Avoidance of the repugnant conclusion is the negation of repugnant conclusion.

Avoidance of the repugnant conclusion: There exist $N \in \mathcal{N}$, $\xi \in \mathbb{R}_{++}$, $\varepsilon \in (0, \xi)$ such that $(N; \xi 1_N)R(M; \varepsilon 1_M)$ for all $M \in \mathcal{N}$ such that $M \supset N$.

We formalize the idea of the very repugnant conclusion introduced by Arrhenius (2003) in our framework as follows.

Very repugnant conclusion: For all $N \in \mathcal{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $M \in \mathcal{N}$, for all $\eta \in \mathbb{R}_{--}$, for all $\varepsilon \in (0, \xi)$, there exists $L \in \mathcal{N}$ such that $L \cap M = \emptyset$, and $(L \cup M; \varepsilon 1_L, \eta 1_M)P(N; \xi 1_N)$.

Avoidance of the very repugnant conclusion is the negation of very repugnant conclusion.

Avoidance of the very repugnant conclusion: There exist $N, M \in \mathcal{N}$, $\xi \in \mathbb{R}_{++}$, $\eta \in \mathbb{R}_{--}$ and $\varepsilon \in (0, \xi)$ such that $(N; \xi 1_N)R(L \cup M; \varepsilon 1_L, \eta 1_M)$ for all $L \in \mathcal{N}$ with $L \cap M = \emptyset$.

Let us recall the standard axiom of strong Pareto.

Strong Pareto: For all $N \in \mathcal{N}$ and $u, v \in \mathbb{R}^N$, if $u_i \geq v_i$ for every $i \in N$ and $u_i > v_i$ for some $i \in N$, then $(N;u)P(N;v)$.

Avoidance of the repugnant conclusion together with strong Pareto implies avoidance of the very repugnant conclusion.

Lemma 1. *If a population principle satisfies strong Pareto and the very repugnant conclusion, then it satisfies the repugnant conclusion.*

---

[1] For and against welfarism, see Blackorby, Bossert, and Donaldson (2002).

[2] See, for example, Blackorby et al. (2002) for an account of lifetime well-being.

[3] R is complete if for all $(N;u), (M;v) \in \mathcal{D}$, $(N;u)R(M;v)$ or $(M;v)R(N;u)$. R is transitive if for all $(N;u), (M;v), (L;w) \in \mathcal{D}$, $(N;u)R(M;v)$ and $(M;v)R(L;w)$ imply $(N;u)R(L;w)$.

[4] See Broome (1993, 2004) for a discussion of neutrality and its normalization to zero.

Proof. Let $N \in \mathcal{N}$, $\xi \in \mathbb{R}_{++}$, $M \in \mathcal{N}$, $\eta \in \mathbb{R}_{--}$ and let $\varepsilon \in (0,\xi)$. By the very repugnant conclusion, there exists $L \in \mathcal{N}$ such that $L \cap M = \emptyset$, $L \cap N = \emptyset$, and $(L \cup M; \varepsilon 1_L, \eta 1_M) P(N; \xi 1_N)$. By strong Pareto, $(L \cup M; \varepsilon 1_L, \varepsilon 1_M) P(L \cup M; \varepsilon 1_L, \eta 1_M)$. By transitivity, $(L \cup M; \varepsilon 1_L, \varepsilon 1_M) P(N; \xi 1_N)$. □

Lemma 1 says that avoidance of the repugnant conclusion along with strong Pareto imply avoidance of the very repugnant conclusion.

A population principle R satisfies generalized utilitarianism if there exists a continuous and increasing transformation $g : \mathbb{R} \to \mathbb{R}$ of utilities with $g(0) = 0$ such that for all $NM \in \mathcal{N}$, for all $u = (u_i)_{i \in N} \in \mathbb{R}^N$, for all $v = (v_i)_{i \in M} \in \mathbb{R}^M$, $(N; u) R(M; v)$ if and only if

$$\sum_{i \in N} g(u_i) \geq \sum_{i \in NM} g(v_i).$$

Sikora (1978) introduces an axiom consisting of strong Pareto and the requirement that adding an individual with a utility level above neutrality should be a social improvement. He calls this axiom Pareto plus. Following Blackorby, Bossert, and Donaldson (2006), we retain strong Pareto as a separate axiom and state Pareto plus as follows.

Pareto plus: For all $N \in \mathcal{N}$, for all $u = (u_i)_{i \in N} \in \mathbb{R}^N$, for all $j \in \mathbb{N} \setminus N$, for all $a \in \mathbb{R}_{++}, (N \cup \{j\}; u, a) P(N, u)$.

The following impossibility result provides a yet another criticism against Pareto plus.

Theorem 1. *There exists no population principle that satisfies generalized utilitarianism, Pareto plus and the avoidance of the very repugnant conclusion.*

Proof. Suppose that there exists a generalized-utilitarian population principle satisfying Pareto plus. Let $M, N \in \mathcal{N}$ be sets of cardinalities m and n, respectively, and let $\varepsilon \in (0,\xi)$. Since $g(\varepsilon) > 0$, one can pick $l \in \mathbb{N}$ large enough to have $lg(\varepsilon) + mg(\eta) > (n+m)g(\xi)$. Let $L \in \mathcal{N}$ be a finite set with cardinality l satisfying $L \cap M = \emptyset$.

Thus, by generalized utilitarianism, $(L \cup M; \varepsilon 1_L, \eta 1_M) P(N \cup M; \xi 1_N, \xi 1_M)$. By repeated application of Pareto plus and transitivity, $(N \cup M; \xi 1_N, \xi 1_M) P(N; \xi 1_N)$. By transitivity, $(L \cup M; \varepsilon 1_L, \eta 1_M) P(N; \xi 1_N)$. Thus, the very repugnant conclusion holds. This completes the proof. □

3 Critical-Level Generalized Utilitarianism

A population principle R satisfies critical-level generalized utilitarianism if there exist a critical-level of utility $\alpha \in \mathbb{R}$ and a continuous and increasing transformation $g : \mathbb{R} \to \mathbb{R}$ of utilities with $g(0) = 0$ such that for all $N, M \in \mathcal{N}$, for all $u = (u_i)_{i \in N} \in \mathbb{R}^N$ for all $v = (v_i)_{i \in M} \in \mathbb{R}^M$, $(N; u) R(M; v)$ if and only if

$$\sum_{i \in N}[g(u_i) - g(\alpha)] \geq \sum_{i \in M}[g(v_i) - g(\alpha)].$$

The following theorem is essentially a strengthening of Theorem 3 (i) in Blackorby, Bossert, and Donaldson (2004).

Theorem 2. *A critical-level generalized utilitarian principle satisfies avoidance of the very repugnant conclusion if and only if the critical level α is positive.*

Proof. Let R be a critical-level generalized utilitarian population principle with a continuous and increasing transformation g of utilities and a critical level α. Then, R satisfies avoidance of very repugnant conclusion if and only if there exist $n, m \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, $\eta \in \mathbb{R}_{--}$ and $\varepsilon \in (0, \xi)$ such that $l[g(\varepsilon) - g(\alpha)] + m[g(\eta) - g(\alpha)] \leq n[g(\xi) - g(\alpha)]$ for all $l \in \mathbb{N}$.

Suppose $\alpha > 0$. Let $n = 1$, $\xi = 2\alpha$, $\varepsilon = \alpha/2$. Pick any $\eta \in \mathbb{R}_{--}$. Clearly, $g(\varepsilon) - g(\alpha) = g(\alpha/2) - g(\alpha) < 0$, $g(\eta) - g(\alpha) < 0$ and $0 < g(2\alpha) - g(\alpha) = g(\xi) - g(\alpha)$.

Hence, $l[g(\varepsilon) - g(\alpha)] + m[g(\eta) - g(\alpha)] < n[g(\xi) - g(\alpha)]$ for all $l \in \mathbb{N}$. Thus R satisfies avoidance of very repugnant conclusion.

Suppose $\alpha \leq 0$. Let $n, m \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, $\eta \in \mathbb{R}_{--}$ and $\varepsilon \in (0, \xi)$. Take $l \in \mathbb{N}$ such that $l > [ng(\xi) - mg(\eta) + (m - n)g(\alpha)]/[g(\varepsilon) - g(\alpha)]$. Clearly, $l[g(\varepsilon) - g(\alpha)] + m[g(\eta) - g(\alpha)] > n[g(\xi) - g(\alpha)]$.

Let $L, M, N \in \mathcal{N}$ be finite sets with cardinality l, m and n, respectively. Clearly, $(L \cup M; \varepsilon 1_L, \eta 1_M) P (N; \xi 1_N)$. Thus, the very repugnant conclusion holds. This completes the proof. □

A social evaluation ordering implies the sadistic conclusion if adding people with negative utilities can be better than adding people with positive utilities. The idea is expressed formally as follows.

Sadistic conclusion: There exist $N \in \mathcal{N}$, $M \in \mathcal{N}$, $L \in \mathcal{N}$, $u \in \mathbb{R}^N$, $v \in \mathbb{R}^M_{--}$ and $w \in \mathbb{R}^L_{++}$ such that $(N \cup M; u, v) P (N \cup L; u, w)$.

The following statement is due to Theorem 3 (ii) in Blackorby, Bossert, and Donaldson (2004). Their argument is designed for critical-level utilitarianism but it does not work for its generalized counterpart. So, we shall provide a proof which invokes continuity of utility transformations.

Theorem 3. *A critical-level generalized utilitarian principle satisfies the sadistic conclusion if and only if the critical level a is non-zero.*

Proof. Suppose $\alpha > 0$. Since $g(0) = 0$ and g is increasing, $g(\alpha) > 0$. Since g is continuous at 0, there exist $v_1 \in \mathbb{R}_{--}$ and $w_1 \in \mathbb{R}_{++}$ such that $g(\alpha) > 2g(w_1) - g(v_1)$.

This inequality is equivalent to the following.

$$[g(v_1) - g(\alpha)] > [g(w_1) - g(\alpha)] + [g(w_1) - g(\alpha)].$$

Let i, j, k, l be distinct natural numbers and let $N = \{i\}$, $M = \{j\}$, $L = \{k, l\}$, $u_i = \alpha, v_j = v_1$ and let $w_k = w_l = w_1$.

Then, $v \in \mathbb{R}^M_{--}$, $w \in \mathbb{R}^L_{++}$ and $(N \cup M; u, v) P (N \cup L; u, w)$.

Suppose $\alpha < 0$. Since $g(0) = 0$ and g is increasing, $g(\alpha) < 0$. Since g is continuous at 0, there exist $v_1 \in \mathbb{R}_{--}$ and $w_1 \in \mathbb{R}_{++}$ such that $-g(\alpha) > g(w_1) - 2g(v_1)$.

This inequality is equivalent to the following.

$$[g(v_1) - g(\alpha)] + [g(v_1) - g(\alpha)] > [g(w_1) - g(\alpha)].$$

Let i, j, k, l be distinct natural numbers and let $N = \{i\}$, $M = \{j, k\}$, $L = \{l\}$, $u_i = \alpha, v_j = v_k = v_1$ and let $w_l = w_1$. Then, $v \in \mathbb{R}^M_{--}$ and $w \in \mathbb{R}^L_{++}$ but $(N \cup M; u, v)$ $P(N \cup L; u, w)$.

For the case $\alpha = 0$, the proof is identical to that of Theorem 3 (ii) in Blackorby, Bossert, and Donaldson (2004). □

4 Concluding Remarks

Though we have established a few results on generalized utilitarianism in this paper, the issues of investigating the robustness of impossibility theorems involving avoidance of repugnant conclusions are still wide open. For instance, Blackorby, Bossert, and Donaldson (2006) establish that there exists no anonymous population principle that satisfies minimal increasingness, weak inequality aversion, Pareto plus and avoidance of the repugnant conclusion. What happens to this impossibility result when we replace avoidance of the repugnant conclusion with avoidance of the very repugnant conclusion? Similar questions can be asked for the impossibility results in Blackorby, Bossert, Donaldson, and Fleurbaey (1998).

Acknowledgments The author is grateful to Walter Bossert and Marc Fleurbaey for helpful suggestions on an earlier version of the paper. The current version benefitted considerably from the detailed comments by a referee.

References

Arrhenius, G. (2000). An impossibility theorem for welfarist axiologies. *Economics and Philosophy*, *16*, 247–266

Arrhenius, G. (2003). The very repugnant conclusion. In K. Segerberg & R. Sliwinski (Eds.), *Logic, law, morality: Thirteen essays in practical philosophy in honor of lennart aquivist* (pp. 167–180). Department of Philosophy, Uppsala University: Uppsala Philosophical Studies

Blackorby, C., Bossert, W., & Donaldson, D. (2002). Utilitarianism and the theory of justice. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare, volume 1* (pp. 543–596). Amsterdam: Elsevier

Blackorby, C., Bossert, W., & Donaldson, D. (2004). Critical-level population principles and the repugnant conclusion. In J. Ryberg & T. Tännsjö (Eds.), *The repugnant conclusion: Essays on population ethics* (pp. 45–59). Dordrecht: Kluwer

Blackorby, C., Bossert, W., & Donaldson, D. (2006). Population ethics and the value of life. In M. McGillivray (Ed.), *Inequality, poverty and well-being* (pp. 8–21). London: Macmillan

Blackorby, C., Bossert, W., Donaldson, D., & Fleurbaey, M. (1998). Critical levels and the (reverse) repugnant conclusion. *Journal of Economics*, *67*, 1–15

Broome, J. (1993). Goodness is irreducible to betterness: The evil of death is the value of life. In P. Koslowski & Y. Shionoya (Eds.), *The good and the economical: Ethical choice in economics and management* (pp. 69–83). Berlin: Springer

Broome, J. (2004). *Weighing lives*. Oxford: Oxford University Press

Parfit, D. (1976). On doing the best for our children. In M. Bayles (Ed.), *Ethics and population* (pp. 100–102). Cambridge: Schenkman

Parfit, D. (1982). Future generations, further problems. *Philosophy and Public Affairs*, *11*, 113–172

Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press

Ryberg, J., & Tännsjö, T. (2004). *The repugnant conclusion: Essays on population ethics*. Dordrecht: Kluwer

Sikora, R. I. (1978). Is it wrong to prevent the existence of future generations? In R. L. Basmann & G. F. Rhodes Jr. (Eds.), *Obligations to future generations* (pp. 112–166). Philadelphia: Temple University Press

On Non-Welfarist Social Ordering Functions

Naoki Yoshihara

1 Introduction

Welfarism is defined as a methodology that evaluates social welfare according to the level of satisfaction with regard to individuals' subjective preferences. For this methodology, the criticisms by Dworkin (1981a, 2000), Sen (1979, 1980), and others are well known. They criticized the limited scope of information used to evaluate social welfare in the aforementioned methodology. Moreover, they criticize the welfarist neutral attitude vis-à-vis the problem of what types of preferences are satisfied. There are types of preferences, such as the utility of individual offensive tastes, that of expensive tastes, that of formation of the adaptive preference, or that of cheaper tastes such as in the case of the 'termed housewife,' all of which should be carefully and distinctively treated in the evaluation of social welfare from an ethical point of view. The point of these critiques is that the welfarist evaluation has no concern for such preferential differences.

The problems with welfarist methodology can emerge in a more acute form within the arguments of welfare economics such as the *hypothetical compensation principle*. For instance, let us take the Kaldor principle, which declares an alternative x to be superior to an alternative y if and only if there is another alternative z which is reached through redistribution from x such that z is better than y according to the Pareto principle. This principle is welfarist in nature, since it evaluates policies based only on satisfaction of individuals' preferences of goods and services, and it is also an extension of the Pareto principle as is clear from its definition. It makes clear judgments on policy changes, based on whether or not there is a possibility of potential Pareto improvement of said changes. It is also well known that the validity of executing a policy according to the hypothetical compensation principle

N. Yoshihara
Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo 186-8603, Japan
e-mail: yosihara@ier.hit-u.ac.jp

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

is possibly confirmed by means of such monetary measures as the *gross national income test*, whenever the change in resource allocation caused by the policy is not radical.

According to the above argument, the notion of social welfare that the hypothetical compensation principle considers as its premise is no less than the sum of subjective satisfaction levels obtained from the consumption of 'marketable' goods and services, which can be evaluated by monetary measures. However, the notion of social welfare in general is broad enough to encompass a wide range of ethical viewpoints. The social welfare that *welfarist*'s welfare economics refers to is not as broad, but it is limited to (market) *economic welfare*.

Against such an argument, the following objection may arise from the welfarist position:

It is true that the social welfare analysis utilized in conventional applied economics concerns only the social welfare as the total sum of satisfaction of individuals' preferences over goods and services, which is convertible to monetary value. However, the concept of social welfare can be extended so as to consider the 'utility' from an outcome other than the private consumption of goods and services, extending the domain of the individual utility function if necessary.

Such an approach is one of the ways to expand the limited informational basis of conventional welfarist's welfare economics. However, even using this approach, we would not be able to avoid the above-mentioned criticisms of Sen and Dworkin, because it treats and evaluates everything, including the private consumption of goods and services as well as intrinsic goods such as friendship, through the prism of the same subjective utility functions. The concept of social welfare in this approach is still corresponding solely to the satisfaction of individual subjective preferences. In contrast, this paper argues that the social welfare should be evaluated, not only from the perspective of subjective preferences or tastes, but also from the perspective of welfare and well-being that cannot be grasped by utilizing such preferences. For instance, the viewpoint of "respect for liberal rights" presented in Sen's Liberal Paradox (1970a,b), and his theory of "functioning and capability" (Sen, 1985) offer such concepts of welfare and well-being.

The criticism of welfarism mentioned above becomes relevant not only in the discussion of the criteria of policy evaluation based on the hypothetical compensation principle, but also in the general discussion of welfarist social welfare functions. A social welfare function associates an ordering over social alternatives with each social choice environment. According to the ordering derived from such a function, the society can identify what the most desired policies are, which would realize the most desired social alternatives.

The basic problem in this context is what type of social welfare function should be constructed, and it is in the course of such discussions that the conventional Bergson–Samuelson (B–S) social welfare functions are perceived as problematic. The very reason for this is that in the B–S social welfare functions, the level of individual satisfaction with their subjective preferences is the sole basis of information. However, the orderings over social alternatives given by social welfare functions should reflect an adequate indicator of individuals' well-being. The criticism of welfarism mentioned so far indicates that individual satisfaction with their subjective preferences is no more than one aspect of welfare and therefore a more *pluralistic* viewpoint is necessary.

To treat such a pluralistic viewpoint appropriately, a more comprehensive framework is necessary. As such, we propose the *extended framework* within which not only welfarist notions of individual well-being, but also non-welfarist notions of consequential values, as well as non-consequential values, can be taken into consideration. The extended framework in this paper takes a pair of feasible allocation and allocation rule as an informational basis for the social evaluation of economic policies, and it also proposes to make use of *extended social ordering functions*, each of which associates a social ordering over the set of pairs of feasible allocations and allocation rules with each economic environment. Within such an extended framework, we propose three basic criteria, each of which, respectively, represents: (1) a value of individual autonomy, (2) a value of non-welfaristic consequentialism, and (3) a value of welfarist consequentialism. Moreover, we examine the possibility of extended social ordering functions which satisfy these three pluralistic values.

Recently, there has been some literature such as Blackorby, Bossert, and Donaldson (2005) and Kaplow and Shavell (2001) which also discuss some sorts of 'extended' social ordering functions satisfying some pluralistic values. In their frameworks of social ordering functions, not only the profile of utility information, but also the profile of non-welfaristic information are taken into account. Then, both papers show that even in such frameworks with non-welfaristic information, the feasible class of social ordering functions is reduced to that of the welfarist types only, whenever the Pareto principle is required. Despite the conclusions of Blackorby, Bossert, and Donaldson (2005) and Kaplow and Shavell (2001), we show in this paper that it is possible to construct a desirable social ordering function that has the properties of the welfarist Pareto principle and the non-welfarist criteria. There is no contradiction between the results of these papers and ours, as discussed in Section 5 of this chapter.

Section 2 introduces the basic framework and the basic three axioms. Section 3 discusses a fundamental incompatibility of these three axioms, and Section 4 explores the possibility of second best extended social ordering functions. Section 5 gives some remarks on the related literature such as Blackorby, Bossert, and Donaldson (2005) and Kaplow and Shavell (2001).

2 Beyond the Welfarist Limitation

The need for the pluralistic approach was argued by Parijs (1992, 1993, 1995), Rawls (1971), and Sen (1980, 1985). Based on the normative theories of these three non-welfarists, we propose three basic criteria.

The first criterion is that individual autonomy in contemporary society should be guaranteed. It is a liberal value that contemporary civil societies respect as an important aspect for evaluating individual well-being. In fact, as opposed to the feudal society and the centralized socialist society where individual autonomy is suppressed, the modern civil society might be characterized as having a certain level of political liberalism in legal systems, a certain level of freedom of choice both in political and economic decision-makings, and a certain level of decentralized decision-making mechanisms such as markets, all of which constitute a necessary condition for the guarantee of individual autonomy. Such a viewpoint would suggest a certain constraint over the class of 'desired' social ordering functions. That is to say, if the social economic system cannot guarantee the decentralization and the freedom of choice in decision-making, the welfare that individuals receive under such social situations will not be highly valued by 'desired' social ordering functions, even if the system may support a sufficient level of individual consumption. Thus, this criterion represents a non-consequential value in nature.

The second criterion is that each and every individual should have as much opportunity to do whatever he might want to do as is feasibly possible. This criterion represents a non-welfaristic consequential value in the sense of the following two points: First, although this criterion pertains to social outcomes in terms of individual well-being, it hinges on an objective notion of individual well-being as opposed to welfarist criteria. Second, this criterion does not concern the realization of individual well-being itself, but rather it pertains to the opportunity to pursue or realize individual well-being. Given these points, theories of distributive justice are relevant in the discussion of what concept of individual well-being is appropriate, and of what types of *equity* notions should be applied to the assignment problem of individual opportunity sets.

The third criterion represents a well-known welfarist consequential value such as the Pareto principle. It is worth noting that the standpoint of non-welfarism does not exclude welfarist notions of well-being. I believe that satisfaction of individual subjective preferences is still an important component of the informational basis used to constitute an overall notion of individual well-being. Thus, the Pareto principle is also taken into consideration as a condition imposed on 'desired' social ordering functions.

With this discussion in mind, the question that arises here is whether it is possible to construct a social ordering function consistent with the different pluralistic criteria mentioned above.

2.1 A Framework of Extended Social Ordering Functions

On the basis of the problems propounded in the previous section, in the following section, the notion of *extended social ordering function* is introduced, which is based on the proposal of Gotoh, Suzumura, and Yoshihara (2005). There are two goods, one of which is an input (labor time) $x \in \mathbb{R}_+$ to be used to produce the other good $y \in \mathbb{R}_+$.[1] There is a set $N = \{1, \ldots, n\}$ of agents, where $2 \leq n < +\infty$. Each agent i's

[1] The symbol $\mathbb{R}_+$ denotes the set of non-negative real numbers.

consumption is denoted by $z_i = (x_i, y_i)$, where x_i denotes his labor time, and y_i the amount of his output. All agents face a common upper bound of labor time $\bar{x}$, where $0 < \bar{x} < +\infty$, and so have the same consumption set $Z \equiv [0, \bar{x}] \times \mathbb{R}_+$.

Each agent i's preference is defined on Z and represented by a utility function $u_i : Z \to \mathbb{R}$, which is continuous and quasi-concave on Z, strictly monotonic (decreasing in labor time and increasing in the share of output) on $\overset{\circ}{Z} \equiv [0, \bar{x}) \times \mathbb{R}_{++}$,[2] and $u_i(x, 0) = 0$ for any $x \in [0, \bar{x}]$. We use $\mathcal{U}$ to denote the class of such utility functions.

Each agent i has a *labor skill*, $s_i \in \mathbb{R}_+$. The universal set of skills for all agents is denoted by $\mathcal{S} = \mathbb{R}_+$. The skill $s_i \in \mathcal{S}$ is i's *effective labor supply* per hour measured in efficiency units. It can also be interpreted as i's *labor intensity* exercised in production. Thus, if the agent's *labor time* is $x_i \in [0, \bar{x}]$ and his skill is $s_i \in \mathcal{S}$, then $s_i x_i \in \mathbb{R}_+$ denotes the agent's *effective labor contribution* to production measured in efficiency units. The production technology is a function $f : \mathbb{R}_+ \to \mathbb{R}_+$, that is continuous, strictly increasing, concave, and $f(0) = 0$. For simplicity, we fix f. Thus, an economy is a pair of profiles $\mathbf{e} \equiv (\mathbf{u}, \mathbf{s})$ with $\mathbf{u} = (u_i)_{i \in N} \in \mathcal{U}^n$ and $\mathbf{s} = (s_i)_{i \in N} \in \mathcal{S}^n$. Denote the class of such economies by $\mathcal{E} \equiv \mathcal{U}^n \times \mathcal{S}^n$.

Given $\mathbf{s} = (s_i)_{i \in N} \in \mathcal{S}^n$, an allocation $\mathbf{z} = (x_i, y_i)_{i \in N} \in Z^n$ is *feasible for* $\mathbf{s}$ if $\sum y_i \leq f(\sum s_i x_i)$. We denote by $Z(\mathbf{s})$ the set of feasible allocations for $\mathbf{s} \in \mathcal{S}^n$. An allocation $\mathbf{z} = (z_i)_{i \in N} \in Z^n$ is *Pareto efficient for* $\mathbf{e} = (\mathbf{u}, \mathbf{s}) \in \mathcal{E}$ if $\mathbf{z} \in Z(\mathbf{s})$ and there does not exist $\mathbf{z}' = (z_i')_{i \in N} \in Z(\mathbf{s})$ such that for all $i \in N$, $u_i(z_i') \geq u_i(z_i)$, and for some $i \in N$, $u_i(z_i') > u_i(z_i)$. We use $P(\mathbf{e})$ to denote the set of Pareto efficient allocations for $\mathbf{e} \in \mathcal{E}$.

To complete the description of how our economy functions, what remains is to specify an *allocation rule* which assigns, to each $i \in N$, how many hours he/she works, and how much share of output he/she receives in return. In this chapter, an allocation rule is a *game form* which is a pair $\gamma = (M, g)$, where $M = M_1 \times \cdots \times M_n$ is the set of admissible profiles of individual *strategies*, and g is the *outcome function* which maps each strategy profile $\mathbf{m} \in M$ into a unique outcome $g(\mathbf{m}) \in Z^n$. For each $\mathbf{m} \in M$, $g(\mathbf{m}) = (g_i(\mathbf{m}))_{i \in N}$, where $g_i(\mathbf{m}) = (g_{i1}(\mathbf{m}), g_{i2}(\mathbf{m}))$ and $g_{i1}(\mathbf{m}) \in [0, \bar{x}]$ and $g_{i2}(\mathbf{m}) \in \mathbb{R}_+$ for each $i \in N$.[3] Let Γ be the set of all possible such allocation rules. Given $\gamma = (M, g) \in \Gamma$ and $\mathbf{e} \in \mathcal{E}$, a *non-cooperative game* $(\gamma, \mathbf{e}) \in \Gamma \times \mathcal{E}$ is obtained.

Throughout this chapter, we will focus on the Nash equilibrium concept in our analysis of the performance of game forms as allocation rules. Given $\gamma = (M, g)$, let $\mathbf{m}_{-i} = (m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n) \in M_{-i} \equiv \times_{j \in N \setminus \{i\}} M_j$ for each $\mathbf{m} \in M$ and $i \in N$. Given an $\mathbf{m}_{-i} \in M_{-i}$ and an $m_i' \in M_i$, $(m_i'; \mathbf{m}_{-i})$ is an admissible strategy profile obtained from $\mathbf{m}$ by replacing m_i with m_i'. Given a game $(\gamma, \mathbf{e}) \in \Gamma \times \mathcal{E}$, $\mathbf{m}^* \in M$ is a (*pure strategy*) *Nash equilibrium* if $u_i(g_i(\mathbf{m}^*)) \geq u_i(g_i(m_i, \mathbf{m}_{-i}^*))$ for each $i \in N$ and each $m_i \in M_i$. The set of all Nash equilibria of $(\gamma, \mathbf{e})$ is denoted by $\mathrm{NE}(\gamma, \mathbf{e})$. A feasible allocation $\mathbf{z}^* \in Z(\mathbf{s})$ is a *(pure strategy) Nash equilibrium allocation* of $(\gamma, \mathbf{e})$ if $\mathbf{z}^* = g(\mathbf{m}^*)$ for some $\mathbf{m}^* \in \mathrm{NE}(\gamma, \mathbf{u})$. The set of all Nash equilibrium allocations of $(\gamma, \mathbf{e})$ is denoted by $\tau(\gamma, \mathbf{e})$.

[2] The symbol $\mathbb{R}_{++}$ denotes the set of positive real numbers.

[3] Note that $g_{i1}(m_i, \mathbf{m}_{-i}) = x_i$ is describing the work-hour supply of an individual $i \in N$ that the outcome function designates corresponding to the strategy $(m_i, \mathbf{m}_{-i}) \in M$.

The domain of social preference relations in this paper is given by pairs of allocations and allocation rules as game forms, which we call *extended social alternatives*. The intended interpretation of an extended social alternative, viz., a pair $(\mathbf{z}, \gamma) \in Z^n \times \Gamma$, is that an allocation $\mathbf{z}$ is attained through an allocation rule γ.[4] Moreover, given $\mathbf{e} \in \mathcal{E}$, an extended social alternative $(\mathbf{z}, \gamma) \in Z^n \times \Gamma$ is *realizable* if $\mathbf{z} \in Z(\mathbf{s}) \cap \tau(\gamma, \mathbf{e})$. Let $\mathcal{R}(\mathbf{e})$ denote the set of realizable extended social alternatives under $\mathbf{e} \in \mathcal{E}$.

What we call an *extended social ordering function* (ESOF) is a mapping $Q : \mathcal{E} \twoheadrightarrow (Z^n \times \Gamma)^2$ such that $Q(\mathbf{e})$ is an ordering on $\mathcal{R}(\mathbf{e})$ for every $\mathbf{e} \in \mathcal{E}$.[5] The intended interpretation of $Q(\mathbf{e})$ is that, for any $(\mathbf{z}^1, \gamma^1)$, $(\mathbf{z}^2, \gamma^2) \in \mathcal{R}(\mathbf{e})$, $((\mathbf{z}^1, \gamma^1), (\mathbf{z}^2, \gamma^2)) \in Q(\mathbf{e})$ holds if and only if realizing a feasible allocation $\mathbf{z}^1$ through an allocation rule γ^1 is at least as good as realizing a feasible allocation $\mathbf{z}^2$ through an allocation rule γ^2 according to the social judgments embodied in $Q(\mathbf{e})$. The asymmetric part and the symmetric part of $Q(\mathbf{e})$ will be denoted by $P(Q(\mathbf{e}))$ and $I(Q(\mathbf{e}))$, respectively. The set of all ESOFs will be denoted by $\mathcal{Q}$.

The notion of extended social ordering functions enables us to treat the criteria of individual autonomy, equitable assignment of opportunities in terms of objective well-being, and the Pareto principle in a unified framework. Within the domain $Z^n \times \Gamma$ of social preference orderings derived from ESOFs, the component of game forms constitutes necessary data for formulating orderings based on the criterion of individual autonomy, whereas the data of feasible allocations is relevant to the remaining two criteria.

In the following part, the above-mentioned three criteria are formalized as axioms applicable to ESOFs.

2.1.1 Individual Autonomy in Terms of Choice of Labor Hours

According to the theory of individual liberty that John Stuart Mill proposed (Mill, 1859), there ought to exist in human life a certain *minimal sphere of personal liberty* that should not be interfered with by anybody other than the person in question. Such a sphere should be socially respected and protected as part of individual rights in a liberal society. The question where exactly to draw the boundary between the sphere of personal liberty and that of social authority is a matter of great dispute, and, indeed, how large of a sphere each individual should be entitled to is a controversial issue. Nevertheless, the notion of the inviolability of a minimal sphere of individual liberal rights seems to be deeply ingrained in our social and political fabric.

Thus, a resource allocation policy would rarely be accepted, if its goal or its implementation were incompatible with this minimal guarantee of individual liberty. Such a viewpoint is relevant to our first axiom of extended social ordering functions.

---

[4] The concept of an extended social alternative was introduced by Pattanaik and Suzumura (1996), capitalizing on the suggestion by Arrow (1963, pp. 89–91).

[5] A binary relation R on a universal set X is a *quasi-ordering* if it satisfies *reflexivity* and *transitivity*. An *ordering* is a quasi-ordering satisfying *completeness* as well.

We will discuss what constitutes the minimal guarantee of individual rights in the context of resource allocations that this paper considers.

In the cases of resource allocation problems, the components of political freedom and the non-economic aspects of individual rights might be assumed to be already established. However, there still remains non-established economic parts of individual rights, which might be either treated as *parameters* or as *variables* for relevant resource allocation problems. For instance, we may view *self-ownership* as such a right guaranteeing individual autonomy. The notion of self-ownership originates from the argument of the *Lockean proviso* of John Locke and was used by Nozick (1974) as the principle to justify private ownership in capitalist societies. Nevertheless, the notion of self-ownership can be connected with two versions of *entitlement principles*, that is, the entitlement principles in the weak sense and in the strong sense, as Parijs (1995) discussed.

The entitlement principle in the weaker sense regards self-ownership as a variable for society. Thus, according to this weaker sense of the principle, self-ownership can be seen as freedom or respect for the decision-making of individuals and identified with political freedom and freedom of choice of occupations, etc.[6] In this version, the notion of self-ownership is entirely consistent with redistribution policies which may induce the reconstruction of a given rights structure to achieve a given distributional goal. This is actually the position that Parijs (1995) takes. On the other hand, *the entitlement principle in the stronger sense* no longer views self-ownership as a control variable, but as a parameter which society respects. This stronger sense of the principle can be identified with the arguments made by John Locke. This principle also made a solid basis for the *original appropriation of unowned external resources*, which was proposed by Libertarians including Locke and Nozick.

We also take the same position as Parijs (1995) regarding the notion of self-ownership, and identify the contents of individual liberal rights within the context of resource allocation problems. First, individual liberal rights guarantee freedom of choice in terms of personal consumption. That is, other than the individual in question, no one else has the right to decide the way to dispose of private goods and leisure time available to him/her. W. l.o.g., we should assume that the right of freedom of choice in consumption, in the context of passive freedom, is presumed to be guaranteed in standard economic models of resource allocation problems.

Second, individual liberal rights contain *the right to freedom from forced labor*. This right consists of the freedom to choose a profession, the freedom to enter into an employment contract, etc. However, in simple economic models like this paper, this right may be reduced to the right to choose labor hours, because there is no difference in profession, and all individuals engage in homogeneous labor.

---

[6] In fact, Parijs (1995) insists, "Though not strictly equivalent to 'basic liberties' or 'human rights' as expressed, for example, in Rawls's first principle of justice or in the constitutions of liberal democracies, self-ownership is closely associated with most of them." (Parijs (1995, p. 235, Notes Chaps. 1 and 8)).

The right to choose labor hours is defined as follows:

Definition 1. (Kranich, 1994). *An allocation rule* $\gamma = (M, g) \in \Gamma$ *is* labor-sovereign *if, for all* $i \in N$ *and all* $x_i \in [0, \bar{x}]$, *there exists* $m_i \in M_i$ *such that, for all* $\mathbf{m}_{-i} \in M_{-i}$, $g_{i1}(m_i, \mathbf{m}_{-i}) = x_i$.

Let Γ_{L} denote the subclass of Γ which consists solely of allocation rules satisfying labor sovereignty. Then:

Labor Sovereignty (LS) *For any* $\mathbf{e} \in \mathcal{E}$ *and any* $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$, *if* $\gamma \in \Gamma_{\mathrm{L}}$ *and* $\gamma' \in \Gamma \setminus \Gamma_{\mathrm{L}}$, *then* $((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q(\mathbf{e}))$.

The axiom LS manifests that the extended social alternative with labor sovereign rule should be given a higher priority than any alternative without it. This manifestation should be implemented regardless of what resource allocations the labor sovereign rule or the non-labor sovereign rule realizes as Nash equilibrium outcomes. This expresses an extremely non-consequential value.

Note that if a society executes a non-labor sovereign rule, then such a society might allow the policy-maker to execute some sort of forced labor. The axiom LS rejects such a society and an economic institution. According to this axiom, even an egalitarian redistribution policy would not be accepted unless it were implemented without using forced labor. I believe that the principle of self-ownership based on the weak sense of entitlement principle, and also, even Rawls's *first principle of justice* (Rawls, 1971) should have the form of LS within this economic model.

2.1.2 Evaluation Based on a Criterion of Distributive Justice

Our next criterion is meant to capture an aspect of non-welfaristic egalitarianism. It hinges on what theories of distributive justice we take, which requires an instrument that incorporates the various criteria of distributive justice.

Such an instrument is given by a mapping $J : \mathcal{E} \twoheadrightarrow Z^n \times Z^n$ which associates a binary relation $J(\mathbf{e}) \subseteq Z(\mathbf{s}) \times Z(\mathbf{s})$ with each economy $\mathbf{e} \in \mathcal{E}$. Denote the class of binary relation mappings by $\mathcal{J}$. Such a binary relation $J(\mathbf{e})$ represents a criterion based on a certain theory of distributive justice and alternative feasible allocations are ranked according to this criterion. For instance, if the mapping J represents Sen's theory of equality of capability, then $J(\mathbf{e})$ provides a ranking over alternative capability assignments available to each economy $\mathbf{e} \in \mathcal{E}$, and the rational choice set, derived from this $J(\mathbf{e})$, is regarded as consisting of the most 'equitable' capability assignments under $\mathbf{e} \in \mathcal{E}$.[7] In this case, the ranking made by J should be invariant with respect to the change in the profile of utility functions: that is, $J(\mathbf{e}) = J(\mathbf{e}')$ holds whenever $\mathbf{s} = \mathbf{s}'$ holds. In contrast, if J represents Dworkinian theory of "equality of resources" (Dworkin, 1981b, 2000), J might not have such an invariance property: that is, $J(\mathbf{e}) \neq J(\mathbf{e}')$ may hold even if $\mathbf{s} = \mathbf{s}'$. Moreover, if J represents the theory of "equality of welfare," then $J(\mathbf{e}) = J(\mathbf{e}')$ should hold for any $\mathbf{e}, \mathbf{e}' \in \mathcal{E}$ with $\mathbf{u} = \mathbf{u}'$. In such a way, this mapping can be universally applicable. Moreover, if J represents the criterion of leximin assignment of opportunity sets suggested by

[7] Gotoh, Suzumura, and Yoshihara (2005) argues this case.

Parijs (1995), then J should rationalize feasible allocations satisfying *undominated diversity* (Parijs, 1995).[8] In any case, if J represents a criterion of distributive justice, it should satisfy at least the following requirement.

Minimal Egalitarianism (ME). *For each* $\mathbf{e} = (\mathbf{u}, \mathbf{s}) \in \mathcal{E}$ *and each* $\mathbf{z}, \mathbf{z}' \in Z(\mathbf{s})$ *such that for any* $i, j \in N$, $s_i = s_j$ *and* $x_i = x_j = x'_i = x'_j$, *if there exist* $i, j \in N$ *such that* $y'_i > y_i \geq y_j > y'_j$ *and* $y_k = y'_k$ *for any* $k \neq i, j$, *then* $(\mathbf{z}, \mathbf{z}') \in P(J(\mathbf{e}))$.

Denote the class of binary relation mappings satisfying ME by $\mathcal{J}^{\mathrm{ME}}$. This axiom says that, any transfer from a poorer person to a richer person, given that these two persons provide the same labor hours and other things remain the same, should be undesirable, which shares the same spirit as the *Pigou–Dalton principle*. Note that various types of distributive justice satisfy ME if each of those theories is formulated as a particular $J \in \mathcal{J}$. In fact, Sen's theory of equality of capability, Dworkin's theory of equality of resources, van Parijs's undominated diversity, and even the *equity as no-envy* (Foley, 1967), respectively, could have their own representations within $\mathcal{J}^{\mathrm{ME}}$.

If $J \in \mathcal{J}^{\mathrm{ME}}$ represents a non-welfarist egalitarianism with objective well-being indices, then J might satisfy the following requirement:

Objective Egalitarianism (OE). *For each* $\mathbf{e} = (\mathbf{u}, \mathbf{s}), \mathbf{e}' = (\mathbf{u}', \mathbf{s}') \in \mathcal{E}$, *if* $\mathbf{s} = \mathbf{s}'$, *then* $J(\mathbf{e}) = J(\mathbf{e}')$.

Denote the class of mappings which satisfy ME and OE by $\mathcal{J}^{\mathrm{MOE}}$. Note that any $J \in \mathcal{J}^{\mathrm{MOE}}$ is invariant with respect to change in the profile of individual utility functions. Thus, for instance, Sen's theory of equality of capability has its representation within $\mathcal{J}^{\mathrm{MOE}}$, as formulated in Gotoh, Suzumura, and Yoshihara (2005). In contrast, the representation of van Parijs's undominated diversity does not belong to $\mathcal{J}^{\mathrm{MOE}}$, since undominated diversity needs information about individual utility functions.

Now, our second axiom on ESOFs is given by means of the binary relation mapping $J \in \mathcal{J}$, as follows:

Respect for *J*-based fairness (*J*-RF). *For any* $\mathbf{e} \in \mathcal{E}$ *and any* $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$, *if* $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, $\mathbf{z}' = (\mathbf{x}', \mathbf{y}')$ *and* $\mathbf{x} = \mathbf{x}'$, *then*:

$$\left((\mathbf{z}, \gamma), \left(\mathbf{z}', \gamma'\right)\right) \in Q(\mathbf{e}) \Leftrightarrow \left(\mathbf{z}, \mathbf{z}'\right) \in J(\mathbf{e});$$
$$\left((\mathbf{z}, \gamma), \left(\mathbf{z}', \gamma'\right)\right) \in P(Q(\mathbf{e})) \Leftrightarrow \left(\mathbf{z}, \mathbf{z}'\right) \in P(J(\mathbf{e})).$$

In evaluating the relative wellness of any two extended alternatives, the axiom J-RF focuses only on the corresponding feasible allocations, and under a certain constraint, it claims that the evaluation by the ESOF over extended alternatives should be consistent with the evaluation by J over the corresponding feasible allocations. Here, the "certain constraint" is given by "$\mathbf{z} = (\mathbf{x}, \mathbf{y})$, $\mathbf{z}' = (\mathbf{x}', \mathbf{y}')$ and $\mathbf{x} = \mathbf{x}'$."

---

[8] Given any economy $\mathbf{e} = (\mathbf{u}, \mathbf{s}) \in \mathcal{E}$, a feasible allocation $\mathbf{z} \in Z(\mathbf{s})$ meets undominated diversity if for any $i, j \in N$, there exists at least one individual $k \in N$ such that $u_k(z_i) \geq u_k(z_j)$.

It may well be asked why J-RF imposes the premise $\mathbf{x} = \mathbf{x}'$. The reasons are twofold. First, the choice of individual labor hours is a matter to be left to individual responsibility, and social value judgements should respect individual choices accordingly. Second, if the requirement of J-RF is applied to ESOFs without the premise $\mathbf{x} = \mathbf{x}'$, this instantly causes the incompatibility with the Paretian axiom, which will be discussed later.

As such, J-RF evaluates the desirability of extended social alternatives only from the viewpoint of J-fairness on resource allocations. Unlike the axiom LS, J-RF represents a consequentialist value. This is because this axiom evaluates the extended alternatives based solely on the evaluation of their corresponding resource allocations.

2.1.3 Evaluation Based on the Welfarist Value

Finally, let us introduce the axiom of ESOFs based on the welfarist value. It is defined as an extension of the standard Pareto principle:

Pareto in Allocations (PA). *For any* $\mathbf{e} \in \mathcal{E}$ *and any* $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$, *if* $u_i(z_i) > u_i(z_i')$ *for all* $i \in N$, *then* $((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q(\mathbf{e}))$, *and if* $u_i(z_i) = u_i(z_i')$ *for all* $i \in N$, *then* $((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in I(Q(\mathbf{e}))$.

The axiom PA also focuses only on the feasible allocation in evaluating the relative wellness of any two extended alternatives, and it claims that the evaluation by the ESOF over extended alternatives should be consistent with the Pareto superiority relation or the Pareto indifference relation over feasible allocations. Thus, by the almost same reason as the case of J-RF, PA also represents a position of consequentialism.

3 Impossibility of ESOFs Satisfying LS, *J*-RF, and PA

Now, we are ready to discuss the existence of ESOFs which simultaneously satisfy the axioms LS, J-RF, and PA. According to the technique introduced in Appendix 1 of this paper, we can see this problem by examining the properties of binary relation functions, each of which, respectively, represents one of the above-mentioned axioms.

Let $Q^L : \mathcal{E} \twoheadrightarrow (Z^n \times \Gamma)^2$ be a binary relation function such that for any $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$, the following holds:

$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q^L(\mathbf{e})) \Leftrightarrow \left[\gamma \in \Gamma_{\mathrm{L}} \ \& \ \gamma' \notin \Gamma_{\mathrm{L}}\right];$$
$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in I(Q^L(\mathbf{e})) \Leftrightarrow (\mathbf{z}, \gamma) = (\mathbf{z}', \gamma').$$

Let $Q^{\mathrm{JF}} : \mathcal{E} \twoheadrightarrow (Z^n \times \Gamma)^2$ be a binary relation function such that for any $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$ with $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ and $\mathbf{z}' = (\mathbf{x}', \mathbf{y}')$, the following holds:

$$((\mathbf{z},\gamma),(\mathbf{z}',\gamma')) \in I\left(Q^{\mathrm{JF}}(\mathbf{e})\right) \Leftrightarrow \mathbf{x}=\mathbf{x}' \text{ and } \left(\mathbf{z},\mathbf{z}'\right) \in I\left(J\left(\mathbf{e}\right)\right);$$
$$((\mathbf{z},\gamma),(\mathbf{z}',\gamma')) \in P\left(Q^{\mathrm{JF}}(\mathbf{e})\right) \Leftrightarrow \mathbf{x}=\mathbf{x}' \text{ and } \left(\mathbf{z},\mathbf{z}'\right) \in P\left(J\left(\mathbf{e}\right)\right).$$

Let $Q^{\mathrm{PA}} : \mathcal{E} \twoheadrightarrow (Z^n \times \Gamma)^2$ be a binary relation function such that for any $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z},\gamma),(\mathbf{z}',\gamma') \in \mathcal{R}(\mathbf{e})$, the following holds:

$$\left((\mathbf{z},\gamma),\left(\mathbf{z}',\gamma'\right)\right) \in P\left(Q^{\mathrm{PA}}\left(\mathbf{e}\right)\right) \Leftrightarrow u_i\left(z_i\right) > u_i\left(z_i'\right) \;\; (\forall i \in N);$$
$$\left((\mathbf{z},\gamma),\left(\mathbf{z}',\gamma'\right)\right) \in I\left(Q^{\mathrm{PA}}\left(\mathbf{e}\right)\right) \Leftrightarrow u_i\left(z_i\right) = u_i\left(z_i'\right) \;\; (\forall i \in N).$$

Let a binary relation function Q be called the *minimal relation function w.r.t. an axiom a* if Q satisfies the axiom a, and for any binary relation function Q' satisfying the axiom a, $Q'(\mathbf{e}) \supseteq Q(\mathbf{e})$ holds. Note that each of Q^{L}, Q^{JF}, and Q^{PA} is the minimal relation function w.r.t. each of the axioms LS, J-RF, and PA.

Thus, if there exists an *ESOF Q* which satisfies these three axioms, $Q(\mathbf{e}) \supseteq Q^{\mathrm{L}}(\mathbf{e}) \cup Q^{\mathrm{JF}}(\mathbf{e}) \cup Q^{\mathrm{PA}}(\mathbf{e})$ holds for each $\mathbf{e} \in \mathcal{E}$. Define Q^{LJP} by $Q^{\mathrm{LJP}}(\mathbf{e}) \equiv Q^{L}(\mathbf{e}) \cup Q^{\mathrm{JF}}(\mathbf{e}) \cup Q^{\mathrm{PA}}(\mathbf{e})$ for each $\mathbf{e} \in \mathcal{E}$. According to Proposition 3 in Appendix 1, there exists an ESOF which satisfies the axioms LS, J-RF, and PA if and only if $Q^{\mathrm{LJP}}(\mathbf{e})$ is *consistent* for each $\mathbf{e} \in \mathcal{E}$.[9] Unfortunately, $Q^{\mathrm{LJP}}(\mathbf{e})$ is not consistent for some $\mathbf{e} \in \mathcal{E}$. This is due to the following property:

Proposition 1. *The union of any two of the relations $Q^{\mathrm{L}}(\mathbf{e})$, $Q^{\mathrm{JF}}(\mathbf{e})$, and $Q^{\mathrm{PA}}(\mathbf{e})$ is inconsistent for some $\mathbf{e} \in \mathcal{E}$ and for any $J \in \mathcal{J}^{\mathrm{ME}}$.*

To begin with, the inconsistency of $Q^{\mathrm{L}}(\mathbf{e}) \cup Q^{\mathrm{JF}}(\mathbf{e})$ is easily confirmed by the fact that $Q^{\mathrm{L}}(\mathbf{e})$ is interested solely in the wellness of allocation rules, whereas $Q^{\mathrm{JF}}(\mathbf{e})$ represents the criterion which judges the wellness of extended alternatives, completely ignoring the wellness of allocation rules. A similar argument can be applied to the case of $Q^{\mathrm{L}}(\mathbf{e}) \cup Q^{\mathrm{PA}}(\mathbf{e})$.

How about the binary relation $Q^{\mathrm{JF}}(\mathbf{e}) \cup Q^{\mathrm{PA}}(\mathbf{e})$? This is related to the issue known as the problem of compatibility between fairness and efficiency in resource allocations, and its answer seems to depend on the criteria of distributive justice J. However, as the following example shows, if $J \in \mathcal{J}^{\mathrm{ME}}$, then $Q^{\mathrm{JF}}(\mathbf{e})$ and $Q^{\mathrm{PA}}(\mathbf{e})$ are incompatible, regardless of what type of distributive justice this J represents.[10]

Example 1. Let $N = \{1,2\}$ and $\bar{x} = 3$. The production function is given by $f(x){=}x$ for all $x \in \mathbb{R}_+$. Define an economic environment $\mathbf{e} = (\mathbf{u},\mathbf{s}) \in \mathcal{E}$ as follows: Let $s_i = 1$ for any $i \in N$. Consider the following four feasible allocations: $\mathbf{z}^* = ((1,1),(1,1))$, $\mathbf{z}^{**} = ((2,2),(2,2))$, $\mathbf{z}^*(\theta) = ((1,1+\theta),(1,1-\theta))$, and $\mathbf{z}^{**}(\theta) = ((2,2-\theta),(2,2+\theta))$, where $\theta \in (0,1)$. The utility function of the individual 1 is assumed to have the following property: for $z = (x,y) \in Z$, if $z = z_1^*(\theta)$ or $z = z_1^{**}$, then

$$u_1(z) = (1-\theta+\varepsilon)\cdot(\bar{x}-x)+y, \text{ where } \varepsilon > 0 \text{ is small enough;}$$

[9] The definition of consistent binary relations is given by Definition 4 in Appendix 1 below.

[10] A similar incompatibility result is also obtained by Fleurbaey and Trannoy (2003).

and if $z = z_1^{**}(\theta)$ or $z = z_1^{*}$, then

$$u_1(z) = (1 - \theta - \varepsilon) \cdot (\overline{x} - x) + y.$$

Also, the utility function of the individual 2 is assumed to have the following property: for $z = (x, y) \in Z$ with $x \in [0, 1)$,

$$u_2(z) = (1 - \theta) \cdot (\overline{x} - x) + y;$$

for $z = (x, y) \in Z$ with $x \in [1, \overline{x}]$, if $z = z_2^{**}(\theta)$ or $z = z_2^{*}$, then

$$u_2(z) = (1 + \theta - \varepsilon) \cdot (\overline{x} - x) + y;$$

and if $z = z_2^{*}(\theta)$ or $z = z_2^{**}$, then

$$u_2(z) = (1 + \theta + \varepsilon) \cdot (\overline{x} - x) + y.$$

Figure 1 illustrates such a situation.

Let γ^*, γ^{**}, $\gamma^*(\theta)$, and $\gamma^{**}(\theta)$ be the allocation rules respectively, in which $\mathbf{z}^*$, $\mathbf{z}^{**}$, $\mathbf{z}^*(\theta)$, and $\mathbf{z}^{**}(\theta)$ become, respectively, Nash equilibrium outcomes under $\mathbf{e} \in \mathcal{E}$. Then, if $J \in \mathcal{J}^{\mathrm{ME}}$, its corresponding $Q^{\mathrm{JF}}(\mathbf{e})$ should have:

$$((\mathbf{z}^*, \gamma^*), (\mathbf{z}^*(\theta), \gamma^*(\theta))) \in P\left(Q^{\mathrm{JF}}(\mathbf{e})\right),$$

$$((\mathbf{z}^{**}, \gamma^{**}), (\mathbf{z}^{**}(\theta), \gamma^{**}(\theta))) \in P\left(Q^{\mathrm{JF}}(\mathbf{e})\right).$$

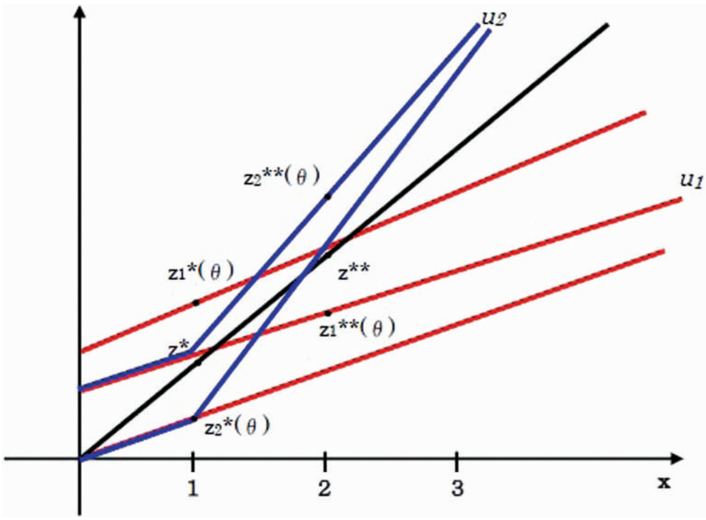


Fig. 1 Example 1 in the consumption space

On the other hand, by the definition of $Q^{\mathrm{PA}}(\mathbf{e})$, we have:

$$((\mathbf{z}^*(\theta), \gamma^*(\theta)), (\mathbf{z}^{**}, \gamma^{**})) \in P\left(Q^{\mathrm{PA}}(\mathbf{e})\right),$$

$$((\mathbf{z}^{**}(\theta), \gamma^{**}(\theta)), (\mathbf{z}^*, \gamma^*)) \in P\left(Q^{\mathrm{PA}}(\mathbf{e})\right).$$

Thus, the binary relation $Q^{\mathrm{JF}}(\mathbf{e}) \cup Q^{\mathrm{PA}}(\mathbf{e})$ is not consistent. □

Thus, this incompatibility can be applied for any J representing any meaningful equity criterion, such as the "equity as no-envy"[11] and Sen's theory of "equality of capability.[12]" This is because any meaningful equity criterion should meet ME.

Note as the above example shows, the incompatibility between $Q^{\mathrm{JF}}(\mathbf{e})$ and $Q^{\mathrm{PA}}(\mathbf{e})$ is obtained by using the *weak Pareto principle* only and without any help of the *Pareto indifference* condition. The Pareto indifference condition is not a crucial factor for this incompatibility.

4 On Possibility of Second-Best Extended Social Ordering Functions

So far, Sect. 3 showed that there is no ESOF which satisfies the three basic axioms, *LS, J-RF, PA*, simultaneously. Then, the next step is to examine the possibility of the second-best ESOFs which satisfy some weaker requirements of the three basic axioms. There are at least two types of methods used to solve this problem. The first method is based on the *pluralistic application of axioms* proposed by Sen and Williams (1982). The second method is based on the *lexicographic application of axioms*. The formal definitions of these approaches are given in Appendix 2.

Here, we focus on the lexicographic application, which sometimes appeared in the literature of normative theories such as Parijs (1995) and Rawls (1971). The lexicographic application takes one *priority order* among axioms, and then makes a ranking between any two alternatives in accordance with the first prior axiom. Then, if the pair of alternatives is non-comparable with respect to the first prior axiom, then the second prior axiom is applied for ranking them. In the following discussion, we will show that even according to this lexicographic application, we cannot yet generally construct a consistent ESOFs. Then, we will consider a further concession to construct the second-best ESOFs. It is a weaker variant of lexicographic application in the sense that the second prior axiom is applied only to a *subset*, not to the whole set, of non-comparable pairs of the first prior axiom. Based on this method, we will show the existence of four types of the second-best ESOFs.

For any $\mathbf{e} \in \mathcal{E}$ and any binary relation $Q(\mathbf{e}) \subseteq (Z^n \times \Gamma)^2$, let $N(Q(\mathbf{e})) \subseteq (Z^n \times \Gamma)^2$ be defined by: for any $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$,

---

[11] The example of the latest successful research is Tadenuma (2002). Also, see Yoshihara (2005).

[12] In fact, Gotoh, Suzumura, and Yoshihara (2005) show that if J represents Sen's theory of "equality of capability," then J-RF is incompatible with PA.

$$\left((\mathbf{z},\gamma),(\mathbf{z}',\gamma')\right) \in N\left(Q\left(\mathbf{e}\right)\right) \Leftrightarrow \left((\mathbf{z},\gamma),(\mathbf{z}',\gamma')\right) \notin Q\left(\mathbf{e}\right) \quad \text{and} \\ \left((\mathbf{z}',\gamma'),(\mathbf{z},\gamma)\right) \notin Q\left(\mathbf{e}\right).$$

Note that this relation is the non-comparable part of $Q\left(\mathbf{e}\right)$. To see the possibility of the second-best ESOFs based on the lexicographic application, let us first examine J-RF *first*-PA *second priority rule*, which is to consider a binary relation function $Q_{\text{lex}}^{J\vdash P} : \mathcal{E} \twoheadrightarrow \left(Z^n \times \Gamma\right)^2$, defined as follows: for any $\mathbf{e} \in \mathcal{E}$,

$$Q_{\text{lex}}^{J\vdash P}\left(\mathbf{e}\right) \equiv Q^{\text{JF}}\left(\mathbf{e}\right) \cup \left[N\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \cap Q^{\text{PA}}\left(\mathbf{e}\right)\right]; \text{ and}$$
$$P\left(Q_{\text{lex}}^{J\vdash P}\left(\mathbf{e}\right)\right) \equiv P\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \cup \left[N\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \cap P\left(Q^{\text{PA}}\left(\mathbf{e}\right)\right)\right].$$

The relation $Q_{\text{lex}}^{J\vdash P}\left(\mathbf{e}\right)$ ranks any two extended alternatives by applying the axiom J-RF in the first place, and if this pair belongs to the non-comparable part of $Q^{\text{JF}}\left(\mathbf{e}\right)$, then PA is applied to rank them. In a similar way, we can also consider PA *first*-**J**-RF *second priority rule*, and define the binary relation function $Q_{\text{lex}}^{P\vdash J} : \mathcal{E} \twoheadrightarrow \left(Z^n \times \Gamma\right)^2$.

Unfortunately, we still obtain the following impossibilities:

Proposition 2. *Both* $Q_{lex}^{J\vdash P}\left(\mathbf{e}\right)$ *and* $Q_{lex}^{P\vdash J}\left(\mathbf{e}\right)$ *are inconsistent for some* $\mathbf{e} \in \mathcal{E}$*, if* $J \in \mathcal{J}^{\text{ME}}$.

This is checked by using the same four feasible allocations and the same economic environment as in Example 1. In fact, we can see in Example 1 that

$$\left(\left(\mathbf{z}^*,\gamma^*\right),\left(\mathbf{z}^*\left(\theta\right),\gamma^*\left(\theta\right)\right)\right) \in N\left(Q^{\text{PA}}\left(\mathbf{e}\right)\right) \text{ and}$$
$$\left(\left(\mathbf{z}^{**},\gamma^{**}\right),\left(\mathbf{z}^{**}\left(\theta\right),\gamma^{**}\left(\theta\right)\right)\right) \in N\left(Q^{\text{PA}}\left(\mathbf{e}\right)\right),$$

which implies that the discussion of inconsistency in Example 1 can be applied to $Q_{\text{lex}}^{P\vdash J}\left(\mathbf{e}\right)$. The same discussion is applied to the binary relation $Q_{\text{lex}}^{J\vdash P}\left(\mathbf{e}\right)$.

As Proposition 2 indicates, we cannot construct any second-best ESWF based on the lexicographic application. To secure the existence of a compatible lexicographic combination of our basic axioms, a further concession seems to be required. As such one, let us consider, for each $\mathbf{e} \in \mathcal{E}$, to choose appropriately a subset $N^*\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right)$ from the whole set of non-comparable parts, $N\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right)$, to make $Q^{\text{JF}}\left(\mathbf{e}\right) \cup \left[N^*\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \cap Q^{\text{PA}}\left(\mathbf{e}\right)\right]$ consistent. Given $J \in \mathcal{J}$ and $\mathbf{x} \in \left[0,\overline{x}\right]^n$, let

$$B\left(J\left(\mathbf{e}\right);\mathbf{x}\right) \equiv \{\left(\mathbf{x},\mathbf{y}\right) \in Z\left(\mathbf{s}\right) \mid \forall \left(\mathbf{x},\mathbf{y}'\right) \in Z\left(\mathbf{s}\right) : \\ \left(\left(\mathbf{x},\mathbf{y}'\right),\left(\mathbf{x},\mathbf{y}\right)\right) \notin P\left(J\left(\mathbf{e}\right)\right)\}.$$

For any $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z},\gamma),(\mathbf{z}',\gamma') \in \mathcal{R}(\mathbf{e})$, let $N^*\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \subsetneq N\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right)$ be given as follows:

$$\left(\left(\mathbf{z},\gamma\right),\left(\mathbf{z}',\gamma'\right)\right) \in N^*\left(Q^{\text{JF}}\left(\mathbf{e}\right)\right) \Leftrightarrow \mathbf{x} \neq \mathbf{x}' \;\&\; \mathbf{z} \in B\left(J\left(\mathbf{e}\right);\mathbf{x}\right) \\ \&\; \mathbf{z}' \in B\left(J\left(\mathbf{e}\right);\mathbf{x}'\right).$$

That is, $((\mathbf{z},\gamma),(\mathbf{z}',\gamma'))$ is a non-comparable pair in the sense of $N^*\left(Q^{\mathrm{JF}}(\mathbf{e})\right)$ if and only if $\mathbf{z}$ and $\mathbf{z}'$ of this pair have mutually different labor supplies, $\mathbf{x}\neq\mathbf{x}'$, and each of them is a $J(\mathbf{e})$-maximal allocation within the same profile of labor supplies. Note that the condition $\mathbf{x}\neq\mathbf{x}'$ implies that $N^*\left(Q^{\mathrm{JF}}(\mathbf{e})\right)$ is a subset of $N\left(Q^{\mathrm{JF}}(\mathbf{e})\right)$. Then, for any $\mathbf{e}\in\mathcal{E}$, let:

$$Q_{\mathrm{lex}}^{*J\vdash P}(\mathbf{e})\equiv Q^{\mathrm{JF}}(\mathbf{e})\cup\left[N^*\left(Q^{\mathrm{JF}}(\mathbf{e})\right)\cap Q^{\mathrm{PA}}(\mathbf{e})\right];$$
$$P\left(Q_{\mathrm{lex}}^{*J\vdash P}(\mathbf{e})\right)\equiv P\left(Q^{\mathrm{JF}}(\mathbf{e})\right)\cup\left[N^*\left(Q^{\mathrm{JF}}(\mathbf{e})\right)\cap P\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\right].$$

The relation $Q_{\mathrm{lex}}^{*J\vdash P}(\mathbf{e})$ ranks any two extended alternatives by applying the axiom J-RF in the first place, and if this pair belongs to the specific non-comparable part, $N^*\left(Q^{\mathrm{JF}}(\mathbf{e})\right)$, then PA is applied to rank them.

Next, let us define a subset $N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$ of $N\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$ to make $Q^{\mathrm{PA}}(\mathbf{e})\cup\left[N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\cap Q^{\mathrm{JF}}(\mathbf{e})\right]$ consistent. For any $\mathbf{e}\in\mathcal{E}$, let $\partial S(\mathbf{e})\equiv\{\overline{\mathbf{u}}\in\mathbb{R}^n\mid\exists\mathbf{z}\in P(\mathbf{e}):u_i(z_i)=\overline{u}_i\ (\forall i\in N)\}$. Then, for each $\overline{\mathbf{u}}\in\partial S(\mathbf{e})$, let us select only one allocation $\mathbf{z}^{\overline{\mathbf{u}}}\in P(\mathbf{e})$ such that for each $i\in N$, $u_i\left(z_i^{\overline{\mathbf{u}}}\right)=\overline{u}_i$. Now, let $P^s(\mathbf{e})\equiv\left\{\mathbf{z}^{\overline{\mathbf{u}}}\right\}_{\overline{\mathbf{u}}\in\partial S(\mathbf{e})}$. By definition, $P^s(\mathbf{e})\subseteq P(\mathbf{e})$. Note that for any $\mathbf{z}^{\overline{\mathbf{u}}},\mathbf{z}^{\overline{\mathbf{u}}'}\in P^s(\mathbf{e})$, $\overline{\mathbf{u}}\neq\overline{\mathbf{u}}'$. Then, for any $\mathbf{e}\in\mathcal{E}$, let $N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\subseteq N\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$ be such that

$$\left((\mathbf{z},\gamma),\left(\mathbf{z}',\gamma'\right)\right)\in N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\Leftrightarrow\mathbf{z},\mathbf{z}'\in P^s(\mathbf{e})\text{ and }\mathbf{z}\neq\mathbf{z}'.$$

That is, $((\mathbf{z},\gamma),(\mathbf{z}',\gamma'))$ is a non-comparable pair in the sense of $N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$ if and only if $\mathbf{z}$ and $\mathbf{z}'$ of this pair are different Pareto efficient allocations, and moreover, their corresponding utility allocations are also different. Note that the latter property follows from $\mathbf{z},\mathbf{z}'\in P^s(\mathbf{e})$. By this property, $\mathbf{z}$ and $\mathbf{z}'$ are Pareto non-comparable, so that $N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$ is actually a subset of $N\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$. Then, for any $\mathbf{e}\in\mathcal{E}$, let:

$$Q_{\mathrm{lex}}^{*P\vdash J}(\mathbf{e})\equiv Q^{\mathrm{PA}}(\mathbf{e})\cup\left[N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\cap Q^{\mathrm{JF}}(\mathbf{e})\right];$$
$$P\left(Q_{\mathrm{lex}}^{*P\vdash J}(\mathbf{e})\right)\equiv P\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\cup\left[N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)\cap P\left(Q^{\mathrm{JF}}(\mathbf{e})\right)\right].$$

The relation $Q_{\mathrm{lex}}^{*P\vdash J}(\mathbf{e})$ ranks over any two extended alternatives by applying the axiom PA in the first place, and if this pair belongs to the specific non-comparable part, $N^*\left(Q^{\mathrm{PA}}(\mathbf{e})\right)$, then J-RF is applied to rank them.

Using these concessive lexicographic binary relation functions, we obtain:

Theorem 1. *Let $J\in\mathcal{J}^{ME}$, and for each $\mathbf{e}\in\mathcal{E}$, $J(\mathbf{e})$ be a continuous quasi-ordering on $Z(\mathbf{s})$*[13] *such that for each $\mathbf{x}\in[0,\overline{x}]^n$, $B(J(\mathbf{e});\mathbf{x})$ is a singleton. Then, there exist at least four ESOFs such that each of which contains either of the following binary relation functions as subrelation mappings:*

(i) $Q_{\mathrm{lex}}^{L\vdash(*P\vdash J)}$; (ii) $Q_{\mathrm{lex}}^{(*P\vdash J)\vdash L}$; (iii) $Q_{\mathrm{lex}}^{L\vdash(*J\vdash P)}$; and (iv) $Q_{\mathrm{lex}}^{(*J\vdash P)\vdash L}$.

[13] A quasi-ordering R is *continuous on* X if for any $x\in X$, its upper and lower contour sets at R is open.

Let $Q^{L\vdash(*P\vdash J)}$ (respectively, $Q^{L\vdash(*J\vdash P)}$) be an ESOF which is obtained as an ordering extension of $Q_{\text{lex}}^{L\vdash(*P\vdash J)}$ (resp. $Q_{\text{lex}}^{L\vdash(*J\vdash P)}$). Note that both $Q^{L\vdash(*P\vdash J)}$ and $Q^{L\vdash(*J\vdash P)}$ are interesting from the viewpoint of non-welfaristic normative theories. Both of them are given by the *weaker* sense of lexicographic application as discussed earlier, and give the first priority to a non-consequential axiom LS rather than the other two consequentialist axioms. Both the Rawlsian two principles of justice combined with the Pareto principle and the Real Libertarianism (Parijs, 1995) combined with the Pareto principle would be formalized as the $Q^{L\vdash(*P\vdash J)}$-type or the $Q^{L\vdash(*J\vdash P)}$-type.

4.1 Rationally Chosen Allocation Rules via ESOFs

In this Section, we characterize allocation rules rationally chosen via $Q^{L\vdash(*P\vdash J)}$ and/or $Q^{L\vdash(*J\vdash P)}$. Given any $Q \in \mathcal{Q}$, the rational choice set $C(Q)$ of allocation rules associated with Q is defined by:

$$\gamma \in C(Q) \Leftrightarrow \forall \mathbf{e} \in \mathcal{E}, \exists \mathbf{z} \in \tau(\gamma, \mathbf{e}) \text{ s.t. } \forall \left(\mathbf{z}', \gamma'\right) \in \mathcal{R}(\mathbf{e}),$$
$$\left((\mathbf{z}, \gamma), \left(\mathbf{z}', \gamma'\right)\right) \in Q(\mathbf{e}).$$

What kinds of allocation rules can be rationally chosen via ESOFs $Q^{L\vdash(*P\vdash J)}$ and/or $Q^{L\vdash(*J\vdash P)}$? If $\gamma \in C\left(Q^{L\vdash(*P\vdash J)}\right) \cup C\left(Q^{L\vdash(*J\vdash P)}\right)$, then what kinds of feasible allocations can this γ implement in Nash equilibria? To examine such questions, let a rationally chosen allocation rule $\gamma \in C\left(Q^{L\vdash(*P\vdash J)}\right) \cup C\left(Q^{L\vdash(*J\vdash P)}\right)$ be called the *first best allocation rule* if for any $\mathbf{e} \in \mathcal{E}$ and any $\mathbf{z} \in \tau(\gamma, \mathbf{e})$, $\mathbf{z}$ is Pareto efficient and $\mathbf{z} \in B(J(\mathbf{e}); \mathbf{x})$. Our particular interest is the existence issue of the first best allocation rule rationalized by $Q^{L\vdash(*P\vdash J)}$ and/or $Q^{L\vdash(*J\vdash P)}$.

Let us call $\gamma = (M, g) \in \Gamma$ an *efficient allocation rule* if, for any $\mathbf{e} \in \mathcal{E}$, $\mathbf{z} \in \tau(\gamma, \mathbf{e})$ implies $\mathbf{z} \in P(\mathbf{e})$. Denote the subclass of Γ which consists solely of efficient allocation rules by Γ_{PE}. Let us call $\gamma = (M, g) \in \Gamma$ a *J-fair allocation rule* if, for any $\mathbf{e} \in \mathcal{E}$, $\mathbf{z} \in \tau(\gamma, \mathbf{e})$ implies $\mathbf{z} \in B(J(\mathbf{e}); \mathbf{x})$. Denote the subclass of Γ which consists solely of J-fair allocation rules by Γ_{JF}. Then:

Theorem 2. *Let $J \in \mathcal{J}^{MOE}$, and for each $\mathbf{e} \in \mathcal{E}$, $J(\mathbf{e})$ be a continuous ordering on $Z(\mathbf{s})$ such that for each $\mathbf{x} \in [0, \bar{x}]^n$, $B(J(\mathbf{e}); \mathbf{x})$ is a singleton. Then, there exists an ESOF $Q^{L\vdash(*J\vdash P)}$ (respectively, $Q^{L\vdash(*P\vdash J)}$) such that for each $\mathbf{e} \in \mathcal{E}$, $Q^{L\vdash(*J\vdash P)}(\mathbf{e}) \supseteq Q_{\text{lex}}^{L\vdash(*J\vdash P)}(\mathbf{e})$ (respectively, $Q^{L\vdash(*P\vdash J)}(\mathbf{e}) \supseteq Q_{\text{lex}}^{L\vdash(*P\vdash J)}(\mathbf{e})$), and $\varnothing \neq C\left(Q^{L\vdash(*J\vdash P)}\right) = \Gamma_{\text{L}} \cap \Gamma_{\text{PE}} \cap \Gamma_{\text{JF}}$ (respectively, $\varnothing \neq C\left(Q^{L\vdash(*P\vdash J)}\right) \supseteq \Gamma_{\text{L}} \cap \Gamma_{\text{PE}} \cap \Gamma_{\text{JF}}$).*

Theorem 2 shows that if $J \in \mathcal{J}^{MOE}$, then the rationally chosen allocation rule via $Q^{L\vdash(*J\vdash P)}$ has the following desired properties: it is labor sovereign, and implements Pareto efficient and J-fair allocations in Nash equilibria. The same property holds for the case of $Q^{L\vdash(*P\vdash J)}$.

This characterization in Theorem 2 is due to the objective egalitarianism of J. If $J \in \mathcal{J}^{\mathrm{MOE}}$, then $B(J(\mathbf{e});\mathbf{x}) = B(J(\mathbf{e}');\mathbf{x})$ for any $\mathbf{x} \in [0,\bar{x}]^n$ and any $\mathbf{e},\mathbf{e}' \in \mathcal{E}$ with $\mathbf{s} = \mathbf{s}'$. This invariance property of $B(J(\cdot);\mathbf{x})$ plays an important role in the existence issue of the first best allocation rules rationalized by $Q^{L\vdash(*J\vdash P)}$ and/or $Q^{L\vdash(*P\vdash J)}$. In contrast, if $J \in \mathcal{J}^{\mathrm{ME}} \backslash \mathcal{J}^{\mathrm{MOE}}$, then the existence of the first best allocation rules rationalized by $Q^{L\vdash(*J\vdash P)}$ and/or $Q^{L\vdash(*P\vdash J)}$ is not necessarily guaranteed. For instance, let $J^{\mathrm{UD}} \in \mathcal{J}^{\mathrm{ME}} \backslash \mathcal{J}^{\mathrm{MOE}}$ represent undominated diversity. Then, there is no first best allocation rule in terms of Pareto efficiency and J^{UD}-fairness, so that the corresponding rationally chosen allocation rule does not have such desired properties.[14]

5 Discussion

In this section, we provide some remarks on the relevant literature of ESOFs. As mentioned in the introduction, Blackorby, Bossert, and Donaldson (2005) and Kaplow and Shavell (2001) also discuss different types of 'extended' social ordering functions.

For instance, Kaplow and Shavell (2001) define any "non-welfarist" axiom as being incompatible with the Pareto Indifference principle. Then, they show that if a social welfare function satisfies continuity and such a "non-welfarist" axiom, then it violates the weak Pareto principle. This is derived from the fact that the continuity of the social welfare function and the weak Pareto principle immediately imply the Pareto Indifference principle.[15] Blackorby, Bossert, and Donaldson (2005) show that if a social welfare function defined over the domain of *multi-profiles* satisfies Universal Domain, Pareto Indifference, and Binary Independence of Irrelevant Alternatives, then it implies Strong Neutrality. Note that Strong Neutrality is regarded as the axiom of Welfarism.

It is well known that, even in the case of B–S social welfare functions with the domains of the utility profiles only, Roberts (1980) and Sen (1977) show that the conventional Arrovian axioms of universal domain, Pareto indifference, and binary independence of irrelevant alternatives together imply strong neutrality. The crucial difference of Blackorby, Bossert, and Donaldson (2005) from Sen (1977) and Roberts (1980) is that the former defines Binary Independence of Irrelevant Alternatives as requiring the social ranking of any two alternatives to depend on not only the utility information but also the non-welfaristic information associated with those two alternatives only. Hence, the independence axiom of Blackorby, Bossert, and Donaldson (2005) is non-welfarist in nature, and it is weaker than the Sen–Roberts independence axiom. Nevertheless, Blackorby, Bossert, and Donaldson (2005) conclude that even in such a framework, the possible social ordering function is only welfarist in nature, if it is required to satisfy the other Arrovian axioms such as

[14] This point is discussed in Yoshihara (2006a).

[15] Kaplow and Shavell (2001) consider the Pareto indifference principle as the definition of welfarism for social welfare functions.

Universal Domain and Pareto Indifference. This seems to provides us with a strong justification of welfarism.

We review the relationship between our approach and the above-mentioned works briefly. First, the welfarist theorem of Blackorby, Bossert, and Donaldson (2005) relies strongly on the axiom of Universal Domain. Such a domain condition cannot directly be applied to the resource allocation problems this paper considers here. For instance, in this paper, all available utility functions are restricted so as to be strongly monotonic and quasi-concave. In fact, as shown in Yoshihara (2006b), if a reasonable domain restriction is imposed, then an ESOF of $Q^{L\vdash(*P\vdash J)}$-type exists and it satisfies the independence axiom of Blackorby, Bossert, and Donaldson (2005). The domain of this ESOF is restricted, because (1) the domain of welfarist information $\mathcal{U}^n$ is restricted to the class of profiles of continuous, strictly monotonic, and quasi-concave utility functions, and (2) the domain of non-welfarist information is also restricted.

Thus, our result on the possibility of the non-welfarist ESOFs is compatible with the result of Blackorby, Bossert, and Donaldson (2005). Moreover, I believe that the universal domain assumption of the non-welfaristic information is not sound from an ethical point of view. This is because a well-being indicator expressing non-welfaristic information should be defined as a binary relation function characterized by a system of axioms,[16] so it needs different formal treatment from the welfarist indicator (individual utility functions) representing capricious subjective preferences.

Second, the conclusion of Kaplow and Shavell (2001) gives us basically the same message as that of Example 1 in this paper. However, Example 1 does not suppose the continuity of social ordering, contrary to the assumption of Kaplow and Shavell (2001). It is also worth noting that this resulting impossibility does not imply a justification of welfarism at all. This is because, as Fleurbaey, Tungodden, and Chang (2003) point out, the Pareto indifference principle and the welfarist axiom are not equivalent. In fact, our ESOF $Q^{(*P\vdash J)\vdash L}$ in Theorem 1 satisfies the weak Pareto principle as well as the Pareto indifference principle, and it also has the properties of the two types of non-welfarism (J-RF and LS). However, this type of function does not meet the continuity axiom. This implies that the real factor inducing the impossibility is not the trade-off between welfarism and non-welfarism, but rather the requirement of continuity.

To summarize this, despite the conclusions of Blackorby, Bossert, and Donaldson (2005) and Kaplow and Shavell (2001), it is eminently possible to construct a desirable social ordering function that has the properties of the welfarist Pareto principle and the non-welfarist criteria.

[16] An example of this is a series of works by Pattanaik and Xu (1990), where the ranking over opportunity sets is characterized by the system of axioms which reflect the viewpoint of "freedom of choice."

6 Conclusion

In the earlier sections, we discussed that the welfarist's framework developed in traditional welfare economics provided us with a rather limited perspective for social evaluation, so a more comprehensive framework would be necessary. As such, we proposed the extended framework within which not only welfarist consequential values, but also non-welfarist consequential values and non-consequential values can be taken into consideration. Moreover, we introduced extended social ordering functions and, as axioms of which, Labor Sovereignty, Respect for J-based Fairness based on non-welfaristic well-being notions, and the Pareto principle. Then, we showed a method of applying these axioms based on a weaker lexicographic approach, by which some consistent extended social ordering functions can be constructed in order to be compatible with the above three values.

Appendix 1

In this Appendix 1, the elementary properties of binary relations are provided, which constitute an analytical technique useful to consider the existence issue of ESOFs. Let X be the universal set of any alternatives and R be a binary relation defined over this set. If R satisfies *completeness* and *transitivity* in particular, we shall call it an *ordering*. Also:

Definition 2. *An axiom a is* represented by a binary relation $R^a \subseteq X \times X$ *if the following condition holds: for any* $\mathbf{x}, \mathbf{x}' \in X$,

$$\begin{aligned}(\mathbf{x},\mathbf{x}') \in R^a &\Leftrightarrow \text{according to the axiom } a, \mathbf{x} \text{ is at least} \\ &\quad\ \text{as desired as } \mathbf{x}'; \\ (\mathbf{x},\mathbf{x}') \in P(R^a) &\Leftrightarrow \text{according to the axiom } a, \mathbf{x} \text{ is strictly} \\ &\quad\ \text{more desirable than } \mathbf{x}'.\end{aligned}$$

In general, the binary relation representing an axiom is not necessarily a complete ordering. In the following discussion, let us denote the representation of the axiom a by R^a. Then, let us see how an ordering $R \subseteq X \times X$ satisfies an axiom in general.

Definition 3. *A binary relation R satisfies a class of axioms $\{a^\lambda\}_{\lambda\in\Lambda}$ if the following condition holds*:

$$R \supseteq \left[\cup_{\lambda\in\Lambda} R^{a^\lambda}\right] \text{ and } P(R) \supseteq \left[\cup_{\lambda\in\Lambda} P\left(R^{a^\lambda}\right)\right].$$

As Definition 3 suggests, a binary relation R satisfies axioms $a^1, \ldots, a^m$ if and only if it contains all of the axiom-representing relations $R^{a^1}, \ldots, R^{a^m}$ as its subrelations.

Given a class of axioms on ordering relations, one interesting problem is to examine whether there exists an ordering relation that satisfies all of these axioms. To discuss this question, the following notion is crucial.

Definition 4. (Suzumura, 1976). *A binary relation $R \subseteq X \times X$ is* consistent *if, for any finite subset $\{x^1, x^2, \ldots, x^t\}$ of X, the following condition does not hold:*

$$[\,(x^1, x^2) \in P(R), \left(x^k, x^{k+1}\right) \in R\ (\forall k = 2, \ldots, t-1)] \Rightarrow \left(x^t, x^1\right) \in R.$$

Proposition 3. *There exists an ordering relation $R \subseteq X \times X$ which satisfies a class of axioms $\{a^{\lambda}\}_{\lambda \in \Lambda}$ if and only if $\left[\cup_{\lambda \in \Lambda} R^{a^{\lambda}}\right]$ is consistent.*

According to Proposition 3, it is sufficient to confirm whether or not the union of the axiom-representing relations $\left\{R^{a^{\lambda}}\right\}_{\lambda \in \Lambda}$ meets the consistency. This condition can be useful when we discuss the existence of ESOFs satisfying some classes of axioms.

Appendix 2

1. The *pluralistic application of axioms* (Sen and Williams, 1982)

Given any two axioms a and b which are mutually incompatible, the *pluralistic application of axioms* is to construct a binary relation $R^{a \cap b} \subseteq X \times X$ which is defined as: $R^{a \cap b} \equiv R^a \cap R^b$ and $P\left(R^{a \cap b}\right) \equiv \left[P\left(R^a\right) \cap R^b\right] \cup \left[R^a \cap P\left(R^b\right)\right]$. Then, $R^{a \cap b}$ becomes consistent whenever R^a and R^b are respectively, consistent. Thus, this kind of second best resolution is to consider an ordering extension of $R^{a \cap b}$.

2. The *lexicographic application of axioms*

Given any binary relation R, let $N(R) \subseteq X \times X$ be defined as follows: for any $\mathbf{x}, \mathbf{x}' \in X$, $(\mathbf{x}, \mathbf{x}') \in N(R) \Leftrightarrow (\mathbf{x}, \mathbf{x}') \notin R$ and $(\mathbf{x}', \mathbf{x}) \notin R$.[17] Given any two axioms a and b which are mutually incompatible, the *lexicographic application of axioms* is to construct a binary relation $R_{\text{lex}}^{a \vdash b} \subseteq X \times X$ which is defined by: for any $x, x' \in X$,

$$\left(x, x'\right) \in R_{\text{lex}}^{a \vdash b} \Leftrightarrow \left(x, x'\right) \in R^a \cup \left[N\left(R^a\right) \cap R^b\right]; \text{ and}$$
$$\left(x, x'\right) \in P\left(R_{\text{lex}}^{a \vdash b}\right) \Leftrightarrow \left(x, x'\right) \in P\left(R^a\right) \cup \left[N\left(R^a\right) \cap P\left(R^b\right)\right].$$

That is, suppose that the society gives a priority to axiom a rather than to axiom b. Then, for any two alternatives, a is applied by $R_{\text{lex}}^{a \vdash b}$ in the first place to make a comparison between them, and b is applied only if these two alternatives are incomparable by a. This is called *axiom a first-axiom b second priority rule*.

[17] The definition of this binary relation is based on Suzumura (2004).

According to Proposition 3, an ordering extension of $R_{\text{lex}}^{a\vdash b}$ is possible whenever $R_{\text{lex}}^{a\vdash b}$ is consistent. Unfortunately, however, the consistency of $R_{\text{lex}}^{a\vdash b}$ is not guaranteed in general. Thus, we need an algorithm to see what properties of the axioms a and/or b can make $R_{\text{lex}}^{a\vdash b}$ consistent.

Suppose that $R_{\text{lex}}^{a\vdash b}$ is not consistent. Our strategy is to choose an appropriate subset $N^*(R^a)$ from $N(R^a)$ such that

$$R_{\text{lex}}^{*a\vdash b} \equiv R^a \cup \left[N^*(R^a) \cap R^b\right]$$

becomes consistent. Then, the problem is to identify what conditions this $N^*(R^a)$ should satisfy so as to make $R_{\text{lex}}^{*a\vdash b}$ consistent. A general solution to this problem is given by Yoshihara (2005), and here we introduce a corollary of this solution given in Yoshihara (2005).

Definition 5. (Yoshihara, 2005). *Given a binary relation $R \subseteq X \times X$, a subset $N^*(R) \subseteq N(R)$ is said to* be connected *if for any $(x,x'), (y,y') \in N^*(R)$, there exists $\left\{z^1,\ldots,z^t\right\} \subseteq X$ such that $z^1 = x'$, $z^t = y$, and $\left(z^k, z^{k+1}\right) \in N^*(R)$ holds for any $k = 1,\ldots,t-1$.*

Proposition 4. (Yoshihara, 2005). *Let R^a be a quasi-ordering over X. Then,* if *the relation $N^*(R^a) \subseteq N(R^a)$ is* transitive *and* connected*, then the relation $R_{\text{lex}}^{*a\vdash b} \subseteq X \times X$ is* consistent *for any quasi-ordering $R^b \subseteq X \times X$.*

Appendix 3: Proofs of Theorems 1 and 2

Proof of Theorem 1. We can see that for any $\mathbf{e} \in \mathcal{E}$, both $N^*\left(Q^{\text{JF}}(\mathbf{e})\right)$ and $N^*\left(Q^{\text{PA}}(\mathbf{e})\right)$ are *connected* and *transitive*, where the definition of connectedness is given in Definition 5 of Appendix 2. Hence, by Proposition 4 of Appendix 2, both $Q_{\text{lex}}^{*J\vdash P}$ and $Q_{\text{lex}}^{*P\vdash J}$ are consistent binary relation functions. □

Lemma 1. *For each $\mathbf{e} \in \mathcal{E}$, $\cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})$ has a closed graph in $Z(\mathbf{s})$.*

Proof. It can be shown in a similar way to Lemma 4 in Gotoh et al. (2005). □

Lemma 2. (Yoshihara, 2000). *For each $\mathbf{s} \in \mathcal{S}^n$, let $h : [0,\bar{x}]^n \to \mathbb{R}_+^n$ be a continuous function such that, for each $\mathbf{x} \in [0,\bar{x}]^n$, $h(\mathbf{x}) = \mathbf{y}$ and $f(\sum s_i x_i) = \sum y_i$. Then, for any $\mathbf{e} = (\mathbf{u},\mathbf{s}) \in \mathcal{E}$, there exists $\mathbf{x}^* \in [0,\bar{x}]^n$ such that $(\mathbf{x}^*, h(\mathbf{x}^*))$ is a* Pareto efficient allocation for $\mathbf{e}$.

Proof. See Proposition 3 in Gotoh, Suzumura, and Yoshihara (2005). □

Lemma 3. *For each $\mathbf{e} = (\mathbf{u},\mathbf{s}) \in \mathcal{E}$, there exists a Pareto efficient allocation $\mathbf{z}^* \in Z(\mathbf{s})$ such that $\mathbf{z}^* \in \cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})$.*

Proof. See Lemma 5 in Gotoh, Suzumura, and Yoshihara (2005). □

Lemma 4. (Yoshihara, 2000). *Given* $J \in \mathcal{J}^{\text{MOE}}$*, for each* $\mathbf{e} \in \mathcal{E}$*, let* $PB^J(\mathbf{e}) \equiv P(\mathbf{e}) \cap \left[\cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})\right]$*. Let* $h : [0,\bar{x}]^n \to \mathbb{R}^n_+$ *be a continuous function such that, for each* $\mathbf{x} \in [0,\bar{x}]^n$*,* $h(\mathbf{x}) = \mathbf{y}$ *and* $f(\sum s_i x_i) = \sum y_i$*. Then, there exists a game form* $\gamma \in \Gamma_{\text{L}}$ *such that, for any* $\mathbf{e} \in \mathcal{E}$*,* $\mathbf{z} \in \tau(\gamma, \mathbf{e})$ *holds if and only if* $\mathbf{z} = (\mathbf{x}, h(\mathbf{x}))$*, and it is Pareto efficient.*

Proof. See Proposition 4 in Gotoh, Suzumura, and Yoshihara (2005). □

Lemma 5. *There exists* $\gamma^* \in \Gamma_{\text{L}}$ *such that* $\tau(\gamma^*, \mathbf{e}) = PB^J(\mathbf{e})$ *for all* $\mathbf{e} \in \mathcal{E}$*.*

Proof. See the proof of Theorem 1 in Gotoh, Suzumura, and Yoshihara (2005).[18] □

Proof of Theorem 2. Given $\mathbf{e} \in \mathcal{E}$, let $S(\mathbf{e})$ be the utility possibility set of feasible allocations, and $\partial S(\mathbf{e})$ be its boundary. Since every utility function is strictly increasing, $\partial S(\mathbf{e})$ is the set of Pareto efficient utility allocations.

(1) Consider the case of $Q^{L\vdash(*J\vdash P)}$. Define an ordering $V(\mathbf{e})$ over $S(\mathbf{e})$ as follows:

(1) If $\bar{\mathbf{u}}, \bar{\mathbf{u}}' \in \partial S(\mathbf{e})$, then $(\bar{\mathbf{u}}, \bar{\mathbf{u}}') \in I(V(\mathbf{e}))$.
(2) For any $\bar{\mathbf{u}}, \bar{\mathbf{u}}' \in S(\mathbf{e})$, there exist $\mu, \mu' \in [1, +\infty)$ such that $\mu \cdot \bar{\mathbf{u}}, \mu' \cdot \bar{\mathbf{u}}' \in \partial S(\mathbf{e})$ and $(\bar{\mathbf{u}}, \bar{\mathbf{u}}') \in V(\mathbf{e})$ if and only if $\mu \leq \mu'$. This ordering $V(\mathbf{e})$ is continuous over $S(\mathbf{e})$.

Define a complete ordering $R_{\mathbf{e},J}$ over $\cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})$ as follows: for any $\mathbf{z}, \mathbf{z}' \in \cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})$, $(\mathbf{z}, \mathbf{z}') \in R_{\mathbf{e},J} \Leftrightarrow (\mathbf{u}(\mathbf{z}), \mathbf{u}(\mathbf{z}')) \in V(\mathbf{u})$. This ordering $R_{\mathbf{e},J}$ is continuous on $\cup_{\mathbf{x}\in[0,\bar{x}]^n} B(J(\mathbf{e});\mathbf{x})$, and its maximal elements constitute $PB^J(\mathbf{e})$. Given $J \in \mathcal{J}^{\text{MOE}}$, let $J(\mathbf{e};\mathbf{x})$ be the restriction of $J(\mathbf{e})$ into the set of feasible allocations with $\mathbf{x}$.

Consider a binary relation $R_{\mathbf{e},J} \cup \left[\cup_{\mathbf{x}\in[0,\bar{x}]^n} J(\mathbf{e};\mathbf{x})\right]$ over $Z(\mathbf{s})$. It is easy to see that this binary relation is consistent, so that there exists an ordering extension $R^*_{\mathbf{e},J}$ of $R_{\mathbf{e},J} \cup \left[\cup_{\mathbf{x}\in[0,\bar{x}]^n} J(\mathbf{e};\mathbf{x})\right]$ by Suzumura's (1976) extension theorem. Based upon this $R^*_{\mathbf{e},J}$, let us consider an ordering function $Q^{L\vdash(*P\vdash J)}$ as follows: for each $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$,

(1) If $\gamma \in \Gamma_{\text{L}}$ and $\gamma' \in \Gamma \setminus \Gamma_{\text{L}}$, then $((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q^{L\vdash(*J\vdash P)}(\mathbf{e}))$.
(2) If either $\gamma, \gamma' \in \Gamma_{\text{L}}$ or $\gamma, \gamma' \in \Gamma \setminus \Gamma_{\text{L}}$, then

$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in Q^{L\vdash(*J\vdash P)}(\mathbf{e}) \Leftrightarrow (\mathbf{z}, \mathbf{z}') \in R^*_{\mathbf{e},J},$$
$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q^{L\vdash(*J\vdash P)}(\mathbf{e})) \Leftrightarrow (\mathbf{z}, \mathbf{z}') \in P(R^*_{\mathbf{e},J}).$$

Note that $Q^{L\vdash(*J\vdash P)}(\mathbf{e})$ is complete and transitive, and $Q^{L\vdash(*J\vdash P)}(\mathbf{e}) \supseteq Q^{L\vdash(*J\vdash P)}_{\text{lex}}(\mathbf{e})$ for each $\mathbf{e} \in \mathcal{E}$, by the definition. Finally, we can see that $C\left(Q^{L\vdash(*J\vdash P)}\right) = \Gamma_{\text{L}} \cap \Gamma_{PE} \cap \Gamma_{\text{JF}} \ni \gamma^*$.

(2) Consider the case of $Q^{L\vdash(*P\vdash J)}$. For each $\bar{\mathbf{u}} \in \partial S(\mathbf{e})$, let us select only one allocation $\mathbf{z}^{\bar{\mathbf{u}}} \in P(\mathbf{e})$ such that for each $i \in N$, $u_i\left(z_i^{\bar{\mathbf{u}}}\right) = \bar{u}_i$, and if for this $\bar{\mathbf{u}} \in \partial S(\mathbf{e})$,

[18] Also, see Yamada and Yoshihara (2007).

there exists $\mathbf{z} \in PB^J(\mathbf{e})$ such that for each $i \in N$, $u_i(z_i) = \bar{u}_i$, then choose such an allocation as $\mathbf{z}^{\bar{\mathbf{u}}}$. Now, let $P^s(\mathbf{e}) \equiv \left\{\mathbf{z}^{\bar{\mathbf{u}}}\right\}_{\bar{\mathbf{u}} \in \partial S(\mathbf{e})}$. By definition, $P^s(\mathbf{e}) \subseteq P(\mathbf{e})$ and $P^s(\mathbf{e}) \cap PB^J(\mathbf{e}) \neq \varnothing$. Note that for any $\mathbf{z}^{\bar{\mathbf{u}}}, \mathbf{z}^{\bar{\mathbf{u}}'} \in P^s(\mathbf{e})$, $\bar{\mathbf{u}} \neq \bar{\mathbf{u}}'$.

Define an ordering $R^0_{\mathbf{e},J}$ over $Z(\mathbf{s})$ as follows: for any $\mathbf{z}, \mathbf{z}' \in Z(\mathbf{s})$,

(i) $(\mathbf{z}, \mathbf{z}') \in I\left(R^0_{\mathbf{e},J}\right)$ if $\mathbf{z}, \mathbf{z}' \in P^s(\mathbf{e}) \cap PB^J(\mathbf{e})$
(ii) $(\mathbf{z}, \mathbf{z}') \in P\left(R^0_{\mathbf{e},J}\right)$ if $\mathbf{z} \in P^s(\mathbf{e}) \cap PB^J(\mathbf{e})$ and $\mathbf{z}' \in P^s(\mathbf{e}) \setminus PB^J(\mathbf{e})$
(iii) $(\mathbf{z}, \mathbf{z}') \in R^0_{\mathbf{e},J} \Leftrightarrow (\mathbf{z}, \mathbf{z}') \in J(\mathbf{e})$ for $\mathbf{z}, \mathbf{z}' \in P^s(\mathbf{e}) \setminus PB^J(\mathbf{e})$
(iv) $(\mathbf{z}, \mathbf{z}') \in P\left(R^0_{\mathbf{e},J}\right)$ if there exist $\mu, \mu' \in [1, +\infty)$ such that $\mu \cdot \mathbf{u}(\mathbf{z}), \mu' \cdot \mathbf{u}(\mathbf{z}') \in \partial S(\mathbf{u})$ and $\mu < \mu'$ and
(v) $(\mathbf{z}, \mathbf{z}') \in R^0_{\mathbf{e},J}$ if there exist $\mu, \mu' \in [1, +\infty)$ such that $\mu \cdot \mathbf{u}(\mathbf{z}), \mu' \cdot \mathbf{u}(\mathbf{z}') \in \partial S(\mathbf{u})$ and $\mu = \mu'$, and $\left(\mathbf{z}^{\mu \cdot \mathbf{u}(\mathbf{z})}, \mathbf{z}^{\mu' \cdot \mathbf{u}(\mathbf{z}')}\right) \in R_{\mathbf{e},J}$ for $\mathbf{z}^{\mu \cdot \mathbf{u}(\mathbf{z})}, \mathbf{z}^{\mu' \cdot \mathbf{u}(\mathbf{z}')} \in P^s(\mathbf{e})$

Denote the set of maximal elements over $Z(\mathbf{s})$ in terms of $R^0_{\mathbf{e},J}$ by $B\left(R^0_{\mathbf{e},J}\right) \subseteq Z(\mathbf{s})$. By definition, $PB^J(\mathbf{e}) \subseteq B\left(R^0_{\mathbf{e},J}\right)$.

Based upon this $R^0_{\mathbf{e},J}$, let us consider an ordering function $Q^{L\vdash(*P\vdash J)}$ as follows: for each $\mathbf{e} \in \mathcal{E}$ and any $(\mathbf{z}, \gamma), (\mathbf{z}', \gamma') \in \mathcal{R}(\mathbf{e})$,

(1) If $\gamma \in \Gamma_{\mathrm{L}}$ and $\gamma' \in \Gamma \setminus \Gamma_{\mathrm{L}}$, then $((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q^{L\vdash(*P\vdash J)}(\mathbf{e}))$
(2) If either $\gamma, \gamma' \in \Gamma_{\mathrm{L}}$ or $\gamma, \gamma' \in \Gamma \setminus \Gamma_{\mathrm{L}}$, then

$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in Q^{L\vdash(*P\vdash J)}(\mathbf{e}) \Leftrightarrow (\mathbf{z}, \mathbf{z}') \in R^0_{\mathbf{e},J},$$
$$((\mathbf{z}, \gamma), (\mathbf{z}', \gamma')) \in P(Q^{L\vdash(*P\vdash J)}(\mathbf{e})) \Leftrightarrow (\mathbf{z}, \mathbf{z}') \in P(R^0_{\mathbf{e},J}).$$

Note that $Q^{L\vdash(*P\vdash J)}(\mathbf{e})$ is complete and transitive, and $Q^{L\vdash(*P\vdash J)}(\mathbf{e}) \supseteq Q^{L\vdash(*P\vdash J)}_{\mathrm{lex}}(\mathbf{e})$ for each $\mathbf{e} \in \mathcal{E}$, by the definition. Finally, we can see that $C\left(Q^{L\vdash(*P\vdash J)}\right) \supseteq \Gamma_{\mathrm{L}} \cap \Gamma_{PE} \cap \Gamma_{\mathrm{JF}} \ni \gamma^*$. $\square$

Acknowledgments I express my special thanks to Kotaro Suzumura for his consistently kind and tolerant support as my supervisor since I was a Ph.D. student under him. I am also thankful to a referee of this paper for his many useful comments. A part of this paper was presented at the International Conference of Rational Choice, Individual Rights, and Non-welfaristic Normative Economics held in Hitotsubashi University, Tokyo, Japan, on March 11th–13th, 2006. Also, another part of this paper was presented at the Workshop on "Real Freedom for All" with Philippe van Parijs held in Ritsumeikan University, Kyoto, Japan, on July 7th, 2006. I am greatly thankful to all of the participants in these meetings, in particular, N. Baigent, K. Basu, M. Fluerbaey, P. Hammond, Y. Sprumont, K. Tadenuma, P. van Parijs, and Y. Xu for their comments. It is worth noting that the main body of this paper is deeply indebted to the joint work with R. Gotoh and K. Suzumura in Gotoh et al. (2005).

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: Wiley

Blackorby, C., Bossert, W., & Donaldson, D. (2005). Multi-profile welfarism: A generalization. *Social Choice and Welfare*, *24*, 253–267

Dworkin, R. (1981a). What is equality? Part 1: Equality of welfare. *Philosophy & Public Affairs*, *10*, 185–246

Dworkin, R. (1981b). What is equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, *10*, 283–345

Dworkin, R. (2000). *Sovereign virtue*. Cambridge, MA: Harvard University Press

Fleurbaey, M., & Trannoy, A. (2003). The impossibility of a paretian egalitarian. *Social Choice and Welfare*, *21*, 243–263

Fleurbaey, M., Tungodden, B. , & Chang, H. (2003). Any non-welfarist method of policy assessment violates the Pareto principle: A comment. *Journal of Political Economy*, *111*, 1382–1385

Foley, D. (1967). Resource allocation and the public sector. *Yale Economic Essays*, *7*, 45–98

Gotoh, R., Suzumura, K., & Yoshihara, N. (2005). Extended social ordering functions for rationalizing fair game forms in the sense of Rawls and Sen. *International Journal of Economic Theory*, *1*, 21–41

Kaplow, L., & Shavell, S. (2001). Any non-welfarist method of policy assessment violates the Pareto principle. *Journal of Political Economy*, *109*, 281–286

Kranich, L. (1994). Equal division, efficiency, and the sovereign supply of labor. *American Economic Review*, *84*, 178–189

Mill, J. S. (1859). On liberty. In M. Warnock (Ed.), *Utilitarianism.* London: Fontana. (republished in 1973)

Nozick, R. (1974). *Anarchy, state and utopia*. Oxford: Basil Blackwell

Parijs, P. van. (1992). Competing justification of basic income. In P. van Parijs (Ed.), *Arguing for basic income* (pp. 1–43). London: Verso

Parijs, P. van. (1993). *Marxism recycled*. Cambridge, UK: Cambridge University Press

Parijs, P. van. (1995). *Real freedom for all: What (if anything) can justify capitalism*. Oxford: Oxford University Press

Pattanaik, P. K., & Suzumura, K. (1996). Individual rights and social evaluation: A conceptual framework. *Oxford Economic Papers*, *48*, 194–212

Pattanaik, P. K., & Xu, Y. (1990). On ranking opportunity sets in terms of freedom of choice. *Recherches Economiques de Louvain*, *56*, 383–390

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press

Roberts, K. (1980). Interpersonal comparability and social choice theory. *Review of Economic Studies*, *47*, 421–439

Sen, A. K. (1970a). *Collective choice and social welfare*. San Francisco: Holden-Day. (republished, North-Holland, Amsterdam, 1979)

Sen, A. K. (1970b). The impossibility of a Paretian liberal. *Journal of Political Economy*, *78*, 152–157

Sen, A. K. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica*, *45*, 1539–1572

Sen, A. K. (1979). Utilitarianism and welfarism. *Journal of Philosophy*, *76*, 463–489

Sen, A. K. (1980). Equality of what? In S. McMurrin (Ed.), *Tanner lectures on human values* (Vol. 1). Cambridge, UK: Cambridge University Press

Sen, A. K. (1985). *Commodities and capabilities*. Amsterdam: North-Holland

Sen, A. K., & Williams, B. (1982). Introduction: Utilitarianism and beyond. In A. K. Sen & B. Williams (Eds.), *Utilitarianism and beyond.* Cambridge, UK: Cambridge University Press

Suzumura, K. (1976). Remarks on the theory of collective choice. *Economica*, *43*, 381–90

Suzumura, K. (2004). *An extension of Arrow's lemma with economic applications* (mimeo). Hitotsubashi University

Tadenuma, K. (2002). Efficiency first or equity first? Two principles and rationality of social choice. *Journal of Economic Theory*, *104*, 462–472

Yamada, A., & Yoshihara, N. (2007). *Triple implementation in production economies with unequal skills by sharing mechanisms*. *International Journal of Game Theory, 36*, 85–106

Yoshihara, N. (2000). *On efficient and procedurally-fair equilibrium allocations in sharing games* (mimeo). IER Discussion Paper 397, Hitotsubashi University

Yoshihara, N. (2005). *A resolution of conflicting claims in constructing social ordering functions* (mimeo). Hitotsubashi University

Yoshihara, N. (2006a). *Normative economic foundation for the theory of welfare state policies* (mimeo). Hitotsubashi University. (Presented at Workshop on "Real Freedom for All" with Van Parijs, Ritsumeikan University, Kyoto, July 7th, 2006)

Yoshihara, N. (2006b). *On non-welfarist social ordering functions* (mimeo). Hitotsubashi University

Part II
Social Choice and Fair Allocations

Monotonicity and Solidarity Axioms in Economics and Game Theory

Yves Sprumont

1 Introduction

An important aspect of the complex notion of fairness in collective choices is that agents should bear responsibility *only* for their own actions. As a corollary, they should be treated "similarly" when a change occurs for which no one is responsible. A minimal condition of "similar" treatment is certainly that nobody benefits from such a change if someone else suffers from it.

The least controversial formulations of this very general *solidarity principle* focus on changes that are unambiguously profitable to society as a whole (in the sense that they always permit a Pareto improvement), or on changes that are clearly detrimental. We refer to the corresponding restricted forms of solidarity as *monotonicity principles.* Thus, in the famous "fair division" problem (discussed in Section 2), a bigger pie for the same group of agents is always better while more agents sharing the same pie is worse: *resource monotonicity* asks that nobody suffer when the pie grows, *population monotonicity* requires that no one benefit from the arrival of new agents.

After the seminal studies of those and related monotonicity properties in the context of bargaining by Kalai (1977), Kalai and Smorodinsky (1975), Thomson (1983a, 1983b), and Thomson and Myerson (1980), among others, the range of models where the properties have been analyzed has expanded tremendously. It now includes transferable and nontransferable utility cooperative games, fair division and exchange economies, problems involving production, public goods, indivisibilities, cost sharing, and more. Thomson (1995, 1999) offers two very complete surveys on population monotonicity and resource monotonicity, respectively.

This essay is meant to be an introduction to these axioms and related solidarity properties; it is not a comprehensive survey. We focus on two particular frameworks, the so-called fair division model and transferable utility cooperative games, but try to provide a relatively detailed treatment of each.

Y. Sprumont
Département de Sciences Économiques and CIREQ, Université de Montréal
e-mail: yves.sprumont@umontreal.ca

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,
Studies in Choice and Welfare.

2 Fair Division

The framework in which the fair division problem is usually formulated is a version of the classical general equilibrium model of exchange economies: the only difference is that property rights are not specified. In contrast with the cooperative games to be studied in the next section, the fair division model is a full-fledged explicit economic model where the information about agents' preferences is carefully separated from the information about physical resources.

There is a finite set L containing $l \geq 2$ goods; a nonempty set $\mathcal{N}$ of nonempty subsets of the positive integers; a nonempty subset $\mathcal{R}$ of the set of all complete; transitive, continuous, monotonic, and convex preference relations on $\mathbb{R}^l_+$; and a subset Ω of $\mathbb{R}^l_{++}$. Monotonicity is understood in the strict sense, and we assume that at least one element of $\mathcal{N}$ contains more than one integer. We refer to $\mathcal{N}$ as the *population domain* and to $\mathcal{R}$ as the *preference domain.* A (*division*) *problem* (or economy) is any list $e = (N, R, \omega)$, where $N \in \mathcal{N}$ is the set of agents, $R \in \mathcal{R}^N$ is their preference profile, and $\omega \in \Omega$ is the collective endowment to be split among them. Agent i's preference $R(i)$ is often written R_i. The indifference and strict preference relations associated with R_i are denoted I_i and P_i, respectively. The set of problems is denoted $\mathcal{E}$; it is fully determined by $\mathcal{N}$, $\mathcal{R}$, and Ω. An allocation (for the problem e) is a vector $x \in \mathbb{R}^{lN}_+$ satisfying the feasibility constraint $\sum_{i \in N} x_i \leq \omega$, where $x_i = (x_i^1, \ldots, x_i^l)$ denotes the bundle allocated to agent i.

An (*allocation*) *rule* F assigns to each problem $e \in \mathcal{E}$ a nonempty set of allocations for e. This set need not be a singleton. Many interesting rules are not single-valued. The Walrasian rule from equal split, for instance, which recommends for each problem e the competitive allocations of the exchange economy obtained from e by splitting the collective endowment equally among the agents, is typically not – unless severe restrictions are imposed on the preference domain $\mathcal{R}$. If we want to be able to discuss such rules, imposing single-valuedness would force us to perform selections from them. This may not be wise because a multivalued rule satisfying a property of interest need not admit any single-valued selection with that property.[1]

2.1 Resource Monotonicity in the Strong Sense

Because we allow for multivalued rules, the general idea of resource monotonicity can be formulated in many different ways. This subsection is devoted to the strongest commonly used requirement.[2]

Resource Monotonicity. If $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are two problems such that $\omega' \geq \omega$, then $x'_i R_i x_i$ for all $x \in F(e)$, all $x' \in F(e')$, and all $i \in N$.

[1] Think, for example, of the upper-hemicontinuity of the Walrasian correspondence on exchange economies with linear preferences.

[2] An even (slightly) stronger property can be defined by insisting that all agents be *strictly* better off when the collective endowment increases.

Resource monotonicity implies the strong restriction that F must be *welfare-wise single-valued*: for all e, all $x, x' \in F(e)$, and all $i \in N$, $x_i I_i x'_i$. This is not true for the weaker condition requiring that if $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are such that $\omega' \geq \omega$, then for all $x \in F(e)$ *there exists* $x' \in F(e')$ such that $x'_i R_i x_i$ for all $i \in N$. This weaker version and the related conditions studied in the next subsection do allow for genuinely multivalued allocation rules.

A simple rule satisfying resource monotonicity consists in splitting the collective endowment equally among all agents: for each $e = (N, R, \omega)$, $F(e) = \{(\frac{\omega}{n}, \ldots, \frac{\omega}{n})\}$. Of course, this may be inefficient. From now on, we insist on (Pareto) efficiency.

Efficiency. For every problem e, every $x \in F(e)$ is efficient.

Resource monotonicity and efficiency are obviously compatible: always allocating the entire collective endowment to, say, agent 1 satisfies both. This dictatorial rule, however, is very unappealing from the distributive viewpoint. Certainly, we would like our rule to meet the standard property of Anonymity. Under welfare-wise single-valuedness, this property reads as follows:[3]

Anonymity. If $N \subset \mathcal{N}$, π is a permutation on the agent set N, and $e = (N, R, \omega), e' = (N, R', \omega)$ are two problems such that $R'_i = R_{\pi(i)}$ for all $i \in N$, then $x'_i I'_i x_{\pi(i)}$ for all $x \in F(e)$, $x' \in F(e')$, and $i \in N$.

This property implies the familiar condition of equivalent treatment of equals: in a given problem, two agents with identical preferences must receive bundles that they judge equally good.

Combining resource monotonicity, efficiency, and anonymity are not much of a challenge if the domain of preferences is small. In the trivial case where $\mathcal{R}$ contains just one preference, for example, equal split of the collective endowment is fully satisfactory. In the same vein, it is important to recall that resource monotonicity is vacuous if no two bundles in Ω are comparable. Throughout the rest of this section, we fix $\Omega = \mathbb{R}^l_+$.

Let us now describe a class of rules satisfying the three axioms introduced above on any preference domain satisfying our basic assumptions. They are based on Pazner and Schmeidler's (1978) concept of egalitarian equivalence. Fix a *numeraire bundle* $\alpha \in \mathbb{R}^l_{++}$. An allocation x for the problem $e = (N, R, \omega)$ is called *egalitarian-equivalent with respect to* α if all agents find their share equivalent to the same multiple of α: there exists a number λ such that $x_i I_i \lambda\alpha$ for all $i \in N$. We call $\lambda\alpha$ the "reference bundle." Pazner and Schmeidler have shown that each problem possesses at least one efficient egalitarian-equivalent allocation with respect to α. The rule F^α that selects all such allocations is welfare-wise single-valued and satisfies the three axioms discussed so far. Resource monotonicity must hold since, when the collective endowment increases, the reference bundle shifts upwards along the ray through α, making everyone better off.

---

[3] The proper formulation of the axiom for an allocation rule that need not be welfare-wise single-valued is as follows. If $N \subset \mathcal{N}$, π is a permutation on the agent set N, and $e = (N, R, \omega), e' = (N, R', \omega)$ are two problems such that $R'_i = R_{\pi(i)}$ for all $i \in N$, then, for all $x \in F(e)$ there exists $x' \in F(e')$ such that $x'_i I'_i x_{\pi(i)}$ for all $i \in N$.

An important weakness of this rule is that it relies on an arbitrary numeraire bundle. Choosing α equal to the collective endowment ω is very tempting but contravenes resource monotonicity unless ad-hoc restrictions are imposed on the preference domain. To secure resource monotonicity, it is crucial that α be independent of ω.

A common interpretation of the rule F^{α} is that it equalizes (at the highest possible level) the numerical representations of the agents' preferences obtained by calibration along the ray through the origin and α. Specifically, defining u_i^{α} by the condition that $u_i^{\alpha}(x_i)\alpha I_i x_i$, the rule F^{α} selects the efficient allocations x at which $u_1^{\alpha}(x_1) = \cdots = u_n^{\alpha}(x_n)$. In fact, any rule defined by first choosing a numerical representation of each agent's preferences and then selecting the efficient allocations equalizing these representations does satisfy resource monotonicity (provided that it is well defined: we must make sure that the numerical representations can be equalized for every problem).

Examples include Thomson's (1994) *equal opportunity equivalent rules*, under which the numerical representations of preferences are calibrated along nested sequences of subsets of $\mathbb{R}_+^l$.

We briefly describe two further examples. Both work (at least) when all indifference sets of every preference R_i in the domain $\mathcal{R}$ cut the axes: for every $x_i \in \mathbb{R}_+^l$ and every $h \in L$, there is some (necessarily unique) number $a^h(x_i, R_i)$ such that $x_i I_i(0,\ldots,0,\ a^h(x_i,R_i),0,\ldots,0)$. Our first example uses the numerical representation u_i^a of R_i given by

$$u_i^a(x_i) = \Pi_{h\in L} a^h(x_i, R_i).$$

It is not difficult to check that every problem admits at least one efficient allocation x at which $u_1^a(x_1) = \cdots = u_n^a(x_n)$. Selecting such allocations defines an anonymous, efficient, resource-monotonic rule. Our second example is based on a numerical representation of preferences due to Neuefeind (1972). Given any preference $R_i \in \mathcal{R}$ and $x_i \in \mathbb{R}_+^l$, simply define $u_i^L(x_i)$ to be the Lebesgue measure of the lower contour set of x_i. Again, selecting in each problem the efficient allocations equalizing these representations yields a rule satisfying the three axioms defined so far.

More work is needed to better understand the class of rules meeting these axioms. Notice that under all the rules described so far, *any* change in the collective endowment affects all agents in a common direction: nobody loses or nobody benefits.

Resource Solidarity. If $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are two arbitrary problems, then (i) $x_i' R_i x_i$ for all $x \in F(e)$, all $x' \in F(e')$, and all $i \in N$, or (ii) $x_i R_i x_i'$ for all $x \in F(e)$, all $x' \in F(e')$, and all $i \in N$.

This powerful property is fairly well understood. We know from the work of Keiding and Moulin (1991) and Sprumont (1996) that it is essentially equivalent to equalize some numerical representation of the agents' preferences. Resource solidarity, however, is by no means an implication of resource monotonicity, efficiency, and anonymity. To see this, note that the Walrasian rule from equal split satisfies resource monotonicity, but not resource solidarity, on the domain of Cobb–Douglas preferences. For instance, if the two-good collective endowment

$\omega = (\omega^1, \omega^2)$ is to be divided between two agents whose preferences admit the numerical representations

$$u_1(x^1, x^2) = \frac{2}{3}\log x^1 + \frac{1}{3}\log x^2,$$
$$u_2(x^1, x^2) = \frac{1}{3}\log x^1 + \frac{2}{3}\log x^2,$$

the Walrasian rule from equal split gives $(\frac{2}{3}\omega^1, \frac{1}{3}\omega^2)$ to agent 1 and the rest to agent 2. Both agents gain when ω increases but if $\omega = (3,6)$ changes to $\omega' = (6,3)$, agent 1 benefits while 2 suffers. For any domain that includes the Cobb–Douglas preferences, applying the Walrasian rule from equal split on the Cobb–Douglas subdomain and, say, the egalitarian-equivalent rule F^α otherwise defines a hybrid rule that satisfies resource monotonicity but violates resource solidarity. Of course, this rule is discontinuous in the preferences under any reasonable definition of continuity.

Splitting the collective endowment equally is an example of a trivially continuous, anonymous, resource-monotonic procedure that fails solidarity, but it is not efficient. It is not clear whether resource monotonicity, efficiency, anonymity, and a reasonable preference continuity condition together imply resource solidarity when the domain $\mathcal{E}$ is large enough.

We now investigate to what extent resource monotonicity is compatible with other important equity criteria. Consider the following three conditions.

Conditional Equal Split. For every problem $e = (N, R, \omega)$ for which $(\frac{\omega}{n}, \ldots, \frac{\omega}{n})$ is efficient, $(\frac{\omega}{n}, \ldots, \frac{\omega}{n}) \in F(e)$.

No Exploitation. For every problem $e = (N, R, \omega)$, every $x \in F(e)$, and every agent $i \in N$, there exists a good $h \in L$ such that $x_i^h \geq \frac{\omega^h}{n}$.

No Domination. For every problem $e = (N, R, \omega)$, every $x \in F(e)$, and all $i, j \in N$, there exists a good $h \in L$ such that $x_i^h \geq x_j^h$.

Conditional equal split requires that equal division of the collective endowment be among the recommended allocations whenever it is efficient. If the crux of the fair division problem is really to reconcile equality with efficiency, this axiom seems rather natural.

The other two conditions are weak forms of general principles that admit other, more demanding, formulations. No exploitation offers to each agent a *guarantee* that is independent of the other agents' preferences: in no circumstances will he get strictly less than the mean endowment of every good. This axiom, introduced by Gaspart (1996) in a context with production, is weaker than the more standard equal split lower bound to be discussed later.

No Domination, finally, is an attempt at rooting fairness in *bilateral comparisons* of bundles. The general idea is that no one should be treated worse than any one else. The very weak requirement expressed here is that no agent receive strictly less of every commodity than any other agent. This minimal condition, due to Moulin and

Thomson (1988), is weaker than the celebrated no envy condition discussed later. No Domination and no exploitation are equivalent in the two-agent case (provided that no part of the collective endowment is destroyed) but are independent as soon as there are at least three agents.

Easy examples show that all the efficient resource-monotonic rules described so far violate conditional equal split, no exploitation, and no domination unless the preference domain is extremely small. As it turns out, this is unavoidable.

Theorem 1. (Maniquet & Sprumont, 2000; Moulin & Thomson, 1988) *If the preference domain $\mathcal{R}$ contains the homothetic preferences, resource monotonicity and efficiency together are incompatible with any of the following principles: conditional equal split, no exploitation, no domination.*

Proof. Suppose $L = N = \{1,2\}$, that is, there are two goods and two agents. Assume that agent 1's preference relation R_1 is strictly convex, homothetic, differentiable, and its marginal rate of substitution is -8 along the ray $x^2 = 3x^1$, -1 along the ray $x^2 = 2x^1$, and $-\frac{1}{8}$ along the ray $x^2 = \frac{1}{3}x^1$. Furthermore, $(2,4)P_1(9,3)$. It is easy to see that the latter requirement is compatible with the data on the marginal rates of substitution. Agent 2's preference relation is symmetric to 1's with respect to the diagonal: $(x^1,x^2)R_2(y^1,y^2)$ if and only if $(x^2,x^1)R_1(y^2,y^1)$.

In the problem e in which $\omega = (6,6)$, giving $(2,4)$ to agent 1 and $(4,2)$ to agent 2 is efficient. Choose an arbitrary allocation $x \in F(e)$. By efficiency, $x_1R_1(2,4)$ or $x_2R_2(4,2)$. If the former statement is true, consider the economy e' obtained by increasing the collective endowment to $\omega' = (18,6)$ and let $x' \in F(e')$. Resource monotonicity demands that $x'_1R_1x_1$, hence $x'_1P_1(9,3) = \frac{\omega'}{2}$. But the efficient allocations in e' coincide with the diagonal of its edgeworth box. Therefore, $x'_1 \gg \frac{\omega'}{2}$ and, consequently, $x'_2 \ll \frac{\omega'}{2}$. This contradicts conditional equal split, no exploitation, and no domination. A similar contradiction obtains in the case where $x_2R_2(4,2)$. $\square$

The incompatibility of resource monotonicity and efficiency with no domination was shown by Moulin and Thomson (1988). Their proof, just like the one given above, consists of a two-agent example and therefore establishes the conflict with no exploitation as well. The incompatibility with conditional equal split is noted in Maniquet and Sprumont (2000). An easy corollary to Theorem 1 is that the Walrasian rule from equal split is not resource-monotonic under the domain restriction we have considered.

2.2 Weakening Resource Monotonicity

In view of the incompatibilities stated in Theorem 1, it is of interest to try and weaken the resource monotonicity axiom. Rather than ask that no one suffer when the collective endowment grows from ω to ω', we could merely demand that nobody's consumption bundle shrink. The strongest requirement in that vein imposes that relation between all allocations chosen at ω and ω'.

More Is Not Less (1). If $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are two problems such that $\omega' \geq \omega$, then, for all $i \in N$, all $x \in F(e)$, and all $x' \in F(e')$, there exists $h \in L$ such that $x_i'^h \geq x_i^h$.

The original version of this axiom is by Geanakoplos and Nalebuff (1988). The name "More Is Not Less" is borrowed from Moulin (1991). The precise conditions used by Geanakoplos and Nalebuff (1988) and Moulin (1991), however, are both weak forms of the condition just stated: we will come to them later.

More Is Not Less (1) does not imply welfare-wise single-valuedness. An interesting rule meeting More Is Not Less (1), efficiency, and anonymity is the *canonical budget-constrained efficient rule* F^{B} adapted from the work of Balasko (1979) by Moulin (1991). In a problem $e = (N, R, \omega), F^{\mathrm{B}}(e)$ contains all the efficient allocations x such that

$$\sum_{h \in L} \frac{x_i^h}{\omega^h} = \frac{l}{n} \text{ for all } i \in N.$$

Thus, the value of all individual bundles is the same at the "canonical" prices $p = (\frac{1}{\omega^1}, \ldots, \frac{1}{\omega^l})$. The set $F^{\mathrm{B}}(e)$ is always nonempty (as shown in Balasko), but may be quite large.

Unfortunately, the rule F^{B} violates the following two classic axioms.

Equal Split Lower Bound. For every problem $e = (N, R, \omega)$, every $x \in F(e)$, and every $i \in N$, $x_i R_i \frac{\omega}{n}$.

No Envy. For every problem $e = (N, R, \omega)$, every $x \in F(e)$, and all $i, j \in N, x_i R_i x_j$.

The equal split lower bound strengthens the no exploitation test. It appears to be the oldest axiom of the fair division literature. Introduced by Steinhaus (1948) and discussed by Dubins and Spanier (1961), it was originally viewed as guaranteeing a fair share of the collective endowment to each agent. More recently, Moulin (1991) showed that it provides the highest symmetric feasible welfare lower bound under the convexity assumption on preferences. This need not be true for nonconvex preferences; indeed, equal split may then be inefficient among identical agents. Moulin and Thomson (1988) defined a weak version of the axiom, the *ε-Lower Bound*, requiring that $x_i R_i \varepsilon \omega$ for all $i \in N, x \in F(e), e \in \mathcal{E}$, where $0 < \varepsilon \leq 1/n$.

The well known no envy property strengthens no domination by demanding that no one prefer anyone else's bundle to his own. The condition is due to Foley (1967). Varian (1976) and Champsaur and Laroque (1981) showed that, in conjunction with efficiency, it characterizes the Walrasian allocations from equal split in smooth connected continuum economies.

Just like conditional equal split, no exploitation, and no domination, the equal split lower bound and its ε-variants, as well as no envy, are restrictions on *allocations* in any given problem. They do not involve comparisons across problems.

The fact that the rule F^{B} violates the equal split lower bound and no envy is only an illustration of more fundamental incompatibilities. It turns out that these two axioms, in conjunction with efficiency, conflict with the More Is Not Less principle. To express these conflicts as sharply as possible, we state two weak versions of the latter principle. Both conditions have been used in the literature. They are due, respectively, to Moulin (1991) and Geanakoplos and Nalebuff (1988).

More Is Not Less (2). If $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are two problems such that $\omega' \geq \omega$, then, for each agent $i \in N$ and every allocation $x \in F(e)$, there exists an allocation $x' \in F(e')$ and a good $h \in L$ such that $x_i'^h \geq x_i^h$.

More Is Not Less (3). If $e = (N, R, \omega)$ and $e' = (N, R, \omega')$ are two problems such that $\omega' \gg \omega$, then, for each agent $i \in N$, there exist an allocation $x \in F(e)$, an allocation $x' \in F(e')$, and a good $h \in L$ such that $x_i'^h \geq x_i^h$.

More Is Not Less (3), called "Bigger is Better" in Geanakoplos and Nalebuff (1988), is the weakest possible formulation of the monotonicity principle.

Theorem 2. (Moulin, 1991; Moulin & Thomson, 1988) *If the preference domain $\mathcal{R}$ contains the homothetic preferences, More Is Not Less (2), Efficiency, and the Equal Split Lower Bound are incompatible.*

The original version of Theorem 2 proved in Moulin and Thomson (1988) used resource monotonicity rather than the weaker condition More Is Not Less (2). The version stated here is by Moulin (1991) who actually replaces the equal split lower bound with any ε-lower bound. Moulin's theorem is stated for rules satisfying anonymity, scale covariance, and neutrality, but these conditions play no role in his proof, as he himself notes. Observe, on the other hand, that More Is Not Less (3) cannot replace More Is Not Less (2) in the statement of Theorem 2, as the rule choosing all efficient allocations satisfying the equal split lower bound shows.

Again, a straightforward corollary to Theorem 2 is that the Walrasian rule from equal split violates More Is Not Less (2) under our domain assumption.

We conclude this section by stating one last important incompatibility result.

Theorem 3. (Geanakoplos & Nalebuff, 1988) *If the preference domain $\mathcal{R}$ contains the homothetic preferences and the number of agents, n, is sufficiently large, More Is Not Less (3), Efficiency, and No Envy are incompatible.*

Geanakoplos and Nalebuff (1988) actually allow *weakly* monotonic preferences (which we rule out), but their argument can be adapted. As they point out, the three axioms in Theorem 3 *are* compatible in the two-agent case: again, selecting all efficient allocations satisfying the equal split lower bound will do. This is because the latter axiom happens to imply no envy in that case (under our convexity assumption on preferences), much like no exploitation implies no domination. The proof of Theorem 3 is markedly different from those of Theorems 1 and 2. It works by first establishing the incompatibility for continuum economies (for which no envy and efficiency essentially characterize Walrasian allocations from equal split) and then using an approximation argument. How large the number of agents must be to derive the result is not known.

A final remark is in order before closing this subsection. The incompatibilities recorded in Theorems 1, 2, and 3 all require preferences with a high degree of complementariness between goods. If all goods are gross substitutes and normal for all agents, the Walrasian rule satisfies Resource Monotonicity. This follows from Polterovich and Spivak (1983), as noted by Moulin and Thomson (1988). Maximal domains over which efficiency and resource monotonicity are compatible with the other fairness conditions discussed above have not been identified.

2.3 Population Monotonicity and Related Properties

Population monotonicity is a dual relative of resource monotonicity. It requires that when the set of agents sharing the same collective endowment expands, none of the agents initially present benefit. The strong form of the axiom is as follows.

Population Monotonicity. If $e = (N, R, \omega)$ and $e' = (N', R', \omega)$ are two problems such that $N \subset N'$ and $R_i = R'_i$ for all $i \in N$, then $x_i R_i x'_i$ for all $x \in F(e)$, all $x' \in F(e')$, and all $i \in N$.

In much the same way as resource monotonicity, population monotonicity forces welfare-wise single-valuedness. As with resource monotonicity, some care must be taken in specifying the set of problems $\mathcal{E}$. Population monotonicity has little bite on a small population domain $\mathcal{N}$.

The two monotonicity axioms are related. In particular, it is easy to see that every Pareto-indifferent, consistent, resource-monotonic rule is population-monotonic. A rule F is said to be Pareto-indifferent if it does not distinguish among welfare-wise equivalent allocations: for every problem e and all allocations x, x' such that $x_i I_i x'_i$ for all $i \in N$, we have $x \in F(e)$ if and only if $x' \in F(e)$. The rule is consistent if the individual bundles assigned to any subset of agents form a fair allocation of the sum of their bundles: for every problem (N, R, ω), $x \in F(N, R, \omega)$, and $S \subset N$, we have $x_S \in F(N, R_S, \sum_{i \in S} x_i)$, where x_S and R_S denote the restrictions of x and R to the subgroup S.

Returning to a procedure suggested in Sect. 2.1, observe that choosing (any nonempty subset of) efficient allocations that are egalitarian-equivalent with respect to the collective endowment defines a rule meeting population monotonicity and efficiency. Such a procedure was already suggested in Pazner and Schmeidler (1978). Formally, we say that F is a *Pazner–Schmeidler rule* if for each problem $e = (N, R, \omega)$, every allocation $x \in F(e)$ is both efficient and egalitarian-equivalent with respect to the collective endowment ω : there exists λ such that $x_i I_i \lambda \omega$ for all $i \in N$. There is only one such rule on any domain of *strictly* convex preferences but there may be several if strict convexity is not assumed. The largest one simply selects all efficient ω-egalitarian-equivalent allocations. All Pazner–Schmeidler rules are welfare-wise single-valued and equivalent.

These rules equalize (at the highest possible level) the numerical representations of the preferences calibrated along the ray through the collective endowment. Because these representations are independent not only of the population N but also of the preference profile R, the Pazner–Schmeidler rules satisfy the following condition.

Extended Population Solidarity. If $e = (N, R, \omega)$ and $e' = (N', R', \omega)$ are two problems such that $R_i = R'_i$ for all $i \in N \cap N'$, then (i) $x_i R_i x'_i$ for all $i \in N \cap N'$ or (ii) $x'_i R_i x_i$ for all $i \in N \cap N'$.

To understand this property, write $M = N \cap N'$. Imagine that the agents in $N \backslash M$ leave N while those in $N' \backslash M$ join in. Since the members of M bear no responsibility for these changes (which are initiated by outsiders of their group), it would not be justified to discriminate among them by making some better off and others worse off.

On top of the extended population solidarity condition, the Pazner–Schmeidler rules clearly satisfy the Equal Split Lower Bound. This is of interest, especially in contrast with the negative result recorded in Theorem 2. It turns out that under a rather mild auxiliary condition dubbed replication invariance, extended population solidarity and the equal split lower bound characterize the Pazner–Schmeidler rules. If k is a positive integer, a *k-replica* of the problem $e = (N, R, \omega)$ is any problem e' such that (i) $\omega' = k\omega$ and (ii) $\mid R'^{-1}(R^0) \mid = k \mid R^{-1}(R^0) \mid$ for every $R^0 \in \mathcal{R}$. Recall that R maps the agent set N into the preference domain $\mathcal{R}$: $R^{-1}(R^0)$ is thus the set of agents in N having preference R^0. Condition (ii) merely says that N' contains exactly k times as many agents having any given preference as N; in particular, there are kn agents in N'. A *replica* of e is any problem e' that is a k-replica of e for some k. Replication invariance requires that the allocations recommended for e' replicate those chosen for e. For welfare-wise single-valued rules, it may be stated as follows.[4]

Replication Invariance. For every problem e, every replica e' of e, every $x \in F(e)$ and $x' \in F(e')$, and every $R^0 \in \mathcal{R}$, $x_i I^0 x'_j$ for all $i \in R^{-1}(R^0)$ and $j \in R'^{-1}(R^0)$.

Theorem 4. (Sprumont & Zhou, 1999) *If the population domain $\mathcal{N}$ contains all finite agent sets, a rule satisfies efficiency, extended population solidarity, the equal split lower bound, and replication invariance if and only if it is a Pazner–Schmeidler rule.*

Proof. The "if" part being straightforward, we prove only its converse. Let F satisfy the three axioms in Theorem 4. As noted earlier, F must be welfare-wise single-valued. Observe that replication invariance implies the following strong version of the anonymity principle: if $e = (N, R, \omega), e' = (N', R', \omega)$ are two problems and π is a bijection from N to N' such that $R'_{\pi(i)} = R_i$ for each $i \in N$, then $x'_{\pi(i)} I_i x_i$ for all $x \in F(e), x' \in F(e')$, and $i \in N$. This is merely because e' is a 1-replica of e.

Suppose now, contrary to the claim, that F is not a Pazner–Schmeidler rule. By efficiency, there must exist a problem $e = (N, R, \omega)$, two preferences $R^1, R^2 \in \mathcal{R}$, and nonnegative numbers λ^1, λ^2 such that

$$x_i I^1 \lambda^1 \omega \gg \lambda^2 \omega I^2 x_j \tag{1}$$

[4] One possible formulation of the axiom for an allocation rule that need not be welfare-wise single-valued is as follows. For any economies e, e' such that e' is a replica of e or e is a replica of e', for all $x \in F(e)$, there exists $x' \in F(e')$ such that, for every $R^0 \in \mathcal{R}$, $x_i I^0 x'_j$ for all $i \in R^{-1}(R^0)$ and $j \in R'^{-1}(R^0)$.

for all $x \in F(e)$, $i \in R^{-1}(R^1)$, $j \in R^{-1}(R^2)$. Let a, b be two positive integers such that

$$\lambda^1 > \frac{a}{b} > \lambda^2. \tag{2}$$

For any positive integer t large enough, construct a problem $e_t = (N_t, R_t, \omega_t)$ such that $\omega_t = ta\omega$, $|N_t| = tb$, and $|R_t^{-1}(R^0)| = |R^{-1}(R^0)|$ for every $R^0 \neq R^1$. As t grows, the number of agents with preference R^1 increases while the numbers of agents with other preferences remain the same as in e. By the feasibility constraint, the average bundle received by the R^1 agents in e_t cannot exceed $ta\omega / |R_t^{-1}(R^1)|$. Since this upper bound goes to $a\omega/b$ as t tends to infinity and since equal sharing is efficient among the R^1 agents because of the preference convexity assumption, we know that for every $\lambda > a/b$ and all t large enough, $\lambda\omega R^1 x_{ti}$ for all $x_t \in F(e_t)$ and $i \in R_t^{-1}(R^1)$. Using (2.1) and (2.2), we conclude that if t is large enough,

$$x_i P^1 x_{tj} \tag{3}$$

for all $x \in F(e), x_t \in F(e_t), i \in R^{-1}(R^1), j \in R_t^{-1}(R^1)$.

Now, consider a ta-replica e' of e. By replication invariance and (2.3), $x'_i P^1 x_{tj}$ for all $x' \in F(e'), x_t \in F(e_t), i \in R'^{-1}(R^1)$, and $j \in R_t^{-1}(R^1)$. But since the collective endowment in e' is the same as in e_t, namely $ta\omega$, we may apply extended population solidarity. After one more "reverse" application of replication invariance, this yields that $x_i P^2 x_{tj}$ for all $x \in F(e)$, $x_t \in F(e_t)$, $i \in R^{-1}(R^2)$, and $j \in R_t^{-1}(R^2)$. Recalling (2.1) and (2.2), we conclude that $\frac{a\omega}{b} P^2 x_{tj}$ for all $x_t \in F(e_t)$ and $j \in R_t^{-1}(R^2)$, violating the equal split lower bound. $\square$

Theorem 4 is a variation on a result of Sprumont and Zhou (1999). Their main result dispenses with replication invariance at the cost of restricting attention to anonymous continuum economies. The authors mention that adding replication invariance yields a characterization in the large anonymous finite case; Theorem 4 shows that anonymity need not be imposed.[5]

Although the Pazner–Schmeidler rules are well defined without convexity assumption on preferences, Theorem 4 collapses in the nonconvex case. This is because the equal split lower bound loses much of its bite in that case, as already noted in Sect. 2.3. While the assumption that all preferences in $\mathcal{R}$ are convex is crucial, we stress that no richness condition is needed: indeed, Theorem 4 is true even if $\mathcal{R}$ contains only one preference.

2.4 Incompatibilities with Population Monotonicity

Returning to the fairness axioms defined earlier, we note that the Pazner–Schmeidler rules satisfy conditional equal split. Since the equal split lower bound implies no envy in the two-agent case, these rules are also envy-free in that case. With more

[5] Another minor improvement is that we do not require *strict* convexity.

agents, unfortunately, simple examples show that the Pazner–Schmeidler rules not only generate envy but also violate no domination. It turns out that this is a general difficulty.

Theorem 5. (Kim, 2004) *Let the preference domain $\mathcal{R}$ contain the homothetic preferences.*

1. If the population domain $\mathcal{N}$ contains all agent sets of cardinalities $n \leq 8$, population monotonicity, efficiency, and no envy are incompatible.

2. If the population domain contains all agent sets of cardinalities $n \leq 52$, population monotonicity, efficiency, and no domination are incompatible.

Proof. We only prove the incompatibility of population monotonicity, efficiency, and no envy. We use a slightly simpler argument than Kim's, which, however, requires agent sets of larger cardinalities (e.g., a population domain containing all sets of cardinalities $n \leq 27$ will do).

There are two goods, whose consumption levels are denoted x and y, and four types of preferences represented by the utility functions

$$\begin{aligned} u_1(x,y) &= ax+y, \\ u_2(x,y) &= x+ay, \\ u_{2'}(x,y) &= x+a'y, \\ u_3(x,y) &= \min\{bx+y, x+by\}, \end{aligned}$$

where $1 < a' < a < b$. The collective endowment is $\omega = (1,1)$.

Consider a society N composed of one agent of type 1 (indexed 1), p agents of type 2 (indexed $2i$, $i = 1,\ldots,p$), one agent of type $2'$ (indexed $2'$), and q agents of type 3 (indexed $3i$, $i = 1,\ldots,q$), where $q < p$ and

$$a' < p < 2a'. \tag{4}$$

Let z be an efficient envy-free allocation for N. Clearly, $z_{3i} = \alpha\omega$ for all $i = 1,\ldots,q$ and some positive number α. We claim that

$$z_{2i} = (0, y_2/p),\ i = 1,\ldots,p,$$

where y_2 is some positive number. To see this, suppose $x_{2i} > 0$ for some i. By efficiency, $y_1 = y_{2'} = 0$, so $\sum y_{2i} = 1 - q\alpha$. To make sure that $2'$ does not envy $2i$, $x_{2'} \geq a' \max y_{2i} \geq a'(1-q\alpha)/p$, whereas preventing 1 from envying $2'$ requires $x_1 \geq x_{2'}$. Combining these two inequalities, $x_1 + x_{2'} \geq 2a'(1-q\alpha)/p > 1 - q\alpha$, contradicting the feasibility constraint $z_1 + z_{2'} + \sum z_{2i} \leq (1-q\alpha)\omega$. Therefore, $x_{2i} = 0$ for all i and no envy among the type 2 agents forces an equal consumption of the other good; it must be positive to prevent them from envying type 3 agents.

Next we claim that

$$z_1 = (x_1, 0),\ \text{where } x_1 > (p-q)/2p.$$

Suppose first that $y_1 > 0$. Efficiency requires $x_{2'} = x_{2i} = 0$ for all i, hence $x_1 = 1 - q\alpha$. By no envy and feasibility, $y_{2'} = y_{2i} < (1-q\alpha)/(p+1)$ for all i, whereas preventing $2'$ from envying 1 requires $a'y_{2'} > 1 - q\alpha$ since $y_1 > 0$. Therefore, $p < a' - 1$, contradicting (4). This shows that $z_1 = (x_1, 0)$. To make sure that 1 does not envy $2'$, we must have $x_1 \geq x_{2'}$, while guaranteeing that $2i$ does not envy $3j$ requires $y_2/p \geq \alpha = (1 - x_1 - x_{2'})/q$. Since $y_2 < 1$, the stated lower bound on x_1 follows.

To finish the proof, let now agent $2'$ and agents $2i$, $i = 1, \ldots, p$ leave society N. The new society, N', thus contains agents 1 and $3i$, $i = 1, \ldots, q$. The only efficient envy-free allocation for N' is equal split. But $u_1(\omega/(q+1)) < u_1(z_1)$, contradicting population monotonicity, if

$$2(a+1)p < a(q+1)(p-q).$$

This inequality is compatible with (4): choose for instance $a' = 9$, $a = q = 10$, and $p = 15$. □

A forerunner of Theorem 5 is in Moulin (1990b), who suggests the incompatibility of population monotonicity, efficiency, and no envy for large enough economies. Kim (2004) provides a precise bound on the number of agents and a detailed proof; he also extends the impossibility result to the no domination principle and establishes incompatibilities involving ε-variants of no domination requiring that no agent receive less than an ε-fraction of anybody else's bundle.

In view of these negative results, it seems natural to weaken the population monotonicity axiom. Instead of requiring that no one benefit from the arrival of new agents, one could ask that nobody's bundle expand in all coordinates. This would be the analogue of the More Is Not Less weakening of resource monotonicity; several versions are possible. This type of requirement has not been systematically studied yet, except to note that the Walrasian rule from equal split violates it: see, for example, Moulin (1995), building on Chichilnisky and Thomson (1987).

The other alternative is again to restrict the preference domain. Fleurbaey (1996) points out that the Walrasian rule from equal split satisfies Population Monotonicity under the gross substitutability and normality restrictions. Maximal domains on which the incompatibilities of Theorem 5 can be avoided are not known.

3 Cooperative Games

We now turn to the model of cooperative games. The data in a cooperative game specify the utility possibility sets of the various subsets of agents. If the underlying preferences are quasi-linear, these sets are just half-spaces of $\mathbb{R}^N$. This is the so-called transferable utility case; we will not consider the more general model of nontransferable utility. Comparing this framework with the fair division problem, we note two key differences. On the one hand, the information about agents' preferences is no longer separated from the information about physical resources; thereby making the environment "poorer." On the other hand, a "fallback" position is now specified for each group of agents, which was absent in the fair division problem (because private endowments were not specified). This complicates the problem and changes its nature; indeed, a key issue is to reward the agents in a way that takes into account what each coalition can guarantee to itself.

The formal model is as follows. Given are a nonempty set $\mathcal{N}$ of nonempty finite subsets of the positive integers and a mapping $\mathcal{V}$ assigning to each $N \in \mathcal{N}$ a nonempty subset $\mathcal{V}(N)$ of $\mathbb{R}^{P(N)}$, where $P(N)$ denotes the set of all nonempty subsets of N. We refer to $\mathcal{N}$ as the *population domain* and to $\mathcal{V}$ as the *domain of worth functions.* A *game* is a pair (N, v), where $N \in \mathcal{N}$ and $v \in \mathcal{V}(N)$. Let $\mathcal{G}$ denote the set of games. The finite set N contains the agents and the worth function v assigns to each nonempty subset S of N, called a coalition, a real number $v(S)$ called its worth. To alleviate notation, we often write i instead of $\{i\}$, ij instead of $\{i, j\}$, and so on.

A payoff vector for the game (N, v) is a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^N$ such that $\sum_{i \in N} x_i = v(N)$:$x_i$ is agent i's share of the worth of N. An (*allocation*) *rule* F assigns to each game (N, v) a nonempty set $F(N, v)$ of payoff vectors for that game. When F is single-valued, we denote by $f(N, v) = (f_1(N, v), ..., f_n(N, v))$ the unique payoff vector belonging to $F(N, v)$.

3.1 Resource Monotonicity

We start with the resource monotonicity principle. We fix N and drop it from our notation: thus, $\mathcal{V}$ stands for $\mathcal{V}(N)$, the set of worth functions on our fixed set of coalitions, and we call any $v \in \mathcal{V}$ a game. A variety of different interesting resource monotonicity conditions can be defined. The closest counterpart to the condition studied in the fair division problem is probably the following axiom: if v and v' are two games such that $v \leq v'$, then $x_i \leq x'_i$ for all $x \in F(v)$, $x' \in F(v')$, and $i \in N$. Splitting $v(N)$ equally meets this axiom; contrary to the fair division problem, efficiency and equality do not conflict. But this egalitarian rule is not compelling here, precisely because it does not take account of the worths of the various coalitions. To a lesser extent, the monotonicity axiom itself lacks appeal for the same reason. From v to v', the relative strength of some agents may deteriorate so much that it becomes unreasonable to require that they be better off.

The central condition in the literature requires that when the worth of a coalition increases, *ceteris paribus,* none of its members get a smaller payoff.

Coalitional Resource Monotonicity (1). If v, v' are two games and S is a coalition such that $v(S) \leq v'(S)$ and $v(T) = v'(T)$ for every coalition $T \neq S$, then $x_i \leq x'_i$ for all $x \in F(v)$, all $x' \in F(v')$, and all $i \in S$.

An interesting equivalent expression of it is as follows.

Individual Resource Monotonicity. If $i \in N$ and v, v' are two games such that $v(S) \leq v'(S)$ for all coalitions S containing i and $v(T) = v'(T)$ for all other coalitions T, then $x_i \leq x'_i$ for all $x \in F(v)$ and all $x' \in F(v')$.

It is obvious that the latter axiom implies coalitional resource monotonicity (1); to see that the converse is also true, simply apply the coalitional version repeatedly. While the coalitional formulation emphasizes the solidarity content of the axiom, the equivalent individual formulation is readily interpreted in incentive terms: see Shubik (1962).

The axiom is clearly satisfied by the *Shapley value* (Shapley, 1953), which selects in every game v the unique payoff vector given by

$$x_i = \sum_{S:i\in S\subset N} \frac{(s-1)!(n-s)!}{n!}(v(S)-v(S\backslash i)),$$

for all $i \in N$: under the premises of individual resource monotonicity, none of the terms in that sum decreases when v is replaced with v'.

Just like resource monotonicity in the fair division problem, coalitional resource monotonicity (1) implies single-valuedness (by choosing $v = v'$). Since this is perhaps not desirable, the following weaker requirement is of interest.

Coalitional Resource Monotonicity (2). If v, v' are two games, S is a coalition such that $v(S) \leq v'(S)$ and $v(T) = v'(T)$ for every coalition $T \neq S$, and $i \in S$, then there exist $x \in F(v)$ and $x' \in F(v')$ such that $x_i \leq x'_i$.

As in Sect. 2, we now ask to what extent the conditions are compatible with other properties of allocation rules. One of the most fundamental stability requirements on a payoff vector x are the well known core constraints, requiring that no coalition be able to improve on x on its own: $\sum_{i\in S} x_i \geq v(S)$ for all coalitions S. The set of payoff vectors satisfying these inequalities is the *core* of the game v. While the core may be empty, it makes sense to ask that a rule chooses core payoffs whenever possible.

Core Principle. If the core of v is nonempty, then $F(v)$ is a subset of it.

As is well known, the Shapley value violates this principle. Unfortunately, this is again an expression of a general difficulty.

Theorem 6. (Aumann, 1985; Young, 1980)

1. If N contains at least four agents and $\mathcal{V}$ includes the games with a nonempty core, coalitional resource monotonicity (1) and the core principle are incompatible.

2. If N contains at least five agents and $\mathcal{V}$ includes the games with a nonempty core, coalitional resource monotonicity (2) and the core principle are incompatible.

Proof. We will only prove statement (1). To do so, we use the "glove game" popularized by Aumann (1985). There are four agents and two completely complementary goods, left and right gloves. Suppose 1 and 2 are endowed with one left glove each while 3 and 4 own one right glove each. If a pair of gloves has value 1, this generates the game v,

$$v(i) = v(12) = v(34) = 0, v(13) = v(14) = v(23) = v(24) = v(ijk) = 1, v(N) = 2,$$

where i, j, k are any three distinct agents. A payoff vector x is in the core of this game if and only if $x = (\lambda, \lambda, 1-\lambda, 1-\lambda)$ for some $\lambda \in [0,1]$. By the core principle, f selects such a payoff.

If $\lambda > 0$, give one more left glove to agent 1. This generates the game v',

$$v'(134) = 2, v'(S) = v(S) \text{ for any } S \neq \{1,3,4\}.$$

The core of this game consists of the unique payoff vector $(0,0,1,1)$, contradicting resource monotonicity.

If $\lambda < 1$, give one more right glove to agent 3, thereby generating the game v'',

$$v''(123) = 2, v''(S) = v(S) \text{ for any } S \neq \{1,2,3\},$$

whose unique core payoff vector is $(1,1,0,0)$, contradicting resource monotonicity again. □

Theorem 6 is in the same spirit as Theorem 2 on fair division. The original version of Theorem 6 proved by Young (1980) was formulated for single-valued rules and required five agents. Young's proof uses only two five-agent games whose cores are singletons and does *not*, in effect, rely on single-valuedness: it is a proof of statement (2) above. The proof of statement (1) given here was provided to me by Aumann; it builds on the discussion of the glove game in Aumann (1985). Observe that it does not use the full force of coalitional resource monotonicity (1) but only the requirement that, under the same premises and for all $x \in S$ and $x \in F(v)$, *there exists* $x' \in F(v')$ such that $x_i \leq x'_i$.

In view of the negative results in Theorem 6, it is natural to investigate weaker resource monotonicity conditions. One weak form, due to Meggido (1974), requires that no agent suffer from an increase in the worth of the "grand" coalition.

Aggregate Resource Monotonicity. If v, v' are two games such that $v(N) \leq v'(N)$ and $v(T) = v'(T)$ for every coalition $T \neq N$, then $x_i \leq x'_i$ for all $x \in F(v)$, all $x' \in F(v')$, and all $i \in N$.

This is a fairly weak axiom. It is compatible with the core principle: Young (1980) showed that the per capita nucleolus satisfies both conditions.

Another possible approach consists in restricting the domain of games under consideration. Coalitional resource monotonicity is compatible with the core principle on suitably restricted classes of games. The most interesting example is that of *convex* games (Shapley, 1971). Roughly speaking, a game v is convex if it exhibits increasing marginal worths: $v(S) - v(S\backslash i) \leq v(T) - v(T\backslash i)$ whenever $i \in S \subset T$. It is well known that the Shapley value is in the core of such games.[6] Another rule satisfying coalitional resource monotonicity and the core principle on convex games is

[6] Shapley (1971) shows that it is, in fact, the barycenter of the core. The Shapley value is known to belong to the core on classes that strictly include the convex games: the quasi-convex games defined in Sprumont (1990) are just an example. See Marin-Solano and Rafels (1996), for a systematic analysis of this issue.

the *egalitarian rule* of Dutta and Ray (1989), which selects the unique payoff vector that is Lorenz-undominated in the core.[7] See also Hougaard, Peleg, and Osterdal (2003).

Finally, returning to (a variant of) the strong monotonicity property described in the opening paragraph of this section, we note that the core principle is compatible on the domain of convex games with the following condition: if v and v' are two games such that $v \leq v'$, then, for every $x' \in F(v')$, there exists $x \in F(v)$ such that $x_i \leq x'_i$ for all $i \in N$. In fact, the core correspondence itself has this property, as shown by Ichiishi (1990).

The Dutta-Ray egalitarian rule triggered recent interest in a particular subclass of allocation rules: those maximizing a fixed social welfare ordering – or perhaps just a partial ordering – subject to the core constraints. Recent contributions include Arin and Inarra (2001), Hougaard, Peleg, and Thorlund-Petersen (2001), Koster (2002), among others.

Requiring that collective choices be based on a fixed social ordering turns out to have far-reaching consequences. Suppose that $W : \mathbb{R}^N \to \mathbb{R}$ is a differentiable, strictly increasing, and strictly concave social welfare function, let $\mathcal{V}^*$ be the set of games with a nonempty core, and define the single-valued allocation rule f on $\mathcal{V}^*$ by the condition that $W(f(v)) \geq W(x)$ for all x in the core of v and all $v \in \mathcal{V}^*$. Hougaard and Osterdal (2005) show that f violates aggregate resource monotonicity.

3.2 Population Monotonicity

The literature on population monotonicity in cooperative games has mainly focused on environments where the arrival of new agents is beneficial to those originally present. The strong form of this "positive" version of the population monotonicity principle is as follows.

Population Monotonicity. If $(N,v),(N',v')$ are two games such that $N \subset N'$ and $v(S) = v'(S)$ for every coalition $S \subset N$, then $x_i \leq x'_i$ for all $x \in F(N,v)$, all $x' \in F(N',v')$, and all $i \in N$.

As in the fair division problem, population monotonicity implies single-valuedness: just choose $(N',v') = (N,v)$ in the definition.

The strength of the axiom depends on the richness of the domain of games under consideration. From now on, we will maintain the assumption that

$$\text{if } (N,v) \in \mathcal{G} \text{ and } S \in P(N), \text{ then } (S,v^S) \in \mathcal{G}, \tag{5}$$

where v^S is the restriction of v to $P(S)$: every subgame of a game in the domain is also in the domain. A population-monotonic rule exists on such a domain only if every game has a nonempty core. Indeed, if $(N,v) \in \mathcal{G}$ and f is population-monotonic,

[7] Uniqueness is proved by Dutta and Ray. If the game is not convex, the core may contain several Lorenz-undominated payoff vectors.

$f_i(S, v^S) \le f_i(N, v)$ whenever $i \in S \in P(N)$. Summing up these inequalities yields that $v(S) \le \sum_{i \in S} f_i(N, v)$ for every $S \in P(N)$, meaning that (N, v) has a nonempty core. Of course, the richness assumption (5) implies at once that (N, v) is in fact "totally balanced": *all* of its subgames have a nonempty core.

The bulk of the research has considered the smallest domains satisfying the richness condition (5), namely, those formed by a single game (N, v) and all of its subgames. A population-monotonic rule on such a minimal domain is known as a *population-monotonic allocation scheme* of the game (N, v) (Moulin, 1990a; Sprumont, 1990). It specifies how the worth of each coalition S should be split among its members, should coalition S form: writing this $f(S)$ instead of $f(S, v^S)$ will cause no confusion because v is fixed. Such minimal domains are naturally generated by economic environments in which resources are fixed and agents are drawn from a fixed set N of potential agents having fixed preferences.

Before discussing the existence of population-monotonic allocation schemes, we stress that population monotonicity is extremely demanding if $\mathcal{G}$ is not a minimal domain. Even if all games (N, v) in $\mathcal{G}$ have a population-monotonic allocation scheme, the population monotonicity axiom may be impossible to meet on $\mathcal{G}$. Consider for instance the games $(\{1,2,3\}, v)$ and $(\{1,2,3\}, w)$ given by $v(i) = v(23) = 0$ for all i, $v(S) = 1$ otherwise, and $w(i) = w(13) = 0$ for all i, $w(S) = 1$ otherwise. The first game has a unique population-monotonic allocation scheme: agent 1 receives the entire worth of every coalition to which he belongs. In the second game, it is agent 2 who receives the full worth of every coalition to which he belongs. Since the restrictions of v and w to $P(\{1,2\})$ are identical, population monotonicity is impossible on the domain consisting of the two games and their subgames.

While every game having a population-monotonic allocation scheme is totally balanced, the converse statement is false. A simple example is again provided by the four-agent "glove game" described earlier, in which 1 and 2 own a left glove and 3 and 4 a right glove. The core of each three-agent subgame consists of a unique payoff vector giving 0 to the two agents owning the same type of glove and 1 to the agent owning the rare type. Population monotonicity therefore implies that everyone should get at least 1 in the four-agent game, which is impossible.

The entire class of games having a population-monotonic allocation scheme is not hard to describe. To alleviate notations and terminology, let us fix the set of agents N and call a worth function v on $P(N)$ a game. A game v is *normalized* if $v(i) = 0$ for all $i \in N$, and *monotonic* if $v(S) \le v(T)$ for any two nested coalitions $S \subset T$. It is *simple* if $v(S)$ is either 0 or 1 for every coalition S and $v(N) = 1$; coalition S is said to be *winning* if $v(S) = 1$. A *vetoer* is an agent who belongs to all winning coalitions and a *veto game* is a normalized simple monotonic game v whose set of vetoers $A(v)$ is nonempty.

The fundamental observation is that every veto game v has a population-monotonic allocation scheme: for each winning coalition S, merely define

$$f_i(S) = \left\{ \begin{array}{ll} \frac{1}{|A(v)|} & \text{if } i \in A(v), \\ 0 & \text{otherwise.} \end{array} \right\}. \tag{6}$$

Monotonicity is obvious and budget balance holds because $A(v) \subset S$ for every winning S.

Going one step further, we note that every normalized game v that is a non-negative linear combination of veto games also has a population-monotonic allocation scheme: just construct the corresponding linear combination of the allocation schemes of the veto games spanning v. As it turns out, the converse is also true.

Theorem 7. (Sprumont, 1990) *A normalized game has a population-monotonic allocation scheme if and only if it is a nonnegative linear combination of veto games.*

Proof. We need only prove the "only if" statement. Let v be a normalized game having a population-monotonic allocation scheme f. The case where $v(N) = 0$ is trivial: the game must then be identically zero and it can obviously be written as a nonnegative linear combination of the veto games by choosing all weights equal to zero. From now on, we assume that $v(N) \neq 0$. Obviously, $v(N) > 0$ and there is at least one $i \in N$ such that $f_i(N) > 0$.

For each $i \in N$ such that $f_i(N) > 0$, define the game v_i on $P(N)$ by $v_i(S) = f_i(S)$ if $S \ni i$ and $v_i(S) = 0$ otherwise. This game is monotonic and $v = \sum_{i \in N: f_i(N)>0} v_i$.

To prove our claim, it remains to express each v_i as a nonnegative linear combination of veto games. Fix i and proceed recursively. Let

$$v_i^1(S) = 1 \text{ if } v_i(S) > 0, \text{ and 0 otherwise,}$$

and

$$\lambda^1 = \min_{S: v_i(S)>0} v_i(S).$$

Because v_i is monotonic, v_i^1 is a veto game. Next, construct the game $v_i - \lambda^1 v_i^1$. If this game is identically zero, we are done. Otherwise, let

$$v_i^1(S) = 1 \text{ if } (v_i - \lambda^1 v_i^1)(S) > 0, \text{ and 0 otherwise,}$$

and

$$\lambda^2 = \min_{S:(v_i - \lambda^1 v_i^1)(S)>0} (v_i - \lambda^1 v_i^1)(S).$$

Continue in that way until $v_i^K = 0$. By construction, $v_i = \sum_{k=1}^{K-1} \lambda^k v_i^k$. Each λ^k is positive, and each v_i^k, $k = 2, \ldots, K-1$, is a veto game provided that v_i^{k-1} is one. Since v_i^1 is a veto game, we are done. □

The normalization restriction in Theorem 7 is not important. In general, a game has a population-monotonic allocation scheme if and only if it is the sum of a nonnegative linear combination of veto games and an additive game v_a given by $v_a(S) = a \mid S \mid$ for all $S \in P(N)$, where a is an arbitrary real number.

In the spirit of the classic Bondareva-Shapley theorem on the nonemptiness of the core, Norde and Reijnierse (2002) provide an interesting dual description of the class of games having a population-monotonic allocation scheme.

Keeping the set of agents N fixed, define a *vector of sub-balanced weights* to be a list of real numbers $((\delta_S)_{S \in \Delta}, (\lambda_T)_{T \in \Lambda})$ with the following properties:

(i) Δ and Λ are disjoint subsets of $P(N)$,
(ii) $\delta_S > 0$ and $\lambda_T > 0$ for all $S \in \Delta$ and $T \in \Lambda$,
(iii) it is possible to assign to each triple $(i,S,T) \in N \times \Delta \times \Lambda$ such that $i \in T \subseteq S$ a nonnegative number $\mu^i_{S,T}$ in such a way that

$$\sum_{T\in\Lambda:i\in T\subseteq S} \mu^i_{S,T} = \delta_S \text{ for each } S \in \Delta \text{ and } i \in S, \text{ and}$$

$$\sum_{S\in\Delta:i\in T\subseteq S} \mu^i_{S,T} = \lambda_T \text{ for each } T \in \Lambda \text{ and } i \in T.$$

For instance, suppose $N = \{1,2,3,4\}$. Let $\Delta = \{\{1,2,3\},\{2,3,4\}\}$, $\Lambda = \{\{1,2\}, \{2,3\}, \{3,4\}\}$, and set $\delta_S = \lambda_T = 1$ for all $S \in \Delta$ and $T \in \Lambda$. Then $((\delta_S)_{S\in\Delta}, (\lambda_T)_{T\in\Lambda})$ is a vector of sub-balanced weights: take $\mu^1_{123,12} = \mu^2_{123,12} = \mu^2_{234,23} = \mu^3_{123,23} = \mu^3_{234,34} = \mu^4_{234,34} = 1$ and $\mu^2_{123,23} = \mu^3_{234,23} = 0$.

The main result of Norde and Reijnierse is the following theorem.

Theorem 8. (Norde and Reijnierse, 2002) *The game (N,v) has a population-monotonic allocation scheme if and only if*

$$\sum_{S\in\Delta} \delta_S v(S) \geq \sum_{T\in\Lambda} \lambda_T v(T)$$

for every vector of sub-balanced weights $((\delta_S)_{S\in\Delta}, (\lambda_T)_{T\in\Lambda})$.

For instance, the inequality corresponding to the particular vector given just before the theorem is

$$v(123) + v(234) \geq v(12) + v(23) + v(34). \tag{7}$$

To check that this inequality is necessary, note that if f is a population-monotonic allocation scheme, then

$$\begin{aligned} v(123) + v(234) &= f_1(123) + f_2(123) + f_3(123) + f_2(234) + f_3(234) + f_4(234) \\ &\geq f_1(12) + f_2(12) + f_3(23) + f_2(23) + f_3(34) + f_4(34) \\ &= v(12) + v(23) + v(34). \end{aligned}$$

For 4-agent games, Norde and Reijnierse identify explicitly the extreme points of the set of sub-balanced weights. These extreme points generate 60 independent linear inequalities of nine different types – one of them being (7) – which are together necessary and sufficient for the existence of a population-monotonic allocation scheme.

Theorems 7 and 8 do not offer any guidelines for recommending a particular population-monotonic allocation scheme in any game that has one. A few remarks on this issue may be useful. Let us call a population-monotonic allocation scheme f of a game v *symmetric* if

$$f_i(S \cup i) = f_j(S \cup j)$$

for every coalition $S \in P(N \backslash \{i, j\})$ and for all agents i, j which are symmetric (in the usual sense that $v(T \cup i) = v(T \cup j)$ for every $T \in P(N \backslash \{i, j\})$). Interestingly, every veto game v possesses a *unique* symmetric population-monotonic allocation scheme f: it is given by formula (6) above. The reason is as follows. For every winning coalition S, $\sum_{i \in S} f_i(S) = 1$. Since $v(N) = 1$, the monotonicity condition requires that $\sum_{i \in S} f_i(N) = 1$. It follows that every agent who is not a vetoer receives a zero share of $v(N)$: for if $f_j(N) > 0$ for some $j \in N \backslash A(v)$, there exists a winning coalition S (not containing j) such that $\sum_{i \in S} f_i(N) < 1$, a contradiction. The monotonicity requirement now implies that every agent who is not a vetoer receives zero in every coalition to which he belongs. Any positive worth is thus split among the vetoers. Since they are symmetric, formula (6) follows. To the extent that non-vetoers receive nothing in spite of making possibly significant contributions to some coalitions, this is a rather extreme and perhaps unappealing scheme. Population monotonicity is extremely constraining.

The observation made in the previous paragraph has further consequences. Let us denote by $\mathcal{V}^*(N)$ the class of normalized games having a population-monotonic allocation scheme. Let us define a *population-monotonic operator* to be a mapping which assigns to each game v in $\mathcal{V}^*(N)$ a population-monotonic allocation scheme $f(\cdot, v)$ of (N, v). Call it *symmetric* if it always selects symmetric allocation schemes. We claim that no symmetric population-monotonic operator can be additive if N contains at least three agents: there exist $v, w \in \mathcal{V}^*(N)$ such that $f(\cdot, v + w) \neq f(\cdot, v) + f(\cdot, w)$. To see this, let $N = \{1, 2, 3\}$ and let u_{12}, u_{23}, u_{123} be the three-agent simple "unanimity games" in which a coalition is winning if and only if it includes $\{1, 2\}, \{2, 3\}$, or $\{1, 2, 3\}$, respectively. Define also the three-agent game w by $w(i) = w(13) = 0$ for all i and $w(S) = 1$ otherwise. This is a veto game which is not a unanimity game; agent 2 is the only vetoer. Using the fact just noted, these four games each have a unique symmetric population-monotonic allocation scheme, yielding

$$f(N, u_{12}) = \left(\frac{1}{2}, \frac{1}{2}, 0\right), \; f(N, u_{23}) = \left(\frac{1}{2}, 0, \frac{1}{2}\right),$$
$$f(N, u_{123}) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), \; f(N, w) = (0, 1, 0).$$

Observe that $f(N, u_{12}) + f(N, u_{23}) \neq f(N, u_{123}) + f(N, w)$ even though $u_{12} + u_{23} = u_{123} + w : f$ is not additive.[8]

Not surprisingly, the difficulty disappears on the class of convex games. Computing the Shapley value of each subgame of a convex game v defines a population-monotonic allocation scheme of this game, as noted among others by Ichiishi (1988), Rosenthal (1990), and Sprumont (1990). Assigning this allocation scheme to every convex game defines a linear symmetric population-monotonic operator

[8] This impossibility is no surprise. Every population monotonic allocation scheme must obey the Dummy principle. With symmetry and additivity, this forces the operator to select the Shapley value in each game and subgame. But this does not yield a population monotonic allocation scheme in a game such as w.

over that class and, in fact, the Shapley value satisfies the very demanding Population Monotonicity axiom defined in the opening paragraph of this section. The Shapley value is not the only population-monotonic rule. Other examples include again the Dutta–Ray solutions, as noted by Dutta (1990), and the families of "sequential Dutta–Ray solutions" and "monotone-path Dutta–Ray solutions" defined in Hokari (2000, 2002).

The discussion in this subsection shows that population monotonicity is a very demanding requirement in the context of cooperative games. An alternative condition would impose that all agents be affected in the same direction by the arrival of newcomers.

Population Solidarity. If $(N,v),(N',v')$ are two games such that $N \subset N'$ and $v(S) = v'(S)$ for every coalition $S \subset N$, then (i) $x_i \leq x'_i$ for all $x \in F(N,v)$, all $x' \in F(N',v')$, and all $i \in N$ or (ii) $x_i \geq x'_i$ for all $x \in F(N,v)$, all $x' \in F(N',v')$, and all $i \in N$.

This condition is weaker than population monotonicity and deserves to be studied.

Acknowledgments This is a somewhat updated version of a survey based on lectures given at the NATO Advanced Study Institute on Game Theory and Resource Allocation, Stony Brook, 1997. I thank R. Aumann, J. Fraysse, F. Maniquet, H. Moulin, and especially W. Thomson for very useful conversations. I thank a referee for very helpful comments. I am grateful to the Social Sciences and Humanities Research Council of Canada for financial support.

References

Arin, J., & Inarra, E. (2001). Egalitarian solutions in the core. *International Journal of Game Theory*, *30*, 187–193

Aumann, R. (1985). On the non-transferable utility value: A comment on the Roth-Shafer examples. *Econometrica*, *53*, 667–677

Balasko, Y. (1979). Budget-constrained Pareto efficient allocations. *Journal of Economic Theory*, *21*, 359–379

Champsaur, P., & Laroque, G. (1981). Fair allocations in large economies. *Journal of Economic Theory*, *25*, 269–282

Chichilnisky, G., & Thomson, W. (1987). The Walrasian mechanism from equal division is not monotonic with respect to variations in the number of agents. *Journal of Public Economics*, *32*, 119–124

Dubins, L., & Spanier, E. (1961). How to cut a cake fairly. *American Mathematical Monthly*, *68*, 17

Dutta, B. (1990). The egalitarian solution and reduced game properties in convex games. *International Journal of Game Theory*, *19*, 153–169

Dutta, B., & Ray, D. (1989). A concept of egalitarianism under participation constraints. *Econometrica*, *57*, 465–479

Fleurbaey, M. (1996). Reward patterns of fair solutions. *Journal of Public Economics*, *59*, 365–395

Foley, D. (1967). Resource allocation and the public sector. *Yale Economic Essays*, *7*, 45–98

Gaspart, F. (1996). *A contribution to the theory of distributive justice* (Ph. D. Thesis). FUNDP, Namur

Geanakoplos, J., & Nalebuff, B. (1988). *On a fundamental conflict between equity and efficiency* (mimeo)

Hokari, T. (2000). Population monotonic solutions on convex games. *International Journal of Game Theory*, *29*, 327–338

Hokari, T. (2002). Monotone-path Dutta–Ray solutions on convex games. *Social Choice and Welfare*, *19*, 825–844

Hougaard, J., & Osterdal, L. (2005). *Monotonicity of social welfare optima* (mimeo) University of Copenhagen.

Hougaard, J., Peleg, B., & Osterdal, L. (2003). *The Dutta–Ray solution on the class of convex games: A generalization and monotonicity properties* (mimeo, forthcoming in International Journal of Game Theory)

Hougaard, J., Peleg, B., & Thorlund-Petersen, L. (2001). On the set of Lorenz maximal imputations in the core of a balanced game. *International Journal of Game Theory*, *30*, 147–165

Ichiishi, T. (1988). *The cooperative nature of the firm* (mimeo). Ohio State University

Ichiishi, T. (1990). Comparative cooperative game theory. *International Journal of Game Theory*, *19*, 139–152

Kalai, E. (1977). Proportional solutions to bargaining situations: Interpersonal utility comparisons. *Econometrica*, *45*, 1623–1630

Kalai, E., & Smorodinsky, M. (1975). Other solutions to Nash's bargaining problem. *Econometrica*, *43*, 513–518

Keiding, H., & Moulin, H. (1991). The solidarity axiom in parametric surplus-sharing problems. *Journal of Mathematical Economics*, *20*, 249–270

Kim, H. (2004). Population monotonic rules for fair allocation problems. *Social Choice and Welfare*, *23*, 59–70

Koster, M. (2002). Hierarchical constrained egalitarianism in TU-games. *Mathematical Social Sciences*, *43*, 251–265

Maniquet, F., & Sprumont, Y. (2000). On resource monotonicity in the fair division problem. *Economics Letters*, *68*, 299–302

Marin-Solano, J., & Rafels, C. (1996). *Convexity versus average convexity: Potential, PMAS, the Shapley value and simple games* (mimeo). Universitat de Barcelona

Meggido, N. (1974). On the non monotonicity of the bargaining set, the kernel, and the nucleolus of a game. *SIAM Journal of Applied Mathematics*, *27*, 355–358

Moulin, H. (1990a). Cores and large cores when population varies. *International Journal of Game Theory*, *19*, 219–232

Moulin, H. (1990b). Fair division under joint ownership: Recent results and open problems. *Social Choice and Welfare*, *7*, 439–452

Moulin, H. (1991). Welfare bounds in the fair division problem. *Journal of Economic Theory*, *54*, 321–337

Moulin, H. (1995). *Cooperative microeconomics*. Princeton: Princeton University Press

Moulin, H., & Thomson, W. (1988). Can everyone benefit from growth? Two difficulties. *Journal of Mathematical Economics*, *17*, 339–345

Neuefeind, W. (1972). On continuous utility. *Journal of Economic Theory*, *5*, 174–176

Norde, H., & Reijnierse, H. (2002). A dual description of the class of games with a population monotonic allocation scheme. *Games and Economic Behavior*, *41*, 322–343

Pazner, E., & Schmeidler, D. (1978). Egalitarian equivalent allocations: A new concept of economic equity. *Quarterly Journal of Economics*, *92*, 671–687

Polterovich, V., & Spivak, V. (1983). Gross substitutability of point-to-set correspondences. *Journal of Mathematical Economics*, *92*, 117–140

Rosenthal, E. (1990). Monotonicity of the core and value in dynamic cooperative games. *International Journal of Game Theory*, *19*, 45–57

Shapley, L. (1953). A value for n-person games. In H. Khun & A. Tucker (Eds.), *Contributions to the theory of games ii, annals of mathematics studies 28.* Princeton, NJ: Princeton University Press

Shapley, L. (1971). Cores of convex games. *International Journal of Game Theory*, *1*, 1–26

Shubik, M. (1962). Incentives, decentralized control, the assignment of joint costs, and internal pricing. *Management Science*, *8*, 325–343
Sprumont, Y. (1990). Population monotonic allocation schemes for cooperative games with transferable utility. *Games and Economic Behavior*, *2*, 378–394
Sprumont, Y.(1996). Axiomatizing ordinal welfare egalitarianism when preferences may vary. *Journal of Economic Theory*, *68*, 77–110
Sprumont, Y., & Zhou, L. (1999). Pazner-Schmeidler rules in large societies. *Journal of Mathematical Economics*, *31*, 321–339
Steinhaus, H. (1948). The problem of fair division. *Econometrica*, *16*, 101–104
Thomson, W. (1983a). The fair division of a fixed supply among a growing population. *Mathematics of Operations Research*, *8*, 319–326
Thomson, W. (1983b). Problems of fair division and the egalitarian solution. *Journal of Economic Theory*, *31*, 211–226
Thomson, W. (1994). Notions of equal, or equivalent, opportunities. *Social Choice and Welfare*, *11*, 137–156
Thomson, W. (1995). Population-monotonic allocation rules. In W. Barnett, H. Moulin, M. Salles, & N. Schofield (Eds.), *Social choice, welfare and ethics* (pp. 79–124). Cambridge: Cambridge University Press
Thomson, W. (1999). Règles dállocation monotones en ressources: une revue de la littérature. *Revue Économique*, *50*, 757–796
Thomson, W., & Myerson, R. (1980). Monotonicity and independence axioms. *International Journal of Game Theory*, *9*, 37–49
Varian, H. (1976). Two problems in the theory of fairness. *Journal of Public Economics*, *5*, 249–260
Young, P. (1980). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *9*, 37–49

To Envy or To Be Envied? Refinements of the Envy Test for the Compensation Problem

Marc Fleurbaey

1 Introduction

The envy test concept is an all-or-nothing notion, and this is problematic when there is no achievable envy-free option. The idea of ranking the "unfair" social states on the basis of how much envy they contain goes back at least to Feldman and Kirman (1974) and Varian (1976), but it is in Suzumura (1981a, b, 1983) that one finds a first systematic study of this issue. More recent contributions to this line of research include Chauduri (1986), Diamantaras and Thomson (1990), Tadenuma (2002), and Tadenuma and Thomson (1995).

One of the contexts where typically envy-free allocations are hard to achieve is when, as noticed in Pazner and Schmeidler (1974) for the production case and Fleurbaey (1994) for the distribution case, individuals have nontransferable personal characteristics to which the no-envy test nonetheless applies. This is sometimes called the "compensation problem," where one tries to compensate inequalities in personal characteristics by counteracting inequalities in transferable resources. This problem is rather different from standard problems of resource allocation because of the presence of nontransferable characteristics, which act as a constraint on redistribution. One typical example is when individuals have unequal levels of skills which give them unequal earning possibilities. Certain other characteristics, such as physical handicaps, may directly affect personal satisfaction and generate inequalities which call for redressing transfers. When inequalities in personal characteristics are huge, or when individuals disagree about the value of their respective characteristics, it may be very hard, or even impossible, to find transfers that eliminate envy between individuals. In Fleurbaey (1994) and Iturbe-Ormaetxe and Nieto (1996) one finds several suggestions about how to weaken the no-envy requirement in order to obtain nonempty solutions to the compensation problem. But most of these

M. Fleurbaey
CERSES, CNRS and University Paris Descartes, and IDEP
e-mail: marc.fleurbaey@univ-paris5.fr

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

solutions fail to be nonempty on the whole domain, and a systematic use of rankings seems not to have been attempted yet in this branch of the literature. This chapter makes an attempt at filling this gap and examines several rankings, which may be of some interest.

Section 2 makes a brief survey of the compensation literature, proposing a few basic criteria for the evaluation of solutions. Section 3 examines rankings based on the number of envy relations, Section 4 deals with rankings that make use of additional information about the population's preferences, and Section 5 is devoted to rankings that involve the degree of envy as measured by the quantity of transfers that would be needed to suppress envy relations. It argues that such rankings are preferable to the others, and also establishes a correspondence between one such ranking and another based on the idea of rationalizing egalitarian competitive equilibria. Section 6 concludes the chapter.

2 A Brief Survey

The compensation problem can be described by the following simple model. The population is $N = \{1,...,n\}$ and every individual $i \in N$ is endowed with two kinds of characteristics: y_i, for which she is not responsible (circumstances), and z_i, for which she is. A profile of characteristics is $(y_N, z_N) = ((y_1, \ldots, y_n), (z_1, \ldots, z_n))$. The sets from which y_i and z_i are drawn, denoted Y and Z, respectively, are assumed to have at least two elements.

Individual i's well-being is denoted u_i and is determined by a function u, which is the same for all individuals:

$$u_i = u(x_i, y_i, z_i),$$

where $x_i \in \mathbb{R}$ is the quantity of money transfer to which the individual is submitted. When $x_i < 0$, the transfer is a tax. The real-valued function u, defined either on $\mathbb{R} \times Y \times Z$ or $\mathbb{R}_+ \times Y \times Z$ depending on the cases (on which more below), is assumed to be continuous and increasing in x_i.

An allocation is denoted $x_N = (x_1, \ldots, x_n)$. The set of feasible allocations is denoted X. The precise definition of X differs in different cases, but typically involves a condition $\sum_{i \in N} x_i \leq \Omega$ for some aggregate endowment $\Omega \in \mathbb{R}$. Given the fact that u is increasing in x_i, this means that allocations such that $\sum_{i \in N} x_i = \Omega$ are all Pareto efficient. This considerably simplifies the analysis.

This is the simplest model in which the compensation problem can be studied, but other models have been studied. In particular, Fleurbaey and Maniquet (1996) and Pazner and Schmeidler (1974) have examined a production model in which agents differ in their productivity. For a general survey on the compensation problem, see Fleurbaey and Maniquet (2002).

Two general concepts of solutions will be useful here. Let $\mathcal{D}$ be the domain of economies $e = (y_N, z_N)$ under consideration. An *allocation rule* is a correspondence

$S : \mathcal{D} \rightarrow\rightarrow X$, such that for all $e \in \mathcal{D}$, $S(e) \subset X$ is the subset of allocations selected by S. A *social ordering function* is a mapping $R : \mathcal{D} \rightarrow \mathcal{R}_X$, where $\mathcal{R}_X$ is the set of complete orderings over X. The expression $x_N\, R(e)\, x'_N$ will mean that x_N is weakly preferred to x'_N, and $P(e)$ and $I(e)$ will denote the corresponding strict preference and indifference relations, respectively. An *allocation rule derived from a social ordering function* is defined by selecting, for each economy, the maximal elements in X for the social ordering defined by the social ordering function for this economy.

Two special cases will be of particular interest. The "distribution" case (Fleurbaey, 1994) is when x_i has to be nonnegative, and there is a fixed amount $\Omega > 0$ to be distributed, that is, when (assuming no waste)

$$X = \left\{ x_N \in \mathbb{R}^N_+ \mid \sum_{i \in N} x_i = \Omega \right\}.$$

An interesting domain for this case is the domain $\mathcal{D}_1$ of economies satisfying, for all $i, j \in N$,

$$u\left(\frac{\Omega}{|N|-1}, y_i, z_i \right) \geq u(0, y_j, z_i).$$

This domain is such that no individual considers his own y_i to be a huge handicap compared to other values of y_j in the population.[1]

The "TU" (transferable utility) case (Bossert, 1995) is when the well-being function is quasi-linear in x,

$$u_i = x_i + v(y_i, z_i),$$

x_i is not bounded below,[2] and there is no external amount of money to be distributed, that is, when

$$X = \left\{ x_N \in \mathbb{R}^N \mid \sum_{i \in N} x_i = 0 \right\}.$$

The distribution case is relevant to situations in which the government has a fixed budget that can be used in order to provide targeted help to particular categories of people, such as disabled individuals, victims of a natural disaster, families with different needs. The TU case is not limited, but is especially relevant, to situations in which individual well-being is itself monetary. The most realistic applications of the TU case are offered by the federalism problem of organizing budget transfers between administrative units (local governments, sectorial administrations, social security agencies, etc.), which are partly responsible for their budget situation.

For the most part we focus here on the TU case, and only briefly mention the differences in results for the distribution case, when relevant.

---

[1] Conditions of this kind are helpful in order to obtain the existence of envy-free allocations in the model where y is an indivisible good that is transferable across agents. See, for example, Maskin (1987).

[2] More realistically, one could impose that $x_i + v(y_i, z_i)$ is bounded below (for instance by zero). We will not study this variant.

The compensation problem consists in neutralizing the impact of circumstances y on well-being while not interfering with inequalities due to differences in responsibility characteristics z. The no-envy condition, due to Foley (1967) and Kolm (1972), is well suited to this purpose if it is applied as follows.

$$\text{No-Envy:} \quad \forall e \in \mathcal{D}, \forall x_N \in S(e), \forall i, j \in N, u(x_i, y_i, z_i) \geq u(x_j, y_j, z_i)\,.$$

Its main drawback, however, is that it is too demanding and is satisfied only on a very small domain. This is connected to the conflictual duality between two principles that it jointly encapsulates, namely, the compensation principle ("neutralize y") and the natural reward principle ("not interfere with z") (Fleurbaey, 1995). Here we focus on a small list of axioms embodying these principles. For the compensation principle:

Equal Well-Being for Equal Responsibility: $\forall e \in \mathcal{D}$, $\forall x_N \in S(e)$, $\forall i, j \in N$ such that $z_i = z_j$,

$$u(x_i, y_i, z_i) = u(x_j, y_j, z_j)\,.$$

Equal Well-Being for Uniform Responsibility: $\forall e \in \mathcal{D}$, $\forall x_N \in S(e)$, if $\forall i, j \in N$, $z_i = z_j$, then

$$\forall i, j \in N,\; u(x_i, y_i, z_i) = u(x_j, y_j, z_j)\,.$$

In the distribution case one can reformulate these axioms in terms of application of the leximin criterion, inequality being allowed when the better-off agent has a zero x. In the above formulation, however, they are nonempty on $\mathcal{D}_1$ (see Lemma 1 in Fleurbaey (1994)).

The dual "natural reward" axioms are the following:

Equal Treatment for Equal Circumstances: $\forall e \in \mathcal{D}$, $\forall x_N \in S(e)$, $\forall i, j \in N$ such that $y_i = y_j$,

$$x_i = x_j.$$

Equal Treatment for Uniform Circumstances: $\forall e \in \mathcal{D}$, $\forall x_N \in S(e)$, if $\forall i, j \in N$, $y_i = y_j$, then

$$\forall i, j \in N,\; x_i = x_j.$$

These four axioms appear to be very basic conditions, and a reasonable requirement for an allocation rule is that it should satisfy at least the two weak axioms ("uniform" case) and one of the strong axioms ("equal" case), knowing that the two strong axioms are incompatible (Fleurbaey, 1994).

Three allocation rules, conceived in terms of weakening the no-envy requirement, have been proposed in Fleurbaey (1994). One, inspired from Daniel (1975) and Feldman and Kirman (1974), selects the allocations with the smallest number of envy occurrences among the "balanced" allocations. A balanced allocation is such that for all $i \in N$, the number of agents he envies equals the number who envy him:

$$\left|\left\{j \in N \mid u(x_i, y_i, z_i) < u(x_j, y_j, z_i)\right\}\right| = \left|\left\{j \in N \mid u(x_i, y_i, z_j) > u(x_j, y_j, z_j)\right\}\right|.$$

Let $B(e) \subseteq X$ denote the subset of balanced allocations, and $E(x_N, e)$ denote the number of envy occurrences in x_N :

$$E(x_N, e) = \left|\left\{(i, j) \in N \mid u(x_i, y_i, z_i) < u(x_j, y_j, z_i)\right\}\right|.$$

Balanced and Minimal Envy ($S_{\mathbf{BME}}$): $\forall e \in \mathcal{D}, \forall x_N \in X,$

$$\begin{aligned} x_N \in S_{\mathrm{BME}}(e) \Leftrightarrow x_N \in B(e) &\text{ and} \\ &\forall x'_N \in B(e),\ E(x'_N, e) \geq E(x_N, e). \end{aligned}$$

A second allocation rule, inspired by Chauduri (1986) and Diamantaras and Thomson (1990), tries to minimize the intensity of envy, this intensity being measured for every agent by the resource needed to make this agent non-envious:

$$I_i(x_N, e) = \min\{\delta \in \mathbb{R} \mid \forall j \in N \setminus \{i\},\ u(x_i + \delta, y_i, z_i) \geq u(x_j, y_j, z_i)\}.$$

The allocation rule is then defined as follows.

Minimax Envy Intensity ($S_{\mathbf{MEI}}$): $\forall e \in \mathcal{D}, \forall x_N \in X,$

$$x_N \in S_{\mathrm{MEI}}(e) \Leftrightarrow \forall x'_N \in F(e),\ \max_{i \in N} I_i(x'_N, e) \geq \max_{i \in N} I_i(x_N, e).$$

The third allocation rule makes use of all agents' opinions about the relative well-being of two agents. It tries to minimize the size of subsets of agents thinking that one agent is worse-off than another agent. It takes inspiration from "undominated diversity" (Parijs, 1990, 1995), which seeks to avoid situations in which one agent is deemed unanimously worse-off than another one, and is related to the family of solutions put forth by Iturbe-Ormaetxe and Nieto (1996), which generalizes van Parijs' idea and seeks to avoid such a unanimity among a subgroup of a given size and containing the worse-off agent. Let

$$N_i^m = \{G \subset N \mid |G| = m,\ i \in G\}.$$

Minimal Unanimous Domination ($S_{\mathbf{MUD}}$): $\forall e \in \mathcal{D}, \forall x_N \in X,$

$$\begin{aligned} &x_N \in S_{\mathrm{MUD}}(e) \Leftrightarrow \exists m \in \{1, \ldots, n\}, \\ &\qquad \begin{cases} \text{(i) } \forall i, j \in N, \forall G \in N_i^m, \exists k \in G,\ u(x_i, y_i, z_k) \geq u(x_j, y_j, z_k), \\ \text{(ii) } \forall p < m, \forall x'_N \in F(e), \exists i, j \in N, \exists G \in N_i^p, \\ \qquad \forall k \in G,\ u(x'_i, y_i, z_k) < u(x'_j, y_j, z_k). \end{cases} \end{aligned}$$

3 Ranking Envy Graphs

A difficulty with S_{BME} is that there do not always exist balanced allocations, so that the domain on which it is defined is restricted. Moreover, the various sufficient conditions of existence specified by Daniel (1975) and Fleurbaey (1994) are not very easy to interpret and apply. Necessary conditions have not been studied to the best of my knowledge.

A more substantial criticism is that one does not see why a lexicographic priority should be given to balancedness of allocations over the number of envy occurrences. If the only balanced allocation has everybody envying everybody, it might be better to prefer an unbalanced allocation with a much smaller number of envy relations.

In fact, the general idea underlying this allocation rule is to examine the graph of envy relations, with a double concern for symmetry and for minimizing the number of relations. A general approach to the problem of ranking envy graphs would probably be more suitable than a narrow focus on balanced allocations.

In Fig. 1, five graphs are represented for a population of four individuals. In case (a), individual 1 is envied by all the others; in case (b), individual 1 envies all the others; in case (c), a cycle of envy occurs; in case (d), an envy relation has been reversed in comparison to case (c); in case (e), this envy relation has been deleted. Although the number of envy relations is smaller in (a) and (b) than in (c), and the same in (d) as in (c), one should probably prefer (c), out of a concern for symmetry. Although (e) is not symmetric, it may not be worse than (c), because it has a strictly smaller graph of envy relations.

This is a difficult case since balancedness in (c) makes it look more equal than (e), but one must be careful to avoid the intuitive illusion that arrows from a transitive "better-off than" relation. In (e), there is no-envy between agents 1 and 2, and this is the relevant test of equality. The arrows from 2 to 1 via 3 and 4 do not mean that 2 is worse-off than 1. It is true that between two agents, reciprocal envy appears better than a one-way envy relation. But this does not necessarily extend to cycles of envy among more agents. Therefore, it is not unreasonable to consider that removing a nonreciprocal envy relation between two agents is always a good thing (given that the only information is the envy graph, a limitation that will be discussed below).

Defining a more precise preference order on envy graphs is a complex matter. Suzumura (1983) proposes a very natural ranking, which applies the reverse leximin criterion to the vector of individual envy indices, where an individual envy index is

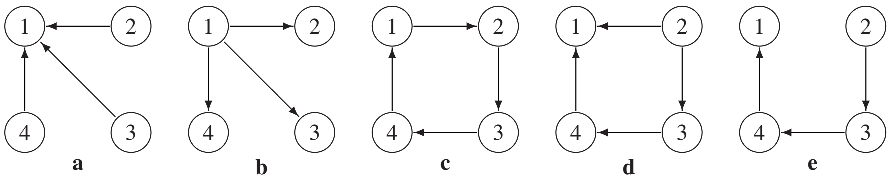


Fig. 1 Envy graphs

simply the number of other agents this individual envies (i.e., the number of outgoing arrows in the graph). The reverse leximin criterion prefers a vector to another if its greatest component is smaller, or if the greatest components are equal in the two vectors but the second greatest component is smaller, and so on. It corresponds to the application of the standard leximin criterion to the opposite vectors.

The symbol $\geq_{\text{lex}}$ appearing in the definition below denotes the standard leximin criterion applied to vectors of real numbers. Namely, $x \geq_{\text{lex}} x'$ if the smallest component of x is greater than the smallest component of x', or they are equal and the second smallest component of x is greater than the second smallest component of x', and so on.

Envious Count criterion (R_{EsC}): Let

$$n_i(x_N) = \left|\left\{ j \in N \mid u(x_i, y_i, z_i) < u(x_j, y_j, z_i)\right\}\right|.$$

For all $x_N, x'_N \in X$, $x_N \, R_{\text{EsC}}(e) \, x'_N$ if and only if

$$-\left(n_i(x_N)\right)_{i\in N} \geq_{\text{lex}} -\left(n_i(x'_N)\right)_{i\in N}.$$

One can define a dual criterion to this one, that relies on the number of agents by whom a given agent is envied.

Envied Count criterion (R_{EdC}): Let

$$n'_i(x_N) = \left|\left\{ j \in N \mid u(x_i, y_i, z_j) > u(x_j, y_j, z_j)\right\}\right|.$$

For all $x_N, x'_N \in X$, $x_N \, R_{\text{EdC}}(e) \, x'_N$ if and only if

$$-\left(n'_i(x_N)\right)_{i\in N} \geq_{\text{lex}} -\left(n'_i(x'_N)\right)_{i\in N}.$$

The envious count and envied count criteria appear to be dual with respect to compensation and natural reward. One can indeed make the following observation. When two agents have the same responsibility characteristics, the worse-off will envy at least all the agents envied by the other plus the other agent himself, which means that his n_i index is greater and the reverse leximin will give absolute priority to him. Similarly, when two agents have the same circumstances, those who envy one of them will systematically envy the agent with the greater x, and the one with the lower x will envy him as well, so that his n'_i will be greater and absolute priority will be put on him (i.e., absolute priority *against* him in this case, in order to reduce the number of agents envying him). This provides the intuition for the following result, which bears on the allocation rules derived from R_{EsC} and R_{EdC}.

Proposition 1. *In the TU case, the allocation rule derived from the envious count criterion satisfies equal well-Being for equal responsibility and the allocation rule derived from the envied count criterion satisfies equal treatment for equal circumstances. This result does not extend to the distribution case.*

Proof. (1) Consider two agents $i, j \in N$ such that $z_i = z_j$ and an allocation x_N such that $u_i > u_j$. Let x'_N be such that $x'_i = x_i - (u_i - u_j)$ and $x'_k = x_k$ for all

$k \neq i$. Then $n_i(x'_N) = n_j(x'_N) = n_j(x_N) - 1$, while $n_k(x'_N) \leq n_k(x_N)$ for all $k \neq i, j$. Since $n_j(x_N) > n_i(x_N)$, the vector $(n_i(x'_N), n_j(x'_N))$ is better for the reverse leximin than $(n_i(x_N), n_j(x_N))$, and since $n_k(x'_N) \leq n_k(x_N)$ for all $k \neq i, j$, the whole vector $(n_k(x'_N))_{k \in N}$ is better than $(n_k(x_N))_{k \in N}$. The allocation x'_N does not belong to X, but the allocation

$$x''_N = x'_N - \frac{1}{|N|} \sum_{i \in N} x'_i$$

does and is such that $n_k(x''_N) = n_k(x'_N)$ for all $k \in N$.

(2) The proof for the envied count criterion and equal treatment for uniform circumstances is similar. Let $y_i = y_j$ and $x_i > x_j$ in allocation x_N. Allocation x'_N is defined by $x'_j = x_i$ and $x'_k = x_k$ for all $k \neq j$. The rest is very similar as above.

(3) Impossibility to extend to the distribution case is a corollary of the next proposition. □

However, these two criteria display no concern for balancedness. For instance, in the examples of Fig. 1, the vectors of indices $n_i(x_N)$ are, respectively, the following:
(a) (0,1,1,1)
(b) (3,0,0,0)
(c) (1,1,1,1)
(d) (0,2,1,1)
(e) (0,1,1,1)
As a consequence, the envious count criterion ranks the five graphs in the following decreasing order:

a e
c
d
b

Indifference between (a) and (e) is due to the fact that this ranking is not sensitive to the balancedness feature of graphs. It only counts the number of outgoing arrows and is indifferent to the direction of these arrows. A similar difficulty is obtained with the envied count criterion, which puts (a) at the bottom, but is indifferent between (b) and (e).[3]

A concern for balancedness can be incorporated by measuring individual situations with respect to envy in terms of an index that depends on n_i and on n'_i. Let

$$d_i(x_N) = D(n_i(x_N), n'_i(x_N))$$

for a function D, the properties of which are discussed below. One can apply the reverse leximin criterion to such indices. The properties of the criterion will then depend on how D ranks various (n_i, n'_i) vectors. Figure 2 shows iso-curves for the D function in the (n_i, n'_i) space.

Panels (1) and (2) illustrate the two extreme cases of the envious count and envied count criteria:

[3] The vectors of indices $n'_i(x_N)$ are, respectively, (a) (3,0,0,0); (b) (0,1,1,1); (c) (1,1,1,1); (d) (2,0,1,1); and (e) (1,0,1,1).

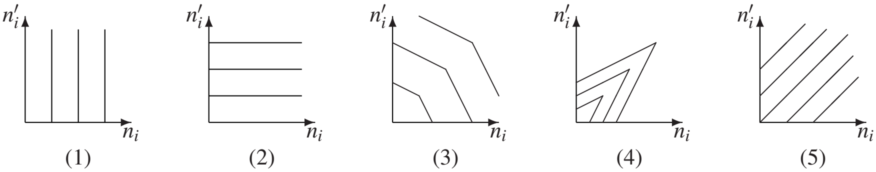


Fig. 2 Iso-curves of D

(1) $D(n_i, n'_i) = n_i$;
(2) $D(n_i, n'_i) = n'_i$.

Panel (3) and (4) correspond to cases in which a concern for balancedness is introduced:

(3) $D(n_i, n'_i) = 2\max\{n_i, n'_i\} + \min\{n_i, n'_i\}$.
(4) $D(n_i, n'_i) = 2\max\{n_i, n'_i\} - \min\{n_i, n'_i\} = \max\{n_i, n'_i\} + |n_i - n'_i|$.

As far as the examples of Fig. 1 are concerned, formula (3) puts (e) above (c) whereas formula (4), displaying a greater concern for balancedness, puts (c) above (e). Panel (5) depicts the extreme case in which only balancedness matters:

(5) $D(n_i, n'_i) = |n_i - n'_i|$.

It turns out that none of these criteria satisfies equal well-being for equal responsibility or equal treatment for equal circumstances in the distribution case. The next proposition shows that there is no hope to find better criteria along these lines. To keep things simple, attention is restricted to "reasonable" criteria that prefer an allocation with only one envy occurrence to any unbalanced allocation with more than n envy occurrences, for n great enough. This restriction seems unquestionable when dealing with criteria that rely only on envy graphs.

Proposition 2. *In the distribution case, no reasonable criterion based on envy graphs satisfies either equal well-being for equal responsibility or equal treatment for equal circumstances.*

Proof. In the distribution case, two agents i, j can be in a situation in which no one envies the other when they have certain x_i^*, x_j^*, whereas at all other allocations at least one envies the other. In such a case, let us say that i and j are "locked" at $\left(x_i^*, x_j^*\right)$. Let us illustrate how this can happen. Let $u(x, y_i, z_i) = u(x, y_j, z_i)$ for all x, and $u(x, y_j, z_j) < u(x, y_i, z_j)$ for all $x \neq x^*$, while $u(x^*, y_j, z_j) = u(x^*, y_i, z_j)$. Then i and j are locked at (x^*, x^*), since there is no envy at (x^*, x^*), whereas for (x_i, x_j) (with at least one different from x^*), i envies j if $x_i < x_j$ and j envies i if $x_i \geq x_j$.

Consider an n-agent population $\{1, \ldots, n\}$ where $z_1 = z_2$ and such that for all pairs of agents $i, j > 1$, i and j are locked at (1,1). Assume that $\Omega = n$, that agents $3, \ldots, n$ never envy agent 1 (whatever the allocation), that $u_1 > u_2$ at allocation $(1, \ldots, 1)$ and that 1 is envied by 2 at this allocation. Necessarily this is the only envy occurence in this allocation. At any other allocation in X, there will be at least

$n-2$ envy occurrences, because at least one of the agents $i > 1$ will have a different x and this will create at least one envy occurrence between him and each one of the others.

Moreover, no allocation in which 1 and 2 do not envy each other is balanced. First note that in such an allocation $u(x_1, y_1, z_1) = u(x_2, y_2, z_2)$, since $z_1 = z_2$. If 2 envies another agent, then 1 envies this other agent as well, but 1 is not envied by 2 in such an allocation, and is never envied by $3, \ldots, n$ in all allocations. In this case 1's situation is unbalanced. If 2 does not envy any other agent, he must be envied by at least one agent $3, \ldots, n$ and his situation is unbalanced.

Therefore, for n great enough, a reasonable criterion will prefer $(1, \ldots, 1)$ to any allocation in which 1 and 2 do not envy each other, and thereby violate equal well-being for equal responsibility.

For equal treatment for equal circumstances, assume $y_1 = y_2$, and all pairs of agents $i, j > 1$ are similarly locked together. Then, for certain preferences, the allocation $(0, 1, \ldots, 1)$ has only one envy occurrence, namely 1 envying 2. The rest of the argument is as above. □

This last result clearly suggests that the information contained in an envy graph is insufficient, and that can be interpreted as being due to the fact that this information is typically insufficient to pinpoint agents with identical y or identical z.

4 Undominated Diversity and Beyond

In this section we turn to a setting with richer information. Recall the S_{MUD} allocation rule, which seeks to minimize the size of the set of agents who unanimously consider that i is worse-off than j (and i is among them), for all pairs (i, j). This allocation rule refines van Parijs' undominated diversity, which is too large in some cases (in particular, it accepts allocations with envy when envy-free allocations exist). It shares with it the drawback that it may happen to be empty in the distribution case. It is, however, nonempty in a rather wide class of situations.

Lemma 1. *The S_{MUD} is nonempty in the TU case. It is also nonempty in the distribution case on $\mathcal{D}_1$.*

Proof. **Distribution case**: Fleurbaey (1994, Prop. 10) proves that, if for all $i, j \in N$, there is $k \in N$ such that

$$u\left(\frac{\Omega}{|N|-1}, y_i, z_k\right) \geq u(0, y_j, z_k),$$

then S_{MUD} is nonempty. This assumption is satisfied on $\mathcal{D}_1$.

TU case: Consider first a modified version of the TU case in which the feasible set is $X^* = \left\{x_N \in \mathbb{R}_+^N \mid \sum_{i \in N} x_i = \Omega\right\}$. Let

$$\Omega = (|N|-1) \max_{i,j,k \in N} \left(v(y_i, z_k) - v(y_j, z_k)\right).$$

With this value of Ω, the above assumption is satisfied, so that there exists an allocation $x_N \in X^*$ such that for all $i, j \in N$, there is $k \in N$ such that

$$u_i(x_i, y_i, z_k) \geq u(x_j, y_j, z_k).$$

Let $\mu = \frac{1}{|N|}\sum_{i\in N} x_i$, and define $x'_i = x_i - \mu$ for all $i \in N$. The allocation x'_N is such that $\sum_{i\in N} x'_i = 0$ and, by the quasi-linearity of u in the TU case, it still holds that for all $i, j \in N$, there is $k \in N$ such that

$$u_i(x_i, y_i, z_k) \geq u(x_j, y_j, z_k).$$

$\square$

In fact, the underlying idea of S_{MUD} is again to rank graphs of envy relations. But, interestingly, instead of simply counting the arrows between individuals, the idea is to assign a number to every envy relation, which is equal to the number of individuals who share the envious' preferences. For instance, suppose i envies j, and there are three other individuals who, with their own responsibility characteristics, would be better-off with j's bundle of external resources and circumstances than with i's. Then the envy arrow from i to j is assigned a value of four. When i does not envy j, no arrow is drawn even if there are some other individuals who would be better-off with j's bundle than with i's. The absence of an arrow is equivalent to a value of zero.

In summary, for every ordered pair (i, j), this procedure gives us a number, equal to zero if i does not envy j, and equal to a positive integer between one and the population size otherwise. "Undominated diversity" is simply the rather special requirement that no pair has a number with the maximal value ($|N|$). The S_{MUD} allocation rule applies the minimax criterion to the list of these numbers (i.e., it minimizes the greatest number), retaining in addition the requirement that no pair has number $|N|$.

A drawback of the minimax criterion is that it neglects the situation of envy relations with a less than maximal number and may therefore accept too much of envy. It appears much more reasonable to apply the reverse leximin criterion to the list of these numbers. Let us call this the "diversity" criterion, since it both extends and refines van Parijs' criterion, and takes account of the diversity of preferences in the population.

Diversity criterion (R_{D}): For any $x_N \in X$, $(i, j) \in N^2$, let

$$n_{ij}(x_N) = \begin{cases} 0 \text{ if } i \text{ does not envy } j, \\ \left|\left\{k \in N \mid u(x_j, y_j, z_k) > u(x_i, y_i, z_k)\right\}\right| \text{ otherwise.} \end{cases}$$

For all $x_N, x'_N \in X$, $x_N \, R_{\text{D}}(e) \, x'_N$ if and only if

$$-\left(n_{ij}(x_N)\right)_{i,j\in N} \geq_{\text{lex}} -\left(n_{ij}(x'_N)\right)_{i,j\in N}.$$

An envy-free allocation corresponds to a list containing only zeros, and will be selected whenever it exists. Similarly, if there exist allocations satisfying the undominated diversity criterion, the selected allocations will be drawn from this subset. An interesting feature of the diversity criterion (already present in undominated diversity) is that it satisfies equal treatment for equal circumstances.

Proposition 3. *The allocation rule derived from the diversity criterion exactly selects the set of envy-free allocations whenever it is nonempty. The allocation rule derived from it satisfies equal well-being for uniform responsibility (on $\mathcal{D}_1$ for the distribution case) and equal treatment for equal circumstances (on $\mathcal{D}_1$ for the distribution case).*

Proof. (1) An envy-free allocation is such that $(n_{ij}(x_N))_{i,j\in N} = 0$, and this dominates any $-(n_{ij}(x'_N))_{i,j\in N} < 0$ for the leximin criterion. Therefore, the set of envy-free allocations is selected whenever it is nonempty.

(2) When $z_i = z_j$ for all $i, j \in N$, the allocation that equalizes well-being across all agents is the only envy-free efficient allocation and is therefore selected whenever it is feasible, which is always true in the TU case, and on $\mathcal{D}_1$ in the distribution case. This proves the satisfaction of equal well-being for uniform responsibility.

(3) That it satisfies equal treatment for equal circumstances is a consequence of the fact that $n_{ij}(x_N) = |N|$ if $x_i < x_j$ while $y_i = y_j$ and that, in the TU case as well as in the distribution case on $\mathcal{D}_1$, by Lemma 1 there always exist (undominated diversity) allocations with $\max(n_{ij}(x_N))_{i,j\in N} < |N|$. In the distribution case, out of $\mathcal{D}_1$, an allocation satisfying equal treatment for equal circumstances is not always selected. Consider a situation with uniform z and $y_1 = y_2$ in which the leximin-utility allocation is $x_N = (3,3,0,\ldots,0)$, while $u(4,y_2,z) = u(0,y_i,z)$ for all $i \neq 1$. Then the allocation $x'_N = (2,4,0,\ldots,0)$ is preferred when $|N| > 3$, because it has

$$\left(n_{ij}(x'_N)\right)_{i,j\in N} = \begin{pmatrix} 0 & |N| & |N| & \cdots & |N| \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

compared to

$$(n_{ij}(x_N))_{i,j\in N} = \begin{pmatrix} 0 & 0 & |N| & \cdots & |N| \\ 0 & 0 & |N| & \cdots & |N| \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

□

The allocation rule associated with R_D is more satisfactory than those of the previous section, and this can be linked to the richer information used by R_D.

The diversity criterion is clearly on the side of the natural reward principle, and one may wonder if a dual criterion can be defined that would embrace the compensation principle instead. The dual criterion does exist, and refers to all the values of

y_k instead of the values of z_k. For a given y_k, it evaluates the situation of an individual i by computing the value of x that would give him, with circumstances y_k, the same utility as in the contemplated allocation. Now consider two individuals i and j. This computation amounts to imagining a situation that is equivalent in terms of well-being but with equal circumstances (y_k) for both individuals. The ideal allocation should then be egalitarian between them, and any inequality observed in x for this imaginary situation does reflect a problem. Now, different computations made with different y_k may yield different answers and this criterion takes account of this possible diversity.

Formally, for any $x_N \in X$, $i,k \in N$, let $x_{ik}(x_N)$ be defined by

$$u(x_i, y_i, z_i) = u(x_{ik}(x_N), y_k, z_i).$$

In the distribution case, it may happen that $u(x_i, y_i, z_i) > u(x, y_k, z_i)$ for all $x \geq 0$, or that $u(x_i, y_i, z_i) < u(x, y_k, z_i)$ for all $x \geq 0$. We focus here, for this case, on the domain $\mathcal{D}_2$ such that for all $i, j \in N$, one has $u(0, y_i, z_i) = u(0, y_j, z_i)$ and there exists $x \geq 0$ such that $u(\Omega, y_i, z_i) < u(x, y_j, z_i)$. On this domain, $x_{ik}(x_N)$ is always well defined.

Compensation Diversity criterion ($R_{\mathbf{CD}}$): For any $x_N \in X$, $(i, j) \in N$, let

$$m_{ij}(x_N) = \begin{cases} 0 \text{ if } i \text{ does not envy } j, \\ \left|\left\{k \in N \mid x_{jk}(x_N) > x_{ik}(x_N)\right\}\right| \text{ otherwise.} \end{cases}$$

For all $x_N, x'_N \in X$, $x_N \, R_{\mathrm{CD}}(e) \, x'_N$ if and only if

$$-\left(m_{ij}(x_N)\right)_{i,j \in N} \geq_{\mathrm{lex}} -\left(m_{ij}(x'_N)\right)_{i,j \in N}.$$

The following statement establishes the connection between this criterion and the compensation principle.

Proposition 4. *The allocation rule derived from the compensation diversity criterion exactly selects the set of envy-free allocations whenever it is nonempty (on $\mathcal{D}_2$ for the distribution case). The allocation rule derived from it satisfies equal well-being for equal responsibility (on $\mathcal{D}_2$ for the distribution case) and equal treatment for uniform circumstances.*

Proof. (1) Notice that $x_{ii}(x_N) \equiv x_i$. When i envies j, one has

$$u(x_i, y_i, z_i) < u(x_j, y_j, z_i),$$

implying that if

$$u_i = u(x_{ij}(x_N), y_j, z_i),$$

as is always obtained in the TU case and on $\mathcal{D}_2$ in the distribution case, then $x_{ij}(x_N) < x_j = x_{jj}(x_N)$ and therefore $m_{ij}(x_N) > 0$. An envy-free allocation is such that $\left(m_{ij}(x_N)\right)_{i,j \in N} = 0$, and this dominates any $-\left(m_{ij}(x'_N)\right)_{i,j \in N} < 0$ for the leximin criterion. Therefore, the set of envy-free allocations is selected whenever it is nonempty.

(2) When $y_i = y_j$ for all $i, j \in N$, the allocation $x_N = 0$ is the only envy-free efficient allocation and is therefore selected. This proves the satisfaction of equal treatment for uniform circumstances.

(3) That it satisfies equal well-being for equal responsibility is a consequence of the fact that $m_{ij}(x_N) = |N|$ if $u_i < u_j$ while $z_i = z_j$ and that, in the TU case as well as in the distribution case on $\mathcal{D}_2$, by a dual to Lemma 1, there always exist allocations with $\max\left(m_{ij}(x_N)\right)_{i,j \in N} < |N|$. □

An interesting difference between diversity and compensation diversity is worth noting. When i envies j, this is recorded by the diversity criterion on the basis of z_i, that is, the preferences of the *envious* agent:

$$u(x_i, y_i, z_k) < u(x_j, y_j, z_k) \text{ for } k = i,$$

whereas with compensation diversity, this is recorded with y_j, that is, the circumstances of the *envied* agent:

$$u(x_i, y_i, z_i) = u(x_{ik}(x_N), y_k, z_i) < u(x_j, y_j, z_i) = u(x_{jk}(x_N), y_k, z_i) \text{ for } k = j.$$

As in the previous section with the envious count and envied count criteria, whether one focuses on the envious or on the envied may contribute to determining whether the criterion falls on the compensation side or on the natural reward side. A similar configuration will again be obtained in the next section.

5 From Envy Intensity to Walras

Although the diversity criteria improve on the envy count criteria of Section 3, they may still be criticized for the restricted information they rely upon. They rank allocations on the basis of a rather poor information, namely, the graphs of envy relations (and of similar preference relations for the diversity criteria). Allocations are made of distributions of resources, which provide a much finer scale for the measurement of envy situations. It is quite unjustifiable to ignore this information and simply focus on zero-one markers of presence or absence of envy relations. In particular, the envy count and diversity criteria are indifferent between any pair of allocations with the same graph, even if one allocation may have much less inequality, that is, a smaller degree of envy, than the other. They may also prefer an allocation with fewer relations of envy but with a very high degree of envy in these relations to another allocation with more envy occurrences but which is in fact much closer to an envy-free situation. In conclusion, looking at the graphs of envy relations, even augmented by n_{ij} numbers, is probably not a very good idea.

The S_{MEI} allocation is based on a finer information and indeed suggests an alternative approach. For every allocation and for every pair of individuals (i, j), compute the number t_{ij} as the smallest amount of external resources such that giving this to i in addition to what he receives in this allocation would prevent him from envying j. If i already does not envy j, this number is typically negative, meaning that one can diminish i's resources without making him envy j. And one always has $t_{ii} = 0$. Let t_{ij} be called the degree of i's envy toward j. The S_{MEI} allocation rule as defined above amounts to retaining the greatest t_{ij} for every i (ignoring t_{ii}) as a measure of his greatest degree of envy (or smallest degree of non-envy if it is negative), and to apply the minimax criterion to the vector of these numbers.[4] This is a rather natural solution, but Fleurbaey (1994, Prop. 9) notices that it satisfies neither equal well-being for equal responsibility nor equal treatment for equal circumstances. This suggests looking for another way to rank distributions $(t_{ij})_{i,j\in N}$.

We examine two other, a priori less intuitive, options which may ultimately be more satisfactory. The first is similar to the above but incorporates t_{ii} in the computation of the greatest degree of envy, so that this number is always nonnegative, and applies the summation operation rather than the minimax. The second solution computes, for every individual, the greatest degree of envy among those who might envy him, and then applies the summation operator. In both cases, the social objective is to minimize the value of these sums. The first is focused on the degree of envy from the standpoint of the envious (the transmitters), while the second takes the viewpoint of those who are envied (the receivers).

For any allocation $x_N \in X$, any pair of agents $i, j \in N$, let $t_{ij}(x_N)$ be the smallest value of t such that

$$u(x_i + t, y_i, z_i) \geq u(x_j, y_j, z_i)$$

and $d_{ij}(x_N)$ be the smallest value of d such that

$$u(x_i, y_i, z_i) \geq u(x_j - d, y_j, z_i).$$

Tadenuma and Thomson (1995), in the context of transferable indivisibles, have considered the two notions of t_{ij} and d_{ij}. The definition of $d_{ij}(x_N)$ should be slightly modified in the distribution case when

$$u(x_i, y_i, z_i) < u(0, y_j, z_i),$$

in which case one can propose to compute $d_{ij}(x_N)$ as the smallest value of $x_j + d$ for d such that

$$u(x_i + d, y_i, z_i) \geq u(0, y_j, z_i).$$

In the TU case, one simply has

$$t_{ij}(x_N) = d_{ij}(x_N) = x_j + v(y_j, z_i) - x_i - v(y_i, z_i).$$

In the distribution case, one may have $t_{ij}(x_N)$ or $d_{ij}(x_N)$ undefined if

$$\lim_{x\to+\infty} u(x, y_i, z_i) < u(x_j, y_j, z_i) \text{ or } u(x_i, y_i, z_i) > \lim_{x\to+\infty} u(x, y_j, z_i).$$

---

[4] In fact, the presentation of S_{MEI} in Sect. 2 was a little simplistic since Fleurbaey (1994) already introduces the t_{ij} and applies the reverse leximin criterion to the vector $(\max_{j\neq i} t_{ij})_{i\in N}$, instead of the minimax.

To avoid this problem, we may restrict our attention to the domain $\mathcal{D}_0$ such that for all $i, j \in N$, there exists $x \geq 0$ such that

$$u(x, y_i, z_i) \geq u(\Omega, y_j, z_i) \text{ and } u(x, y_j, z_i) \geq u(\Omega, y_i, z_i).$$

In all cases, the following three statements are equivalent: (1) $t_{ij}(x_N) > 0$; (2) $d_{ij}(x_N) > 0$; and (3) i envies j. One always has $t_{ii}(x_N) \equiv d_{ii}(x_N) \equiv 0$.

The two social ordering functions are formally defined as follows.

Envious Intensity ($R_{\mathbf{EsI}}$): For all $x_N, x'_N \in X$, $x_N \, R_{\mathrm{EsI}}(e) \, x'_N$ if and only if

$$\sum_{i \in N} \max_{j \in N} t_{ij}(x_N) \leq \sum_{i \in N} \max_{j \in N} t_{ij}(x'_N).$$

Envied Intensity ($R_{\mathbf{EdI}}$): For all $x_N, x'_N \in X$, $x_N \, R_{\mathrm{EdI}}(e) \, x'_N$ if and only if

$$\sum_{j \in N} \max_{i \in N} d_{ij}(x_N) \leq \sum_{j \in N} \max_{i \in N} d_{ij}(x'_N).$$

The quantity $\max_{j \in N} t_{ij}(x_N)$ measures how much must be added to x_i for envious i to get rid of envy, while $\max_{i \in N} d_{ij}(x_N)$ measures how much must be deducted from x_j for envied j not to be envied any more.

The allocation rules derived from these rankings appear to have more interesting properties than S_{MEI}.

Proposition 5. *The allocation rules derived from the envious and envied intensity criteria both exactly select the set of envy-free allocations whenever it is nonempty (on $\mathcal{D}_0$ for the distribution case). The allocation rule derived from envious intensity satisfies equal well-being for uniform responsibility (on $\mathcal{D}_1$ for the distribution case) and equal treatment for equal circumstances (on $\mathcal{D}_0$ for the distribution case). The allocation rule derived from envied intensity satisfies equal well-being for equal responsibility (on $\mathcal{D}_2$ for the distribution case) and equal treatment for uniform circumstances.*

Proof. (1) For all $x_N \in X$, all $i \in N$,

$$\max_{j \in N} t_{ij}(x_N) \geq t_{ii}(x_N) \equiv 0,$$

and $\max_{j \in N} t_{ij}(x_N) > 0$ if and only if i is envious, so that one has

$$\sum_{i \in N} \max_{j \in N} t_{ij}(x_N) = 0$$

if and only if x_N is envy-free. The same can be said about $\sum_{j \in N} \max_{i \in N} d_{ij}(x_N)$. Therefore, the set of envy-free allocations is selected by either allocation rule whenever it is nonempty.

(2) By the same argument as in Proposition 3, step 2 (respectively, Proposition 4, step 2), this implies that the allocation rule derived from envious intensity

(respectively, envied intensity) satisfies equal well-being for uniform responsibility (respectively, equal treatment for uniform circumstances).

(3) Now let us turn to envious intensity and equal treatment for equal circumstances. Consider two agents i, j such that $y_i = y_j$, and suppose, by way of contradiction, that there is an allocation x_N minimizing $\sum_{k\in N} \max_{l\in N} t_{kl}(x_N)$, with $x_i > x_j$. The fact that $x_i > x_j$ implies that $t_{ji}(x_N) > 0$ and that, for all $k \neq j$,

$$\max_{l\in N} t_{kl}(x_N) \geq t_{ki}(x_N) > t_{kj}(x_N).$$

Take δ such that

$$0 < \delta < \frac{1}{|N|} \min_{k\in N} \left(\max_{l\in N} t_{kl}(x_N) - t_{kj}(x_N) \right)$$

and $\delta < (x_i - x_j) / |N|$. Construct a new allocation such that $x'_k = x_k - \delta$ for all $k \neq j$, and $x'_j = x_j + (|N| - 1)\,\delta$. Notice that one still has $x'_i > x'_j$ and therefore, for all $k \neq j$,

$$\max_{l\in N} t_{kl}(x'_N) > t_{kj}(x'_N).$$

Consider $k, l \neq j$. One has

$$u(x_k + t_{kl}(x_N), y_k, z_k) \geq u(x_l, y_l, z_k).$$

One also has either

$$u(x_k - \delta + t_{kl}(x'_N), y_k, z_k) = u(x_l - \delta, y_l, z_k) < u(x_l, y_l, z_k),$$

implying $t_{kl}(x_N) > -\delta + t_{kl}(x'_N)$, or $t_{kl}(x'_N) = -(x_k - \delta)$, implying $t_{kl}(x_N) \geq -\delta + t_{kl}(x'_N)$ since one always has $t_{kl}(x_N) \geq -x_k$. One therefore has

$$\max_{l\in N} t_{kl}(x'_N) \leq \max_{l\in N} t_{kl}(x_N) + \delta$$

for all $k \neq j$.

Now consider j and $k \neq j$ envied by j in x'_N (at least i is envied by j). One has

$$u(x_j + (|N| - 1)\,\delta + t_{jk}(x'_N), y_j, z_j) = u(x_k - \delta, y_k, z_j) < u(x_k, y_k, z_j)$$

and

$$u(x_k, y_k, z_j) \leq u(x_j + t_{jk}(x_N), y_j, z_j),$$

implying

$$(|N| - 1)\,\delta + t_{jk}(x'_N) < t_{jk}(x_N).$$

For any $k \neq j$ that is not envied by j in x'_N, one has

$$t_{jk}(x'_N) \leq 0 < x_i - x_j - |N|\,\delta \leq \max_{l\in N} t_{jl}(x_N) - |N|\,\delta.$$

Therefore,

$$\max_{l\in N} t_{jl}(x'_N) < \max_{l\in N} t_{jl}(x_N) - (|N|-1)\,\delta.$$

Summing up over all agents, one obtains

$$\sum_{k\in N} \max_{l\in N} t_{kl}(x'_N) < \sum_{k\in N} \max_{l\in N} t_{kl}(x_N) + (|N|-1)\,\delta - (|N|-1)\,\delta,$$

contradicting the assumption that x_N minimizes $\sum_{i\in N} \max_{j\in N} t_{ij}(x_N)$.

(4) Finally, envied intensity and equal well-being for equal responsibility. Consider two agents i,j such that $z_i = z_j$, and suppose, by way of contradiction, that there is an allocation x_N minimizing $\sum_{k\in N} \max_{l\in N} d_{lk}(x_N)$, with $u_i > u_j$. The fact that $u_i > u_j$ implies that $d_{ji}(x_N) > 0$ and that, for all $k \in N$,

$$\max_{l\in N} d_{lk}(x_N) \geq d_{jk}(x_N) > d_{ik}(x_N).$$

Take $\delta > 0$ such that

$$u(x_i - (|N|-1)\,\delta, y_i, z_i) > u(x_j + \delta, y_j, z_j).$$

Construct a new allocation such that $x'_i = x_i - (|N|-1)\,\delta$ and for all $k \neq i$, $x'_k = x_k + \delta$. One still has $u'_i > u'_j$ and therefore, for all $k \in N$,

$$\max_{l\in N} d_{lk}(x'_N) > d_{ik}(x'_N).$$

Consider $k,l \neq i$. One has (in the TU case as well as in the distribution case for the domain $\mathcal{D}_2$)

$$u(x_l, y_l, z_l) = u(x_k - d_{lk}(x_N), y_k, z_l),$$

and

$$u(x_l + \delta, y_l, z_l) = u(x_k + \delta - d_{lk}(x'_N), y_k, z_l) > u(x_l, y_l, z_l),$$

implying $\delta - d_{lk}(x'_N) > -d_{lk}(x_N)$, that is, $d_{lk}(x'_N) < d_{lk}(x_N) + \delta$. One therefore has

$$\max_{l\in N} d_{lk}(x'_N) \leq \max_{l\in N} d_{lk}(x_N) + \delta$$

for all $k \neq i$.

Now consider i and $k \neq i$. One has

$$u(x_k + \delta, y_k, z_k) = u(x_i - (|N|-1)\,\delta - d_{ki}(x'_N), y_i, z_k) > u(x_k, y_k, z_k)$$

and

$$u(x_k, y_k, z_k) = u(x_i - d_{ki}(x_N), y_i, z_k),$$

implying

$$d_{ki}(x'_N) < d_{ki}(x_N) - (|N|-1)\,\delta.$$

Therefore,

$$\max_{l\in N} d_{li}(x'_N) < \max_{l\in N} d_{li}(x_N) - (|N|-1)\,\delta.$$

Summing up over all agents, one obtains

$$\sum_{k\in N} \max_{l\in N} d_{lk}(x'_N) < \sum_{k\in N} \max_{l\in N} d_{lk}(x_N) + (|N|-1)\,\delta - (|N|-1)\,\delta,$$

contradicting the assumption that x_N minimizes $\sum_{k\in N} \max_{l\in N} d_{lk}(x_N)$. □

The envied intensity criterion is a little mysterious because it is not written in terms of indices of personal situations (being envied is not a characteristic of one's situation but rather a token of the others' situations), contrary to the envious intensity that transparently measures how envious every agent is and constructs a synthetic measure of this. The envied intensity criterion, however, can be related to a more orthodox social ordering function. Let $(q_i)_{i\in N} \in \mathbb{R}^N$ be a vector of prices for y_N, in a virtual market in which agents could buy bundles (x,y). The budget constraint for $i \in N$ on this market is such that a bundle (x,y_j) is affordable if

$$x + q_j = I_i,$$

where I_i denotes i's personal wealth. Let e_i denote i's expenditure function:

$$e_i(u_i, q_N) = \min\left\{x + q_j \mid (x,j) \in \mathbb{R}\times N \text{ and } u(x,y_j,z_i) \geq u_i\right\}.$$

We now define the Egalitarian Walras social ordering function. Let Q be the subset of q_N such that $\sum_{i\in N} q_i = 0$.

Egalitarian Walras ($\boldsymbol{R}_{\mathbf{EW}}$): For all $x_N, x'_N \in X$, $x_N\, R_{\mathrm{EW}}(e)\, x'_N$ if and only if

$$\max_{q_N\in Q} \min_{i\in N} e_i(u(x_i,y_i,z_i), q_N) \geq \max_{q_N\in Q} \min_{i\in N} e_i(u(x'_i,y_i,z_i), q_N).$$

This social ordering function is an adaptation to this model of a function introduced in Fleurbaey and Maniquet (2008) for the fair division context in order to rationalize the egalitarian competitive equilibrium. Notice that this social ordering function can be used to rank all allocations, not just the feasible ones.

Consider the virtual market in which circumstances y are tradable. This is just the model of allocation of large indivisibles as studied, for example, in Svensson (1983), with a number of indivisible goods equal to the number of agents. One can easily extend the definition of the above social ordering function in order to consider possibilities of permutations in y_N : $(x_N,y_N)\; R_{\mathrm{EW}}(e)\; (x'_N,y'_N)$ if and only if

$$\max_{q_N\in Q} \min_{i\in N} e_i(u(x_i,y_i,z_i), q_N) \geq \max_{q_N\in Q} \min_{i\in N} e_i(u(x'_i,y'_i,z_i), q_N).$$

In this market a competitive equilibrium is an allocation (x_N,y_N) associated to a price vector q such that for all $i \in N$, (x_i,y_i) maximizes $u(x,y_j,z_i)$ over the set of bundles (x,y_j) satisfying the budget constraint $x + q_j = I_i$. It is egalitarian if $I_i = I_j$ for all $i,j \in N$.

Let $\Pi(y_N)$ denote the set of permutations of y_N and let $(x_N, y_N, q_N) \in X \times \Pi(y_N) \times Q$ be any allocation and price vector. If for all $i \in N$,

$$x_i + q_i = e_i(u(x_i, y_i, z_i), q_N),$$

then this is a competitive equilibrium. More generally, one always has

$$e_i(u(x_i, y_i, z_i), q_N) \leq x_i + q_i$$

for all $i \in N$, with at least one strict inequality if this is not a competitive equilibrium. By construction one has

$$\sum_{i \in N} (x_i + q_i) = \Omega,$$

implying that one always has

$$\max_{q_N \in Q} \min_{i \in N} e_i(u(x_i, y_i, z_i), q_N) \leq \frac{\Omega}{|N|},$$

with equality if and only if (x_N, y_N, q_N) is an egalitarian competitive equilibrium. This shows that the Egalitarian Walras social ordering function rationalizes the egalitarian competitive equilibrium (in the sense that it exactly selects the set of egalitarian equilibria whenever it is non-empty), which, in the particular context of indivisibles, coincides with the set of envy-free and efficient allocations (Svensson, 1983).

Let us now focus on the distribution case for the domain $\mathcal{D}_2$ and on the TU case. In these two cases one can simply define $d_{ij}(x_N)$ by the equation

$$u(x_j - d_{ij}(x_N), y_j, z_i) = u_i.$$

One then computes

$$e_i(u(x_i, y_i, z_i), q_N) = \min_{j \in N} (x_j - d_{ij}(x_N) + q_j).$$

One therefore has

$$\begin{aligned} \min_{i \in N} e_i(u(x_i, y_i, z_i), q_N) &= \min_{i,j \in N} (x_j - d_{ij}(x_N) + q_j) \\ &= \min_{j \in N} \left(x_j + q_j - \max_{i \in N} d_{ij}(x_N) \right), \end{aligned}$$

implying that for all $x_N \in X$, $q_N \in Q$,

$$\min_{i \in N} e_i(u(x_i, y_i, z_i), q_N) \leq \frac{1}{|N|} \sum_{j \in N} \left(x_j + q_j - \max_{i \in N} d_{ij}(x_N) \right).$$

Since $\sum_{j \in N} (x_j + q_j) = \Omega$, this can be simplified into

$$\min_{i\in N} e_i(u(x_i,y_i,z_i),q_N) \le \frac{\Omega}{|N|} - \frac{1}{|N|}\sum_{j\in N}\max_{i\in N} d_{ij}(x_N).$$

Now let, for all $j \in N$,

$$q_j = -\left(x_j - \max_{i\in N} d_{ij}(x_N)\right) + \frac{\Omega}{|N|} - \frac{1}{|N|}\sum_{k\in N}\max_{i\in N} d_{ik}(x_N).$$

By construction $q_N \in Q$. Moreover, for all $j \in N$,

$$x_j + q_j - \max_{i\in N} d_{ij}(x_N) = \frac{\Omega}{|N|} - \frac{1}{|N|}\sum_{k\in N}\max_{i\in N} d_{ik}(x_N),$$

implying that

$$\begin{aligned}\min_{i\in N} e_i(u(x_i,y_i,z_i),q_N) &= \min_{j\in N}\left(x_j + q_j - \max_{i\in N} d_{ij}(x_N)\right)\\ &= \frac{\Omega}{|N|} - \frac{1}{|N|}\sum_{k\in N}\max_{i\in N} d_{ik}(x_N).\end{aligned}$$

Since we have seen above that this is an upper bound for $\min_{i\in N} e_i(u(x_i,y_i,z_i),q_N)$ when q_N varies, one actually has

$$\max_{q_N\in Q}\min_{i\in N} e_i(u(x_i,y_i,z_i),q_N) = \frac{\Omega}{|N|} - \frac{1}{|N|}\sum_{j\in N}\max_{i\in N} d_{ij}(x_N),$$

which establishes an exact equivalence (for a fixed value of Ω) between the egalitarian Walras and the envied intensity criteria in the TU case and on the domain $\mathcal{D}_2$ for the distribution case. It is somewhat surprising that a maximin social ordering function can be equivalent to a purely additive criterion. But one observes in the above computations that the maximin criterion of the egalitarian Walras function is what triggers the focus on the maximal intensity retained in the computation of $\max_{i\in N} d_{ij}(x_N)$ for the envied intensity criterion.

Proposition 6. *The egalitarian Walras and the envied intensity criteria are equivalent in the TU case. In the distribution case, they are equivalent on the domain* $\mathcal{D}_2$.

Regarding envious intensity, one can similarly establish an equivalence with the social ordering function that evaluates an allocation x_N by computing

$$\max_{q\in Q}\min_{i\in N}(x_i + q_i - \max_{j\in N} t_{ij}(x_N)).$$

This amounts to measuring the agents' wealth $x_i + q_i$ and deducting from it their maximal degree of envy. One notices here that the duality between compensation and natural reward appears to be related to the duality of consumer theory. Applied to the market for (transferable) indivisibles, this social ordering function also rationalizes the equal competitive equilibrium. Contrary to egalitarian Walras, it does not satisfy the Pareto principle in that context and therefore appears less interesting.

6 Conclusion

Six main no-envy rankings have been examined in this chapter, in the context of the compensation problem. Apart from the intrinsic interest of these solutions, which still deserves further assessment, the main conceptual insight obtained here may be that the well-know duality in the compensation problem between the compensation principle and the natural reward principle is related to a duality between focussing on the envied and focussing on the envious. But this connection is not simple since, for instance, the compensation principle is satisfied by the envious count criterion and the envied intensity criterion, that is, two criteria focussing on a different side of the envy relation. The key observation underlying these two facts is that when $z_i = z_j$ and $u_i > u_j$, agent j envies all the agents envied by i with greater intensity than i, implying that, for all $k \neq i$, i is never such that $d_{ik} = \max_{l \in N} d_{lk}$, and that, since j also envies i in top of the others envied by i, one has $n_j > n_i$.

A gap that this chapter may highlight is that there is a lack of axiomatic framework for the study of social ordering functions in the compensation problem. The evaluation of rankings that has been performed here was concerned with satisfying axioms of allocation rules and therefore focused on the subsets selected by the contemplated rankings. It is not very difficult to formulate axioms for rankings that bear a close relation to the axioms presented here. For instance, one can think of the following variants of the above axioms, applying to a social ordering function R:

Transfer for Equal Responsibility: $\forall e \in \mathcal{D}$, $\forall x_N, x'_N \in X$, $\forall i, j \in N$ such that $z_i = z_j$, if $x'_i - x_i = x_j - x'_j$ and

$$u\left(x'_i, y_i, z_i\right) > u\left(x_i, y_i, z_i\right) \geq u\left(x_j, y_j, z_j\right) > u\left(x'_j, y_j, z_j\right)$$

while $x'_k = x_k$ for all $k \neq i, j$, then $x_N\, P(e)\, x'_N$.

Transfer for Equal Circumstances: $\forall e \in \mathcal{D}$, $\forall x_N, x'_N \in X$, $\forall i, j \in N$ such that $y_i = y_j$, if $x'_i - x_i = x_j - x'_j$ and

$$x'_i > x_i \geq x_j > x'_j$$

while $x'_k = x_k$ for all $k \neq i, j$, then $x_N\, P(e)\, x'_N$.

Of all the social ordering functions studied in this chapter, only envious intensity and envied intensity come close to satisfying such axioms – and this illustrates again the usefulness of a fine measure of the degree of envy – but they satisfy only weak versions of the axioms, involving a weak preference $x_N\, R(e)\, x'_N$. Some leximin version of these two rankings should be invented in order to cope with this problem. This is rather easy for the envied intensity criterion, for which a maximin interpretation (egalitarian Walras) has already been provided. For envious intensity, the solution is less obvious, and in particular the maximin ranking underlying S_{MEI} appears to be of no help in this matter. A more systematic study of these issues is left for another occasion.

Acknowledgments I am grateful to Y. Sprumont, K. Suzumura, K. Tadenuma, and a referee for very helpful comments and to the audience at the Conference in Honor of K. Suzumura for useful reactions.

References

Bossert, W. (1995). Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, *17*, 1–17

Chauduri, A. (1986). Some implications of an intensity measure of envy. *Social Choice and Welfare*, *3*, 255–270

Daniel, T. E. (1975). A revised concept of distributional equity. *Journal of Economic Theory*, *11*, 94–109

Diamantaras, D., & Thomson, W. (1990). A refinement and extension of the no-envy concept. *Economics Letters*, *33*, 217–222

Feldman, A. M., & Kirman, A. (1974). Fairness and envy. *American Economic Review*, *64*, 996–1005

Fleurbaey, M. (1994). On fair compensation. *Theory and Decision*, *36*, 277–307

Fleurbaey, M. (1995). Equality and responsibility. *European Economic Review*, *39*, 683–689

Fleurbaey, M., & Maniquet, F. (1996). Fair allocation with unequal production skills: The no-envy approach to compensation. *Mathematical Social Sciences*, *32*, 71–93

Fleurbaey, M., & Maniquet, F. (2008). Utilitarianism versus fairness in welfare economics. In M. Fleurbaey, M. Salles, & J. A. Weymark (Eds.), *Justice, political liberalism and utilitarianism: Themes from harsanyi and rawls.* Cambridge: Cambridge University Press, 263–280

Fleurbaey, M., & Maniquet, F. (2002). Compensation and responsibility. Forcoming in K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare, vol. 2.* Amsterdam: North-Holland

Foley, D. (1967). Resource allocation and the public sector. *Yale Economic Essays*, *7*, 45–98

Iturbe-Ormaetxe, I., & Nieto, J. (1996). On fair allocations and monetary compensations. *Economic Theory*, *7*, 125–138

Kolm, S. C. (1972). *Justice et equité*. Paris: Editions du CNRS

Maskin, E.(1987). On the fair allocation of indivisible goods. In G. R. Feiwel (Ed.), *Arrow and the foundations of the theory of economic policy* (pp. 341–349). London: Macmillan

Parijs, P. V. (1990). Equal endowments as undominated diversity. *Recherches Economiques de Louvain*, *56*, 327–355

Parijs, P. V. (1995). *Real freedom for all*. Oxford: Oxford University Press

Pazner, E., & Schmeidler, D. (1974). A difficulty in the concept of fairness. *Review of Economic Studies*, *41*, 441–443

Suzumura, K.(1981a). On Pareto-efficiency and the no-envy concept of equity. *Journal of Economic Theory*, *25*, 367–379

Suzumura, K. (1981b). On the possibility of “fair” collective choice rule. *International Economic Review*, *22*, 351–364

Suzumura, K. (1983). *Rational choice, collective decisions, and social welfare*. Cambridge: Cambridge University Press.

Svensson, L. G. (1983). Large indivisibles: An analysis with respect to price equilibrium and fairness. *Econometrica*, *51*, 939–954

Tadenuma, K. (2002). Efficiency first or equity first? Two principles and rationality of social choice. *Journal of Economic Theory*, *104*, 462–472

Tadenuma, K., & Thomson, W. (1995). Refinements of the no-envy solution in economies with indivisible goods. *Theory and Decision*, *39*, 189–206

Varian, H. (1976). Two problems in the theory of fairness. *Journal of Public Economics*, *5*, 249–260

Choice-Consistent Resolutions of the Efficiency-Equity Trade-Off

Koichi Tadenuma

1 Introduction

It is not rare that multiple criteria are applied to make individual or social decisions. In the context of resource allocation problems, most prominent criteria are efficiency and equity of allocations. *Pareto efficiency* is probably the most widely accepted criterion among economists, but it is silent about the distributional equity of allocations. On the other hand, several concepts of equity have been proposed and extensively studied in welfare economics. Two of them are central: *no-envy* (Foley, 1967 and Kolm, 1972) and *egalitarian-equivalence* (Pazner and Schmeidler, 1978). We say that an allocation is *envy-free* if no agent prefers the consumption bundle of any other agent to his own and that an allocation is *egalitarian-equivalent* if there is a consumption bundle, called the *reference bundle*, such that every agent is indifferent between the bundle and his own.

However, as Kolm (1972) pointed out, there is a fundamental conflict between the Pareto efficiency criterion and the no-envy criterion. There often exist two allocations x and y such that x is Pareto superior to y, whereas x is not envy-free but y is. If these two allocations $\{x,y\}$ are the only policy options available at the time, we cannot attain an allocation that is *both* Pareto efficient in $\{x,y\}$ *and* envy-free, but we have to choose *either* the efficient allocation *or* the envy-free allocation. The same kind of conflict also arises between Pareto efficiency and egalitarian-equivalence, as shown by Tadenuma (2005).

If two criteria of decision-making are incompatible simultaneously, we have to give priority to one criterion. That is, we take one criterion as the first and the other as the second, and apply them in the lexicographic order. This chapter formalizes the idea in a standard framework of social choice correspondences in economic environments. Given the first and the second criterion, we require that choice should

K. Tadenuma
Department of Economics, Hitotsubashi University, Kunitachi, Tokyo 186-8601, Japan
e-mail: koichi.t@econ.hit-u.ac.jp

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

always be made from the allocations satisfying the first criterion whenever there are any. The second criterion should then be applied when the first criterion is not at all effective as a guide for selection, namely, either when *all* the available allocations satisfy the first criterion or when there is *no* such allocation at all. A stronger version of this condition may be obtained by requiring that *all* the allocations satisfying the second criterion should be selected in the latter case.

Besides the socially desirable properties of selected allocations, another important requirement for social choice correspondences is *choice-consistency*. Especially, *path independence* is crucial. It implies "the independence of the final choice from the path to it" (Arrow, 1963, p. 120). Path independence is an indispensable property of social choice rules. Were it violated, some arbitrary agenda controls could affect the final choice, which is clearly undesirable.

Another natural choice-consistency condition, which is weaker than path independence, is *contraction consistency*. This says that if an allocation is chosen from a set S of available allocations, then it should also be chosen from any subset T of S as long as it is still available.

The purpose of this chapter is to examine possibility of consistent choices under the efficiency-first and equity-second principle or the equity-first and efficiency-second principle. We show several impossibility theorems on the existence of social choice correspondences satisfying the efficiency-first and equity-second principle with the concepts of no-envy and egalitarian-equivalence, and contraction consistency. However, if we restrict the range of reference bundles for egalitarian-equivalence to a fixed ray from the origin, then there exists a social choice correspondence satisfying the efficiency-first and equity-second-as-egalitarian-equivalence principle and path independence. But even for this case, the stronger versions of properties representing the principle is incompatible with path independence.

Turning to the equity-first and efficiency-second principle, we also obtain impossibility and possibility results on the existence of social choice correspondences satisfying the principle and choice-consistency properties. It turns out that the borderline between possibility and impossibility is quite subtle. If our equity criterion selects only allocations with no-envy at all, then there exists a social choice correspondence satisfying the equity-first principle and path independence. Moreover, we obtain a characterization of the social choice correspondence by using the stronger versions of properties representing the principle and path independence. However, if we select allocations with "minimal-envy" according to the measure of envy-instances introduced by Suzumura (1996), which is based on the *set-inclusions of envy relations*, then no social choice correspondence satisfies the equity-first principle and contraction consistency together. In contrast, if the equity criterion selects allocations with "least-envy" in the sense that the *number* of envy-instances is the smallest (Feldman and Kirman, 1974), then compatibility with path independence is retained.

The rest of this chapter is organized as follows. Section 2 defines basic concepts and notation, and Section 3 introduces various properties of social choice correspondences that represent the efficiency-first principle or the equity-first principle, and choice-consistency. In Section 4, we review the fundamental conflict between efficiency and equity. Section 5 examines choice-consistency of the social choice correspondences satisfying the efficiency-first principle, and Section 6 turns to the equity-first principle. In Section 7, we study the equity-first principle with the notion of minimal-envy. Section 8 contains concluding remarks.

2 Basic Definitions and Notation

There are n agents and m infinitely divisible goods, where n and m are some integers with $n,m \geq 2$. Let $N = \{1,\ldots,n\}$ be the set of agents. Denoting by $\mathbb{R}$ the set of real numbers, the set $\mathbb{R}^m_+$ is the consumption set of each agent. Let $\mathcal{R}$ be the class of preference relations on $\mathbb{R}^m_+$ that are reflexive, transitive, complete, continuous, and monotonic. Each agent $i \in N$ is endowed with a preference relation $R_i \in \mathcal{R}$. The strict preference relation and the indifference relation of agent i are denoted by P_i and I_i, respectively. A list of preference relations, $(R_i)_{i \in N} \in \mathcal{R}^n$, is called a preference profile, and denoted by R_N.

An *allocation* is a vector $x = (x_1,\ldots,x_n) \in \mathbb{R}^{mn}_+$ where each $x_i = (x_{i1},\ldots,x_{im}) \in \mathbb{R}^m_+$ is a consumption bundle of agent $i \in N$. The set of all allocations is denoted by X. We set $X = \mathbb{R}^{mn}_+$ for simplicity of presentation. However, we might impose a resource constraint on X. For example, given a total amount of resources $\Omega \in \mathbb{R}^m_+$, define the set of all feasible allocations with no free disposal as $X := \{x \in \mathbb{R}^{mn}_+ \mid \sum_{i=1}^n x_i = \Omega\}$. All the results in this chapter hold on this more restricted set of allocations. Let $\mathcal{S}$ be the set of all non-empty finite subsets of X.

Let a preference profile $R_N \in \mathcal{R}^n$ be given. An allocation $x \in X$ is *weakly Pareto superior to* an allocation $y \in X$ for R_N if $x_i\ R_i\ y_i$ for all $i \in N$. We write $x \succsim_{P(R_N)} y$ if x is weakly Pareto superior to y. Let $\succ_{P(R_N)}$ be the strict part of $\succsim_{P(R_N)}$.[1] An allocation $x \in X$ is *Pareto superior to* an allocation $y \in X$ for R_N if $x \succ_{P(R_N)} y$. For each $S \in \mathcal{S}$, an allocation $x \in S$ is *Pareto efficient in S for R_N* if there is no allocation $y \in S$ such that $y \succ_{P(R_N)} x$. Let $P(R_N,S)$ be the set of Pareto efficient allocations in S for R_N.

An allocation $x \in X$ is *envy-free for R_N* if $x_i\ R_i\ x_j$ for all $i,j \in N$. For each $S \in \mathcal{S}$, let $F(R_N,S)$ be the set of envy-free allocations in S for R_N.[2] An allocation $x \in X$ is *egalitarian-equivalent for R_N* if there is a consumption bundle $x_0 \in \mathbb{R}^m_+$ such that for all $i \in N$, $x_i\ I_i\ x_0$. Then, the bundle x_0 is called a *reference bundle for x*. For each $S \in \mathcal{S}$, let $E(R_N,S)$ be the set of egalitarian-equivalent allocations in S for R_N. Particular subclasses of egalitarian-equivalent allocations are often studied in the literature. Let $\bar{r} \in \mathbb{R}^m_{++}$ be a given vector. An allocation $x \in X$ is *egalitarian-equivalent for a fixed reference ray with $\bar{r}$ for R_N* or simply *$\bar{r}$-egalitarian-equivalent for R_N* if there is a real number $t \in \mathbb{R}$ such that for all $i \in N$, $x_i\ I_i\ t\bar{r}$. For each $S \in \mathcal{S}$, let $E_{\bar{r}}(R_N,S)$ be the set of $\bar{r}$-egalitarian-equivalent allocations in S for R_N.

---

[1] Given a binary relation $\succsim$, its strict part $\succ$ is defined as $x \succ y \Leftrightarrow x \succsim y$ and $y \not\succsim x$.

[2] The capital letter **F** stands for **F**reedom from envy.

A *social choice correspondence* is a set-valued function $\varphi : \mathcal{R}^n \times \mathcal{S} \to \mathcal{S}$ such that $\varphi(R_N, S) \subseteq S$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$. A social choice correspondence is interpreted as follows. Each $S \in \mathcal{S}$ is the set of available allocations, which may be termed an *environment* following (Arrow, 1963, p. 15). Then $\varphi(R_N, S)$ is the set of socially desirable allocations in the given environment S when the preferences of the agents are R_N. A fundamental example of a social choice correspondence is the *Pareto correspondence*, denoted by P, which associates with each $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$ the set of all Pareto efficient allocations in S for R_N. The *no-envy correspondence*, the *egalitarian-equivalence correspondence*, and the *$\bar{r}$-egalitarian-equivalence correspondence*, denoted by F, E, and $E_{\bar{r}}$, respectively, can be defined analogously.

A remark should be in order on the domain of social choice correspondences. The domain consists of the *preferences domain* $\mathcal{R}$ and the *alternatives domain* $\mathcal{S}$. As in many contributions in the literature of social choice theory, we assume that $\mathcal{S}$ is the class of all finite subsets of X. Our major interest here is not in investigating what are "optimal" allocations *in the set of all technologically feasible allocations*. There are many situations in which only a finite number of policy options are at issue at any one time. In such situations, we are rather interested in examining "consistency" of social choices at different times, or under expansions, contractions, or partitions of alternatives available at hand. To that end, our choice of $\mathcal{S}$ would be appropriate.

3 The Axioms

This section introduces a variety of desirable properties of social choice correspondences, which we call "axioms." In the rest of the chapter, we denote by φ a social choice correspondence.

The first axiom is familiar. It means that we should always select from Pareto efficient allocations whenever they exist.[3]

Pareto Efficiency. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if $P(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq P(R_N, S)$.

The next three axioms require that only equitable allocations be chosen whenever there are any.

No-Envy. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if $F(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq F(R_N, S)$.

Egalitarian-Equivalence. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if $E(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq E(R_N, S)$.

$\bar{r}$-Egalitarian-Equivalence. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if $E_{\bar{r}}(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq E_{\bar{r}}(R_N, S)$.

[3] Note that, in our present model, there always exists a Pareto efficient allocation in $S \in \mathcal{S}$ since S is finite. To keep symmetry in the definitions of axioms concerning efficiency and equity, and to present a definition which may be applicable to other contexts where there may not exist Pareto efficient allocations, we include the condition that $P(R_N, S) \neq \emptyset$ in the definition of this axiom.

Even if the efficiency criterion is taken as the first principle for social choice, equity criteria should be used when the efficiency criterion is not at all effective as a guide for selection: either when all the available allocations are efficient or when no available allocation is so.

P-Conditional No-Envy. [4] For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if (i) $P(R_N, S) = S$ or $P(R_N, S) = \emptyset$ and (ii) $F(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq F(R_N, S)$.[5]

The next axiom strengthens P-Conditional No-Envy. It means that if either all the available allocations are efficient or no available allocation is efficient, then *all* the envy-free allocations should be recommended. In other words, it claims that we should not discriminate between allocations that equally satisfy the efficiency and equity criteria defined explicitly as axioms.[6]

P-Conditional No-Envy Inclusion. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if (i) $P(R_N, S) = S$ or $P(R_N, S) = \emptyset$ and (ii) $F(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) = F(R_N, S)$.

By simply replacing the correspondence F with each of the correspondences E and $E_{\bar{r}}$ in the above definitions, we define *P-Conditional Egalitarian-Equivalence* and *P-Conditional $\bar{r}$-Egalitarian-Equivalence*, respectively, and their corresponding stronger versions.

Turning now to the equity-first and efficiency-second principle, we define the counterparts of the above axioms. Let a social choice correspondence $\Psi \in \{F, E, E_{\bar{r}}\}$ be given. (The correspondence Ψ is one of the three "equity correspondences.") If the equity criterion described by Ψ is accepted as the first selection principle, we may still apply the efficiency criterion when all the available allocations are equitable or when there is no equitable allocation at all.

Ψ-Conditional Pareto Efficiency. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if (i) $\Psi(R_N, S) = S$ or $\Psi(R_N, S) = \emptyset$ and (ii) $P(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) \subseteq P(R_N, S)$.

As an example, setting $\Psi = F$, we obtain the axiom, *F-Conditional Pareto Efficiency*. A strengthening of Ψ-Conditional Pareto Efficiency is the following.

Ψ-Conditional Pareto Inclusion. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if (i) $\Psi(R_N, S) = S$ or $\Psi(R_N, S) = \emptyset$ and (ii) $P(R_N, S) \neq \emptyset$, then $\varphi(R_N, S) = P(R_N, S)$.

---

[4] The capital letter **P** stands for the Pareto correspondence.

[5] In fact, since every $S \in \mathcal{S}$ is finite, $P(R_N, S)$ is never empty. Hence, in the hypothesis, the case where $P(R_N, S) = \emptyset$ is vacuous. Despite this fact, we chose to write the axiom in this way because we would like to present axioms representing the efficiency-first principle and the equity-first principle in a symmetrical way. See the following axiom called Ψ-Conditional Pareto Efficiency.

[6] The issue of whether we should take the stronger version of conditional equity or efficiency may be related with the issue of *full* vs. *partial* implementation of social choice correspondences. Thomson (1996) made an argument in support of full implementation as follows: "one should make sure that the complete list of desired properties of correspondences have been identified, and then identify the class of correspondences satisfying them. If all the properties are satisfied only by F (that is, if F is characterized by these properties), then full implementation of F is indeed what we should be after" (p. 135).

Next, we introduce several choice-consistency axioms. The first one is called *Path Independence*, which is due to Arrow (1963) and Plott (1973), and may be described as follows. Let S be the set of available allocations, and $\{S_1, S_2\}$ be a partition of S. Suppose that we first choose desirable allocations $\varphi(S_i)$ from each $S_i (i = 1, 2)$, and next make the final choice from $\varphi(S_1) \cup \varphi(S_2)$, that is, from the "winners" of the first round. Then, *Path Independence* requires that for all partitions of S, the final choice should be the same, and hence the choice be independent of the way how to partition S. Therefore, path independent social choice rules are immune to any manipulation through an agenda control.

Path Independence. For all $R_N \in \mathcal{R}^n$ and all $S_1, S_2 \in \mathcal{S}$, $\varphi(R_N, S_1 \cup S_2) = \varphi(R_N, \varphi(R_N, S_1) \cup \varphi(R_N, S_2))$.

Path Independence implies the following choice-consistency condition, which was introduced by Chernoff (1954). Its intuitive meaning is also clear: Suppose that an allocation x is chosen from a set S_1, and then the set of available alternatives is contracted to $S_2 \subset S_1$, but the allocation x is still available. Then, this allocation should be selected from the set S_2 as well.

Contraction Consistency. For all $R_N \in \mathcal{R}^n$ and all $S_1, S_2 \in \mathcal{S}$ with $S_2 \subseteq S_1$, $S_2 \cap \varphi(R_N, S_1) \subseteq \varphi(R_N, S_2)$.

Ever since Arrow (1951), it has been a central issue in social choice theory whether social choice correspondences are *rationalizable*, that is, the choice described by the social choice correspondence from each set of available alternatives could be obtained by maximization of some "well-behaved" social preference relation. The question itself is of much theoretical interest, and moreover it is worth examining because various rationalizability conditions are logically related to choice-consistency conditions.

Let $\succ$ be an irreflexive and asymmetric binary relation on X, the interpretation of which is a strict social preference relation.[7] For each $S \in \mathcal{S}$, denote by $M_\succ(S)$ the set of maximal elements of $\succ$ in S:

$$M_\succ(S) := \{x \in S \mid \text{There exists no } y \in S \text{ such that } y \succ x\}.$$

Quasi-Transitive Rationalizability.[8] For every $R_N \in \mathcal{R}^n$, there exists an irreflexive, asymmetric, and transitive binary relation $\succ_{(R_N)}$ on X such that for all $S \in \mathcal{S}$, $\varphi(R_N, S) = M_{\succ_{(R_N)}}(S)$.

We say that a binary relation $\succ$ has a *cycle* if there exist a positive integer K and K allocations $x^1, \ldots, x^K$ such that $x^k \succ x^{k+1}$ for all k, with $1 \leq k \leq K-1$ and $x^K \succ x^1$. The binary relation $\succ$ is *acyclic* if it has no cycle.[9]

---

[7] It will be convenient for us to present the results by strict social preference relations $\succ$. However, we could alternatively use the reflexive and complete social preference relations $\succsim$ induced from $\succ$ as follows: For all $x, y \in X$, $x \succsim y$ if and only if $y \succ x$ does not hold.

[8] The term "quasi-transitivity" is due to Sen (1970), which means transitivity of strict social preference relations.

[9] Note that if $\succ$ is acyclic, then by definition, it is irreflexive and asymmetric.

$$
\begin{array}{ccc}
\text{Quasi-Transitive Rationalizability} & \Longrightarrow & \text{Acyclic Rationalizability} \\
\Downarrow & & \Downarrow \\
\text{Path Independence} & \Longrightarrow & \text{Contraction Consistency}
\end{array}
$$

Fig. 1 Logical relations of choice-consistency and rationalizability conditions

Acyclic Rationalizability. For every $R_N \in \mathcal{R}^n$, there exists an acyclic binary relation $\succ_{(R_N)}$ on X such that for all $S \in \mathcal{S}$, $\varphi(R_N, S) = M_{\succ_{(R_N)}}(S)$.

The conditions introduced above have the following logical relations.[10] Quasi-Transitive Rationalizability implies both Acyclic Rationalizability and Path Independence, and each of the two conditions, Acyclic Rationalizability and Path Independence, implies Contraction Consistency. The converse of each statement does not hold true. Hence, Contraction Consistency may be considered as the minimal requirement of choice-consistency of social choice correspondences. It is also a necessary condition for any kind of rationalizability by a single binary relation, but it is not a sufficient condition even for Acyclic Rationalizability.[11]

Figure 1 summarizes the logical relations between the axioms. Each arrow indicates the direction of logical implication.

Our final axiom is an obvious requirement: Social choice rules should be able to select some allocations for any environment.

Non-Emptiness. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, $\varphi(R_N, S) \neq \emptyset$.

4 Conflicts Between Efficiency and Equity

This section reviews the fundamental conflict between the Pareto efficiency criterion and the equity criteria. First, we observe the conflict between efficiency and no-envy, which was first pointed out by Kolm (1972).

Example 1. There are two agents $N = \{1, 2\}$ and two goods $\{1, 2\}$. The preferences of the agents are represented by the utility functions

$$
\begin{aligned}
u_1(x_{11}, x_{12}) &= x_{11} x_{12}, \\
u_2(x_{21}, x_{22}) &= 2x_{21} + x_{22}.
\end{aligned}
$$

[10] See Suzumura (1983, Chap. 3).

[11] As a counter-example for the last claim, consider the social choice correspondence defined as follows: Choose $x^0 \in X$. Define the correspondence C by

$$
\begin{aligned}
C(R_N, S) &= \{x^0\} \text{ if } x^0 \in S \text{ and } |S| \geq 3, \\
C(R_N, S) &= S \ \text{ otherwise.}
\end{aligned}
$$

It can be checked that the correspondence C satisfies Contraction Consistency, but cannot be rationalized by any social preference relation.

Let $x = (x_1, x_2) = ((1, \frac{46}{5}), (9, \frac{4}{5}))$ and $y = (y_1, y_2) = ((\frac{11}{5}, \frac{22}{5}), (\frac{39}{5}, \frac{28}{5}))$. Then, since $u_1(y_1) > u_1(x_1)$ and $u_2(y_2) > u_2(x_2)$, the allocation y is Pareto superior to the allocation x. However, x is envy-free because $u_1(x_1) > u_1(x_2)$ and $u_2(x_2) > u_2(x_1)$, whereas y is not since $u_1(y_2) > u_1(y_1)$. Now let $S = \{x, y\}$. Then, both the sets $P(R_N, S)$ and $F(R_N, S)$ are nonempty, but the intersection of the two sets is empty.

This example can be extended to the case of any finite numbers of agents and goods. Hence, we have the following impossibility.

Theorem 1. (Kolm, 1972; Suzumura, 1981a). *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, and No-Envy.*

Next we show that the same kind of conflict may also arise between the Pareto efficiency criterion and the equity-as-egalitarian-equivalence criterion.

Example 2. There are two agents $N = \{1, 2\}$ and two goods $1, 2$. The preferences of the agents are represented by the utility functions[12]

$$\begin{aligned} u_1(x_{11}, x_{12}) &= \min\{x_{11}, x_{12}\}, \\ u_2(x_{21}, x_{22}) &= x_{21} + 3x_{22}. \end{aligned}$$

Let $x = (x_1, x_2) = ((3, 11), (9, 1))$ and $y = (y_1, y_2) = ((8, 8), (4, 4))$. Let $\bar{r} = (1, 1)$. Then, the allocation y is Pareto superior to the allocation x. On the other hand, since $u_1(x_1) = 3 = u_1(3\bar{r})$ and $u_2(x_2) = 12 = u_2(3\bar{r})$, the allocation x is $\bar{r}$-egalitarian-equivalent, with $3\bar{r}$ being the reference bundle. However, y is not egalitarian-equivalent because for any bundle z_0 such that $u_1(z_0) = u_1(y_1)$, $z_0 \geq (8, 8)$ and hence $u_2(z_0) \geq u_2(8, 8) > u_2(4, 4) = u_2(y_2)$. Let $S = \{x, y\}$. Then, both the sets $P(R_N, S)$ and $E(R_N, S) = E_{\bar{r}}(R_N, S)$ are nonempty, but the intersection of the two sets is empty.

Theorem 2. (Tadenuma, 2005). (i) *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, and Egalitarian-Equivalence.* (ii) *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, and $\bar{r}$-Egalitarian-Equivalence.*

5 The Efficiency-First Principle

The results in the previous section show that we cannot always select an allocation that is both Pareto efficient and equitable. Therefore, in cases where the two criteria are conflicting with each other, we have to give priority to one of them. In this section, we adopt the efficiency criterion as the first principle, keeping the requirement of Pareto Efficiency on social choice correspondences. As for equity criteria, however, we only require their conditional versions.

---

[12] We use Leontief preferences only for easy calculations. An example can be constructed with smooth and strictly monotonic preferences.

By the definitions of axioms, Pareto Efficiency and P-Conditional No-Envy (or P-Conditional Egalitarian-Equivalence, P-Conditional $\bar{r}$-Egalitarian-Equivalence) together are compatible with Non-Emptiness. We examine with which choice-consistency conditions these axioms are compatible. To present the results, we introduce several social preference relations.

5.1 No-Envy as the Second Criterion

Let $R_N \in \mathcal{R}^n$ be given. We define the *equity-as-no-envy superior relation*, denoted $\succ_{F(R_N)}$, as follows: For all $x, y \in X$, $x \succ_{F(R_N)} y$ if and only if x is envy-free and y is not. Recall that $\succ_{P(R_N)}$ denotes the Pareto superior relation.

Given $R_N \in \mathcal{R}^n$, define the binary relation $\succ_{PF(R_N)}$ on X as follows: For all $x, y \in X$, $x \succ_{PF(R_N)} y$ if and only if (i) $x \succ_{P(R_N)} y$ or (ii) $x \not\succ_{P(R_N)} y$, $y \not\succ_{P(R_N)} x$, and $x \succ_{F(R_N)} y$.

Under the social preference relation $\succ_{PF(R_N)}$, we first apply the Pareto criterion when we rank any two allocations. Then, only when the Pareto criterion does not give a strict ranking between the two, we apply the equity-as-no-envy criterion.

The next lemma clarifies the relation between the social choice correspondences satisfying Pareto Efficiency, P-Conditional No-Envy, and Contraction Consistency, and the social preference relation $\succ_{PF(R_N)}$.

Lemma 1. *If a social choice correspondence φ satisfies Pareto Efficiency, P-Conditional No-Envy, and Contraction Consistency, then $\varphi(R_N, S) \subseteq M_{\succ_{PF(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$.*

Proof. Suppose that a social choice correspondence φ satisfies Pareto Efficiency, P-Conditional No-Envy, and Contraction Consistency. Let $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$ be given. Suppose, on the contrary, that there exists $x \in S$ such that $x \in \varphi(R_N, S)$ but $x \notin M_{\succ_{PF(R_N)}}(S)$. Then, there exists $y \in S$ such that $y \succ_{PF(R_N)} x$. Because $x \in \varphi(R_N, S) \subseteq P(R_N, S)$ by Pareto Efficiency, $y \not\succ_{P(R_N)} x$. Hence, $y \succ_{PF(R_N)} x$ holds only if $x \not\succ_{P(R_N)} y$ and $y \succ_{F(R_N)} x$. Let $S' = \{x, y\}$. Then, $S' \subseteq S$ and $P(R_N, S') = S'$. By P-Conditional No-Envy, $\varphi(R_N, S') \subseteq F(R_N, S')$. Thus, $x \notin \varphi(R_N, S')$. This means, however, that φ violates Contraction Consistency, which is a contradiction. □

It is well-known that for any binary relation $\succ$, the set $M_{\succ}(S)$ is nonempty for all $S \in \mathcal{S}$ if and only if $\succ$ is acyclic (Sen, 1970). Hence, it follows from the above lemma that there exist social choice correspondences satisfying the three axioms and non-emptiness only if $\succ_{PF(R_N)}$ is acyclic. Unfortunately, the social preference relation $\succ_{PF(R_N)}$ may have a cycle.

Proposition 1. (Tadenuma, 2002). *There exists a preference profile $R_N \in \mathcal{R}^n$ such that $\succ_{PF(R_N)}$ has a cycle.*

Proof. For simplicity of presentation, we consider a two-agent and two-good economy. Similar examples can be constructed for the case of any numbers of agents and goods. Let $N = \{1,2\}$ be the set of agents. Assume that agent $i \in N$ has the preference relation R_i on $\mathbb{R}^2_+$ that is represented by the following utility function:

$$u_1(x_{11}, x_{12}) = x_{11}x_{12},$$
$$u_2(x_{21}, x_{22}) = 2x_{21} + x_{22}.$$

Define four allocations x, y, z, and w by $x = ((1,9),(9,1))$, $y = ((3,6),(7,4))$, $z = ((2,8),(8,2))$, and $w = ((2,7),(8,3))$. Then, $x \succ_{PF(R_N)} y$ since $x \not\succ_{P(R_N)} y$, $y \not\succ_{P(R_N)} x$, and $x \succ_{F(R_N)} y$. Since $y \succ_{P(R_N)} z$, we have $y \succ_{PF(R_N)} z$. Because $z \not\succ_{P(R_N)} w$, $w \not\succ_{P(R_N)} z$, and $z \succ_{F(R_N)} w$, we have $z \succ_{PF(R_N)} w$. Finally, $w \succ_{PF(R_N)} x$ follows from the fact that $w \succ_{P(R_N)} x$. Thus, the relation $\succ_{PF(R_N)}$ has a cycle. □

From Lemma 1 and Proposition 1, the next impossibility theorem follows.

Theorem 3. *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional No-Envy, and Contraction Consistency.*

We have argued that Path Independence is an indispensable property of social choice correspondences. However, since Path Independence implies Contraction Consistency, we have the following impossibility as a corollary of Theorem 3.

Corollary 1. *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional No-Envy, and Path Independence.*

5.2 Egalitarian-Equivalence as the Second Criterion

Next, we adopt egalitarian-equivalence as the concept of equity instead of no-envy. The analyses will go parallel to those in the previous subsection. We define the *equity-as-egalitarian-equivalence superior relation* $\succ_{E(R_N)}$ as $x \succ_{E(R_N)} y$ if and only if x is egalitarian-equivalent and y is not. Then, define the binary relation $\succ_{PE(R_N)}$ on X as $x \succ_{PE(R_N)} y$ if and only if (i) $x \succ_{P(R_N)} y$ or (ii) $x \not\succ_{P(R_N)} y$, $y \not\succ_{P(R_N)} x$, and $x \succ_{E(R_N)} y$.

Just like Lemma 1, we can show that if a social choice correspondence φ satisfies Pareto Efficiency, P-Conditional Egalitarian-Equivalence, and Contraction Consistency, then $\varphi(R_N, S) \subseteq M_{\succ_{PE(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$. Hence, whether the social preference relation $\succ_{PE(R_N)}$ is acyclic or not is a crucial question for the existence of social choice correspondence satisfying the three axioms and Non-Emptiness. However, $\succ_{PE(R_N)}$ may have a cycle.

Proposition 2. (Tadenuma, 2005). *There exists a preference profile $R_N \in \mathcal{R}^n$ such that $\succ_{PE(R_N)}$ has a cycle.*

Proof. For simplicity of presentation, we consider a two-agent and two-good economy. Let $N = \{1,2\}$ be the set of agents. Assume that agent 1's preference relation

R_1 on $\mathbb{R}^2_+$ is represented by a Leontief utility function:

$$u_1(x_{11}, x_{12}) = \min\{x_{11}, x_{12}\}.$$

Agent 2's preference relation R_2 is represented by the following piece-wise linear utility function:

$$u_2(x_{21}, x_{22}) = x_{21} + 20, \text{ if } x_{22} \geq x_{21} \text{ and } x_{22} \geq 20,$$
$$u_2(x_{21}, x_{22}) = x_{21} + x_{22}, \text{ if } x_{22} \geq x_{21} \text{ and } x_{22} \leq 20,$$
$$u_2(x_{21}, x_{22}) = 2x_{22}, \text{ if } x_{22} \leq x_{21}.$$

Define four allocations $x, y, z,$ and w by $x = ((18,9),(10,19))$, $y = ((12,10),(16,18))$, $z = ((23,11),(5,17))$, and $w = ((17,15),(11,13))$. Then, observe the following facts: (1) $y \succ_{P(R_N)} x$; (2) $w \succ_{P(R_N)} z$; (3) $y \not\succ_{P(R_N)} z$ and $z \not\succ_{P(R_N)} y$; (4) $x \not\succ_{P(R_N)} w$ and $w \not\succ_{P(R_N)} x$; (5) x is egalitarian-equivalent with a reference bundle $(9,20)$ since $u_1(x_1) = 9 = u_1(9,20)$ and $u_2(x_2) = 29 = u_2(9,20)$; (6) z is egalitarian-equivalent with a reference bundle $(11,11)$; (7) y is not egalitarian-equivalent because for all $a_0 \in \mathbb{R}^2_+$ such that $u_2(a_0) = u_2(y_2) = 34$, $a_0 \geq (13,13)$, and hence $u_1(a_0) \geq u_1(13,13) = 13 > 10 = u_1(y_1)$; (8) w is not egalitarian-equivalent since for all $b_0 \in \mathbb{R}^2_+$ such that $u_1(b_0) = u_1(w_1) = 15$, $b_0 \geq (15,15)$, and thus $u_2(b_0) \geq u_2(15,15) = 30 > 24 = u_2(w_2)$.

By (1), we have $y \succ_{PE(R_N)} x$. It follows from (3), (6), and (7) that $z \succ_{PE(R_N)} y$. By (2), $w \succ_{PE(R_N)} z$. Finally, from (4), (5), and (8) together, we have $x \succ_{PE(R_N)} w$. Thus, the relation $\succ_{PE(R_N)}$ has a cycle. □

By Proposition 2, we have the following impossibility results.

Theorem 4. *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional Egalitarian-Equivalence, and Contraction Consistency.*

Corollary 2. *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional Egalitarian-Equivalence, and Path Independence.*

5.3 $\bar{r}$-Egalitarian-Equivalence as the Second Criterion

We have reached an impossibility again with egalitarian-equivalence as the second criterion. In this subsection, we adopt a more restricted concept of equity than egalitarian-equivalence, namely $\bar{r}$-egalitarian-equivalence. Let us recall that the reference bundles of $\bar{r}$-egalitarian-equivalent allocations must lie in the given ray from the origin, while there is no such restriction in the definition of (general) egalitarian-equivalent allocations. With $\bar{r}$-egalitarian-equivalence as the second criterion, we have a positive result as shown next.

Theorem 5. *There exists a social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional $\bar{r}$-Egalitarian-Equivalence, and Path Independence.*

Proof. The proof relies on the social preference relation introduced by Pazner and Schmeidler (1978). For each $R_N \in \mathcal{R}^n$, define a binary relation $\succsim_{PS(R_N)}$ on $\mathbb{R}^{mn}_+$ as follows. For all $x, y \in X$, $x \succsim_{PS(R_N)} y$ if and only if

$$\min_{i \in N} \min\{\lambda_i \in \mathbb{R} \mid \lambda_i \bar{r} \, R_i \, x_i\} \geq \min_{i \in N} \min\{\lambda_i \in \mathbb{R} \mid \lambda_i \bar{r} \, R_i \, y_i\}.$$

Let $\succ_{PS(R_N)}$ be the strict part of $\succsim_{PS(R_N)}$. Define the social choice function φ_{PS} by $\varphi_{PS}(R_N, S) = M_{\succ_{PS(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$. It is easy to check that φ_{PS} satisfies Non-Emptiness, Pareto Efficiency, and P-Conditional $\bar{r}$-Egalitarian-Equivalence. Since $\succsim_{PS(R_N)}$ is transitive and complete, and transitive rationalizability implies Path Independence, it follows that φ_{PS} satisfies Path Independence. □

We argued that if explicitly defined criteria of efficiency and equity should be the only selection principles, then the stronger version of conditional efficiency and equity should be accepted. That is, if the first principle selects all or none, then the allocations satisfying the second principle should be all taken, and no discrimination between them should be introduced. Next we examine the compatibility of this stronger version with choice-consistency.

We define the *equity-as-$\bar{r}$-egalitarian-equivalence superior relation* $\succ_{E_{\bar{r}}(R_N)}$ as $x \succ_{E_{\bar{r}}(R_N)} y$ if and only if x is $\bar{r}$-egalitarian-equivalent and y is not.

Theorem 6. *There exists a social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional $\bar{r}$-Egalitarian-Equivalence inclusion, $E_{\bar{r}}$-Conditional Pareto Inclusion, and Contraction Consistency.*

Proof. For each $R_N \in \mathcal{R}^n$, define the binary relation $\succ_{PE_{\bar{r}}(R_N)}$ as follows. For all $x, y \in X$, $x \succ_{PE_{\bar{r}}(R_N)} y$ if and only if (i) $x \succ_{P(R_N)} y$ or (ii) $x \not\succ_{P(R_N)} y$, $y \not\succ_{P(R_N)} x$, and $x \succ_{E_{\bar{r}}(R_N)} y$. Define the social choice correspondence $\varphi_{PE_{\bar{r}}}$ by $\varphi_{PE_{\bar{r}}}(R_N, S) = M_{\succ_{PE_{\bar{r}}(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$. By Tadenuma (2005, Prop. 5), $\succ_{PE_{\bar{r}}(R_N)}$ is acyclic, and hence $\varphi_{PE_{\bar{r}}}$ satisfies non-emptiness. It is clear that $\varphi_{PE_{\bar{r}}}$ satisfies Pareto Efficiency. To check that it satisfies P-Conditional $\bar{r}$-Egalitarian-Equivalence Inclusion, let $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$ be such that $P(R_N, S) = S$ and $E_{\bar{r}}(R_N, S)$ is non-empty. Clearly, $\varphi_{PE_{\bar{r}}}(R_N, S) \subseteq E_{\bar{r}}(R_N, S)$. Suppose that there exists $x \in E_{\bar{r}}(R_N, S)$ such that $x \notin \varphi_{PE_{\bar{r}}}(R_N, S)$. Then, there is $y \in S$ with $y \succ_{PE_{\bar{r}}(R_N)} x$. Since $x, y \in P(R_N, S)$, y cannot be Pareto superior to x. Because $x \in E_{\bar{r}}(R_N, S)$, y cannot be equity-as-$\bar{r}$-egalitarian-equivalence superior to x, either. This is a contradiction. Hence, $E_{\bar{r}}(R_N, S) \subseteq \varphi_{PE_{\bar{r}}}(R_N, S)$. Similarly, we can show that $\varphi_{PE_{\bar{r}}}$ satisfies $E_{\bar{r}}$-Conditional Pareto Inclusion. Finally, since $\varphi_{PE_{\bar{r}}}(R_N, \cdot)$ is rationalizable by the binary relation $\succ_{PE_{\bar{r}}(R_N)}$, it satisfies Contraction Consistency. □

However, if we strengthen the requirement of choice-consistency, we have another impossibility.

Theorem 7. *There exists no social choice correspondence that satisfies Non-Emptiness, Pareto Efficiency, P-Conditional $\bar{r}$-Egalitarian-Equivalence, $E_{\bar{r}}$-Conditional Pareto Inclusion, and Path Independence.*

Proof. Consider the economy defined in the proof of Proposition 2. Let $\bar{r} = (9,20)$. Define three allocations x, y, and w by $x = ((18,9),(10,19)), y = ((12,10),(16,18))$, and $w = ((17,15),(11,13))$. Then, the allocation x is $\bar{r}$-egalitarian-equivalent, but the other two allocations are not. On the other hand, $y \succ_{P(R_N)} x$, but $y \not\succ_{P(R_N)} w$ and $w \not\succ_{P(R_N)} y$. Similarly, $x \not\succ_{P(R_N)} w$ and $w \not\succ_{P(R_N)} x$.

Suppose that there exists a social choice correspondence φ that satisfies Pareto Efficiency, P-Conditional $\bar{r}$-Egalitarian-Equivalence, $E_{\bar{r}}$-Conditional Pareto Inclusion, Path Independence, and Non-Emptiness. It follows from Pareto Efficiency, P-Conditional $\bar{r}$-Egalitarian-Equivalence, and Path Independence (which implies Contraction Consistency) that $\varphi(R_N, S) \subseteq M_{\succ_{PE_{\bar{r}}(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$.

Let $S_1 = \{x,y,w\}$, $S_2 = \{x,y\}$, and $S_3 = \{w\}$. Then, $M_{\succ_{PE_{\bar{r}}(R_N)}}(S_1) = \{y\}$ and $M_{\succ_{PE_{\bar{r}}(R_N)}}(S_2) = \{y\}$. By the above relation and Non-Emptiness of φ, we have

$$\varphi(R_N, S_1) = \{y\} \text{ and } \varphi(R_N, S_2) = \{y\}. \tag{1}$$

By Non-Emptiness, $\varphi(R_N, S_3) = \{w\}$. Hence,

$$\varphi(R_N, \varphi(R_N, S_2) \cup \varphi(R_N, S_3)) = \varphi(R_N, \{y, w\}). \tag{2}$$

Observe that $E_{\bar{r}}(R_N, \{y,w\}) = \emptyset$ and $P(R_N, \{y,w\}) = \{y,w\}$. Since φ satisfies $E_{\bar{r}}$-Conditional Pareto Inclusion, we have

$$\varphi(R_N, \{y,w\}) = \{y,w\}. \tag{3}$$

It follows from (1), (2), and (3) that $\varphi(R_N, \varphi(R_N, S_2) \cup \varphi(R_N, S_3)) \neq \varphi(R_N, S_1)$. This means that φ violates Path Independence, which is a contradiction. □

6 The Equity-First Principle

In this section, we reverse the order of application of the efficiency and equity criteria. That is, we first select from equitable allocations, and if the equity criterion is not effective as a guide for selection either because *all* the available allocations are equitable or because *no* allocation is equitable, then we choose from efficient allocations. In the following, we consider the equity-as-no-envy criterion. However, essentially the same results hold true with egalitarian-equivalence or $\bar{r}$-egalitarian-equivalence.

To identify the social choice correspondences satisfying No-Envy, F-Conditional Pareto Efficiency, and Contraction Consistency, it is useful to introduce a social preference relation. Given $R_N \in \mathcal{R}^n$, define the binary relation $\succ_{FP(R_N)}$ on X as follows: For all $x, y \in X$, $x \succ_{FP(R_N)} y$ if and only if (i) $x \succ_{F(R_N)} y$ or (ii) $x \not\succ_{F(R_N)} y$,

$y \not\succ_{F(R_N)} x$, and $x \succ_{P(R_N)} y$. Under this social preference relation, we first apply the equity-as-no-envy criterion to rank any two allocations, and when the two allocations are not strictly ranked by the criterion because both are envy-free or neither is, we invoke the efficiency criterion.

The next lemma is the counterpart of Lemma 1.

Lemma 2. *If a social choice correspondence φ satisfies No-Envy, F-Conditional Pareto Efficiency, and Contraction Consistency, then $\varphi(R_N,S) \subseteq M_{\succ_{FP(R_N)}}(S)$ for all $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$.*

Proof. Suppose that a social choice correspondence φ satisfies No-Envy, F-Conditional Pareto Efficiency, and Contraction Consistency. Let $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$ be given. Suppose, on the contrary, that there exists $x \in S$ such that $x \in \varphi(R_N,S)$ but $x \notin M_{\succ_{FP(R_N)}}(S)$. Then, there exists $y \in S$ such that $y \succ_{FP(R_N)} x$. We distinguish two cases.

Case 1: $F(R_N,S) \neq \emptyset$.
Because $x \in \varphi(R_N,S) \subseteq F(R_N,S)$ by No-Envy, $y \succ_{FP(R_N)} x$ holds only if both x and y are envy-free, and y is Pareto superior to x. Let $S' = \{x,y\}$. Then, $S' \subseteq S$, and $F(R_N,S') = S'$. By F-Conditional Pareto Efficiency, $\varphi(R_N,S') \subseteq P(R_N,S') = \{y\}$. Thus, $x \notin \varphi(R_N,S')$. This means that φ violates Contraction Consistency, which is a contradiction.

Case 2: $F(R_N,S) = \emptyset$.
Then, $y \succ_{FP(R_N)} x$ holds only if y is Pareto superior to x. Let $S' = \{x,y\}$. Then, $S' \subseteq S$ and $F(R_N,S') = \emptyset$. By F-Conditional Pareto Efficiency, $\varphi(R_N,S') \subseteq P(R_N,S') = \{y\}$. The rest of the argument is the same as Case 1. $\square$

The next result should be contrasted with Proposition 1.

Proposition 3. *For all $R_N \in \mathcal{R}^n$, $\succ_{FP(R_N)}$ is transitive.*

Proof. Let $R_N \in \mathcal{R}^n$ be given. To lighten notation, we simply write $\succ_{FP}$, $\succ_P$, and F for $\succ_{FP(R_N)}$, $\succ_{P(R_N)}$, and $F(R_N,X)$, respectively. Assume that $x \succ_{FP} y$ and $y \succ_{FP} z$. By $x \succ_{FP} y$, (1) $x \in F$ and $y \notin F$ or (2) $[[x \in F$ and $y \in F]$ or $[x \notin F$ and $y \notin F]]$ and $x \succ_P y$. By $y \succ_{FP} z$, (3) $y \in F$ and $z \notin F$ or (4) $[[y \in F$ and $z \in F]$ or $[y \notin F$ and $z \notin F]]$ and $y \succ_P z$. (1) and (3) are incompatible. If (1) and (4) hold, then we must have $x \in F$ and $z \notin F$. Hence, $x \succ_{FP} z$. Similarly, if (2) and (3) hold, then $x \in F$ but $z \notin F$, and we have $x \succ_{FP} z$. If (2) and (4) hold, then either $x,y,z \in F$ and $x \succ_P y \succ_P z$ or $x,y,z \notin F$ and $x \succ_P y \succ_P z$. Since the relation $\succ_P$ is transitive, we have $x \succ_{FP} z$. $\square$

Define the social choice correspondence φ_{FP} by

$$\varphi_{FP}(R_N,S) = M_{\succ_{FP(R_N)}}(S) \text{ for all } (R_N,S) \in \mathcal{R}^n \times \mathcal{S}.$$

This correspondence takes the following values:

$$\begin{aligned} \varphi_{FP}(R_N,S) &= P(R_N,F(R_N,S)) \;\text{ if } F(R_N,S) \neq \emptyset, \\ \varphi_{FP}(R_N,S) &= P(R_N,S) \;\text{ if } F(R_N,S) = \emptyset. \end{aligned}$$

Note that for all $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$, $P(R_N,F(R_N,S)) \supseteq P(R_N,S) \cap F(R_N,S)$ and there exists $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$ such that $P(R_N,F(R_N,S)) \neq P(R_N,S) \cap F(R_N,S)$.

We next show that the social choice correspondence φ_{FP} is characterized by No-Envy, the stronger versions of Conditional Pareto Efficiency and No-Envy, and Path Independence.

Theorem 8. *A social choice correspondence φ satisfies Non-Emptiness, No-Envy, P-Conditional No-Envy Inclusion, F-Conditional Pareto Inclusion, and Path Independence if and only if $\varphi = \varphi_{FP}$.*

Proof. First, we show that φ_{FP} satisfies the five axioms. It is easy to check that φ_{FP} satisfies No-Envy, F-Conditional Pareto Inclusion, and P-Conditional No-Envy Inclusion. By Proposition 3, φ_{FP} satisfies Quasi-Transitive Rationalizability, which implies Path Independence. Since any quasi-transitive binary relation has maximal elements in any finite set, Non-Emptiness follows.

Next we show that φ_{FP} is the unique social choice correspondence that satisfies the five axioms. Suppose, on the contrary, that there is a social choice correspondence φ with $\varphi \neq \varphi_{FP}$ that satisfies the five axioms. Then, there is $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$ such that

$$\varphi(R_N,S) \neq \varphi_{FP}(R_N,S). \tag{1}$$

If $F(R_N,S) = \emptyset$, then by F-Conditional Pareto Inclusion, $\varphi(R_N,S) = P(R_N,S)$. On the other hand, it can be checked that $\varphi_{FP}(R_N,S) = M_{\succ_{FP(R_N)}}(S) = P(R_N,S)$. Hence, we have $\varphi(R_N,S) = \varphi_{FP}(R_N,S)$, a contradiction. Thus, $F(R_N,S) \neq \emptyset$. Since F-Conditional Pareto Inclusion implies F-Conditional Pareto Efficiency, and Path Independence implies Contraction Consistency, it follows from Lemma 2 that

$$\varphi(R_N,S) \subseteq M_{\succ_{FP(R_N)}}(S) = \varphi_{FP}(R_N,S). \tag{2}$$

Because $F(R_N,S) \neq \emptyset$, we have

$$M_{\succ_{FP(R_N)}}(S) = P(R_N,F(R_N,S)). \tag{3}$$

It follows from (1), (2), and (3) that there exists $x^* \in S$ such that $x^* \in P(R_N,F(R_N,S))$ but $x^* \notin \varphi(R_N,S)$. Define $S' := \{x^*\} \cup \{y \in S \mid y \in F(R_N,S)$ and $x^* \succ_{P(R_N)} y\} \cup [S \backslash F(R_N,S)]$. By Lemma 2, $\varphi(R_N,S') \subseteq M_{\succ_{FP(R_N)}}(S') = \{x^*\}$. By Non-Emptiness, we have $\varphi(R_N,S') = \{x^*\}$. Define $S'' := S \backslash S'$. Again from Lemma 2, it follows that

$$\varphi(R_N,S'') \subseteq M_{\succ_{FP(R_N)}}(S'') = P(R_N,F(R_N,S'')). \tag{4}$$

Claim: $P(R_N,F(R_N,S'')) \subset P(R_N,F(R_N,S))$.
Let $z \in P(R_N,F(R_N,S''))$. Then, $z \in F(R_N,S'') \subset F(R_N,S)$. Suppose that $z \notin P(R_N,F(R_N,S))$. Then, there exists $w \in P(R_N,F(R_N,S))$ such that $w \succ_{P(R_N)} z$. If $w = x^*$, then $z \in S'$ and hence $z \notin S''$, which is a contradiction. Thus, $w \neq x^*$.

But then, $w \in S''$ and so $z \notin P(R_N, F(R_N, S''))$, which contradicts $z \in P(R_N, F(R_N, S''))$. Therefore, we must have $z \in P(R_N, F(R_N, S))$. Thus, the claim has been proved.

It follows from (4) and the above claim that $\varphi(R_N, S'') \subseteq P(R_N, F(R_N, S))$. Hence, $\varphi(R_N, S') \cup \varphi(R_N, S'') = \{x^*\} \cup \varphi(R_N, S'') \subseteq P(R_N, F(R_N, S))$. Therefore, $P(R_N, \varphi(R_N, S') \cup \varphi(R_N, S'')) = \varphi(R_N, S') \cup \varphi(R_N, S'')$. Then, by P-Conditional No-Envy Inclusion, we conclude that $\varphi(R_N, \varphi(R_N, S') \cup \varphi(R_N, S'')) = F(R_N, \varphi(R_N, S') \cup \varphi(R_N, S'')) = \varphi(R_N, S') \cup \varphi(R_N, S'') = \{x^*\} \cup \varphi(R_N, S'')$. But since $x^* \notin \varphi(R_N, S) = \varphi(R_N, S' \cup S'')$ and φ satisfies Path Independence, we must have $x^* \notin \varphi(R_N, \varphi(R_N, S') \cup \varphi(R_N, S''))$. This is a contradiction.

Therefore, there is no social choice correspondence φ with $\varphi \neq \varphi_{FP}$ that satisfies the five axioms together in the statement of the theorem. $\square$

7 Minimal-Envy and Choice-Consistency

In the previous sections, our equity-as-no-envy criterion made only "all-or-nothing" selection: an allocation is selected if there is no envy at all, whereas it is not selected if there is at least one instance of envy. However, among allocations with envy, the instances of envy may differ greatly. In such cases, it should be desirable to select allocations with "minimal" instances of envy.

In this section, we introduce a measure of envy at allocations, which is due to Suzumura (1996). Based on the measure, we define the notion of *minimal-envy*. Then, as in the foregoing section, we examine the choice-consistency properties of social choice correspondences satisfying the efficiency first or the equity first principle with minimal-envy as the concept of equity.

For each $R_N \in \mathcal{R}^n$, and each $x \in X$, define the set $H(R_N, x) \subset N \times N$ by

$$H(R_N, x) = \{(i, j) \in N \times N \mid x_j \, P_i \, x_i\}.$$

The set $H(R_N, x)$ is the set of all *instances of envy* at x. Following Suzumura (1996), we define the binary relation $\succ_{F_{\min}(R_N)}$ as follows: for all $x, y \in X$, $x \succ_{F_{\min}(R_N)} y$ if and only if $H(R_N, x) \subsetneq H(R_N, y)$.

In a similar way to defining Pareto efficiency, we can define the notion of *minimal-envy*. Given $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, an allocation $x \in S$ is *envy-minimal in S for R_N* if there is no allocation $y \in S$ such that $H(R_N, y) \subsetneq H(R_N, x)$. Let $F_{\min}(R_N, S)$ be the set of envy-minimal allocations in S for R_N. By simply replacing $F(R_N, S)$ with $F_{\min}(R_N, S)$ in the definitions of axioms in Sect. 3, we can define axioms, *Minimal-Envy*, *P-Conditional Minimal-Envy*, and $F_{\min}$-*Conditional Pareto Efficiency*.

Reexamining the proofs of Lemma 1, Proposition 1, and Theorem 3, we find that there is no social choice correspondence that satisfies *Non-Emptiness*, *Pareto Efficiency*, *P-Conditional Minimal-Envy* and *Contraction Consistency*. Moreover, with this concept of equity, even the equity-first principle may contradict the minimum requirement of choice-consistency, which we now turn to.

For each $R_N \in \mathcal{R}^n$, define the binary relation $\succ_{F_{\min}P(R_N)}$ as follows: For all $x, y \in X$, $x \succ_{F_{\min}P(R_N)} y$ if and only if (i) $x \succ_{F_{\min}(R_N)} y$, or (ii) $x \not\succ_{F_{\min}(R_N)} y$, $y \not\succ_{F_{\min}(R_N)} x$, and $x \succ_{P(R_N)} y$. Then, as in Lemma 2, we can show that if a social choice correspondence φ satisfies *Non-Emptiness*, *Minimal-Envy*, $F_{\min}$-*Conditional Pareto Efficiency*, and *Contraction Consistency*, then $\varphi(R_N, S) \subseteq M_{\succ_{F_{\min}P(R_N)}}(S)$ for all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$.

Proposition 4. *There exist a preference profile $R_N \in \mathcal{R}^n$ such that $\succ_{F_{\min}P(R_N)}$ has a cycle.*

Proof. Consider an economy with four agents and two goods. Let $N = \{1,2,3,4\}$ be the set of agents. Assume that the agents' preference relations are represented by the following utility functions:

$$u_1(x_{11}, x_{12}) = x_{11}x_{12},$$
$$u_3(x_{31}, x_{32}) = x_{31} + x_{32},$$

and

$$u_2 = u_1 \text{ and } u_4 = u_3.$$

Consider the following three allocations, x, y, and z:

$$(x_1, x_2, x_3, x_4) = ((6,0), (0,6), (3,3), (3,3)),$$
$$(y_1, y_2, y_3, y_4) = ((0.5,1), (0.5,6), (8.5,0), (2.5,5)),$$
$$(z_1, z_2, z_3, z_4) = ((1,1), (2,2), (5,4), (4,5)).$$

Observe that

$$z \succ_{P(R_N)} y \succ_{P(R_N)} x$$

and

$$H(R_N, x) = \{(1,3), (1,4), (2,3), (2,4)\},$$
$$H(R_N, y) = \{(1,4), (2,4), (1,2), (4,3)\},$$
$$H(R_N, z) = \{(1,3), (1,4), (2,3), (2,4), (1,2)\}.$$

Since $H(R_N, x) \not\subseteq H(R_N, y)$, $H(R_N, y) \not\subseteq H(R_N, x)$, and $y \succ_{P(R_N)} x$, we have $y \succ_{F_{\min}P(R_N)} x$. Similarly, $z \succ_{F_{\min}P(R_N)} y$ holds true. However, it follows from $H(R_N, x) \subsetneq H(R_N, z)$ that $x \succ_{F_{\min}P(R_N)} z$. Thus, there is a cycle for $\succ_{F_{\min}P(R_N)}$. □

In a framework of abstract social choice, Suzumura (2004) defines the binary relation $\succsim_{F_S P(R_N)}$ as follows: For all $x, y \in X$, $x \succsim_{F_S P(R_N)} y$ if and only if (i) $H(R_N, x) \subseteq H(R_N, y)$ or (ii) $H(R_N, x) \not\subseteq H(R_N, y)$, $H(R_N, y) \not\subseteq H(R_N, x)$, and $x \succsim_P y$. Let $\succ_{F_S P(R_N)}$ be the strict part of $\succsim_{F_S P(R_N)}$. Suzumura (2004) presented an example such that $\succsim_{F_S P(R_N)}$ is acyclic but is not *consistent*. Let us recollect that a binary relation $\succsim$ on X is *consistent* if for any integer $k \geq 3$, there exists no finite set $\{x^1, x^2, \ldots, x^k\} \subseteq X$ such that $x^1 \succ x^2$, $x^2 \succsim x^3, \ldots, x^{k-1} \succsim x^k$, and $x^k \succsim x^1$.

Notice a difference between the definitions of $\succ_{F_{\min}P(R_N)}$ and $\succ_{F_S P(R_N)}$. Just as all other lexicographic compositions of two criteria studied in this chapter and

Tadenuma (2002, 2005), the relation $\succ_{F_{\min}P(R_N)}$ invokes the second criterion (the Pareto superior relation in this case) when the first criterion judges the two alternatives to be indifferent or noncomparable, and if there is a strict ranking between the two by the second criterion, it is adopted. In contrast, the relation $\succ_{F_S P(R_N)}$ does not make any strict ranking when the two alternatives are indifferent for the first criterion. For example, if $H(R_N,x) = H(R_N,y)$ and $x \succ_{P(R_N)} y$, then $x \succ_{F_{\min}P(R_N)} y$ but $x \not\succ_{F_S P(R_N)} y$.

A motivation for our lexicographic composition of two binary relations is that we should apply the second criterion to evaluate desirability of two allocations whenever they are noncomparable *or* equally good by the first criterion, so that we could have a more fine-grained social ranking of allocations. If the set of pairs of envious and envied agents is exactly the same for the two allocations, and one of them is Pareto superior to the other, why do not we choose the former? Similarly, if the two allocations give the same utility for every agent, and one of them is more equitable than the other, then we should select the more equitable.

In spite of this difference in the definitions, the proof of Proposition 4 actually shows that the relation $\succ_{F_S P(R_N)}$ also has a cycle for the preference profile because the proof does not depend on the case where the set of instances of envy is exactly the same for the two allocations. Thus, Proposition 4 strengthens Suzumura's result in two respects. First, we show that the relation $\succsim_{F_S P(R_N)}$ is not even acyclic. Note that acyclicity is a weaker condition than consistency. Second, we establish the result on the domain of classical exchange economies with no free disposal of goods and continuous, convex and strictly monotonic preferences. Even on this much restricted domain, we cannot avoid a cycle of the strict part of $\succsim_{F_S P(R_N)}$.

From Proposition 4, we obtain another impossibility theorem. The proof is analogous to those for the previous results, and it is omitted.

Theorem 9. *There exists no social choice correspondence that satisfies Non-Emptiness, Minimal-Envy, $F_{\min}$-Conditional Pareto Efficiency, and Contraction Consistency.*

Feldman and Kirman (1974) introduced a different measure of envy at allocations, which simply counts the number of instances of envy. Based on the measure, we may define the notion of the *least envy*.

For each $(R_N,S) \in \mathcal{R}^n \times \mathcal{S}$, we say that an allocation $x \in S$ has the *least envy in S for R_N* if there is no allocation $y \in S$ such that $\#H(R_N,y) < \#H(R_N,x)$.[13] Let $F_{\text{least}}(R_N,S)$ be the set of allocations that have the least envy in S. Then, we can similarly define axioms, *Least-Envy*, *P-Conditional Least-Envy*, and *F_{least}-Conditional Pareto Efficiency*.

As in Theorems 3 and 8, we can show that (i) there exists no social choice correspondence that satisfies *Non-Emptiness*, *Pareto Efficiency*, *P-Conditional Least-Envy*, and *Contraction Consistency*; and (ii) there exists a social choice correspondence that satisfies *Non-Emptiness*, *Least-Envy*, *F_{least}-Conditional Pareto Efficiency*, and *Path Independence*. The proofs are similar, and we omit them.

[13] For any set A, $\#A$ denotes the cardinality of A.

8 Concluding Remarks

In his two seminal works, Suzumura (1981a, b) considered a class of abstract social choice problems and examined possibility of constructing social choice correspondences satisfying the following conditions on efficiency and equity.

Fairness Extension. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, $\varphi(R_N, S) = P(R_N, S) \cap F(R_N, S)$ if $P(R_N, S) \cap F(R_N, S) \neq \emptyset$.

This axiom requires that if there are Pareto efficient and equitable allocations, then they should all be selected.

Fairness Inclusion. For all $(R_N, S) \in \mathcal{R}^n \times \mathcal{S}$, if $P(R_N, S) \cap F(R_N, S) = \emptyset$, $y \in \varphi(R_N, S)$, and $x \in S$ is Pareto superior to y or equity-as-no-envy superior to y for R_N, then $x \in \varphi(R_N, S)$.

This means that if some allocation is selected, and there are allocations that are superior to the former *either* in the Pareto principle *or* in the equity criterion, then the latter allocations should also be selected.

A basic difference of these axioms from ours is that they treat the efficiency criterion and the equity criterion with equal weight, whereas our axioms give priority to one of the two criteria. Indeed, there is no logical relation between Fairness Extension or Fairness Inclusion and any one or any combination of our axioms concerning efficiency and equity. Moreover, combined with the requirement of Non-Emptiness, Fairness Inclusion is incompatible with either of our axioms Pareto Efficiency and No-Envy. To see this, let us reconsider the case of fundamental conflict between the Pareto criterion and the equity-as-no-envy criterion as in Sect. 4. In Example 1, the allocation y is Pareto superior to the allocation x, whereas x is equity-as-no-envy-superior to y. Then, if a social choice correspondence φ satisfies Non-Emptiness and Fairness Inclusion, then $\varphi(\{x, y\}) = \{x, y\}$. That is, any correspondence satisfying this axiom avoids selection in face of the fundamental conflict. To the contrary, the correspondences satisfying our efficiency-first or equity-first axioms do make a selection in the case of the fundamental conflict, depending upon which criterion should be placed first.

Another difference between Suzumura (1981a, b) and this chapter lies in the domain of social choice problems. While he considers a class of abstract social choice problems with no restrictions on individual preferences except rationality, we study the class of canonical economic problems of distributing infinitely divisible goods among n agents with preferences that satisfy all standard assumptions in economics. For the kind of axioms studied in this chapter, with more restrictions on the domain of problems, the more cases for compatibility among required conditions may arise. To put it in other words, impossibility results obtained in this chapter straightforwardly extend to the unrestricted domain.

This work started with the simple question: Which criterion should we take first to select socially desirable allocations, the efficiency criterion or the equity criterion? We have represented two alternative principles in the form of axioms, and examined choice-consistency of the social choice correspondences satisfying these axioms. Our results show that the existence of path independent or contraction consistent social choice correspondences depends not only on which philosophical position we take but also on what is the precise notion of equity.

There are many cases in the real life in which we must consider multiple criteria in individual or social decision-making problems. To explore general conditions for lexicographic compositions of multiple criteria to satisfy various degrees of choice-consistency may be an interesting topic in future research.

Acknowledgments This work is dedicated to Kotaro Suzumura with my deep gratitude for his invaluable support and advice over many years. I also thank a referee on this paper for many useful comments. Financial support from the Ministry of Education, Culture, Sports, Science and Technology, Japan, through the grant for the 21st Century Center of Excellence Program on the Normative Evaluation and Social Choice of Contemporary Economic Systems and the Grant-in-Aid for Scientific Research (B) No. 18330036 is gratefully acknowledged.

References

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York Wiley (1st ed. 1951)
Chernoff, H. (1954). Rational selection of decision functions. *Econometrica*, *22*, 442–443
Feldman, A. M., and Kirman, A. (1974). Fairness and envy. *American Economic Review*, *64*, 996–1005
Foley, D. (1967). Resource allocation and the public sector. *Yale Economic Essays*, *7*, 45–98
Kolm, S.-C. (1972). *Justice et equité*. Paris: Centre National de la Recherche Scientifique
Pazner, E. A., and Schmeidler, D. (1978). Egalitarian equivalent allocations: A new concept of economic equity. *Quarterly Journal of Economics*, *92*, 671–687
Plott, C. R. (1973). Path independence, rationality, and social choice. *Econometrica*, *41*, 1075–1091
Sen, A. K. (1970). *Collective choice and social welfare*. San Francisco: Holden-Day
Suzumura, K. (1981a). On Pareto-efficiency and the no-envy concept of equity. *Journal of Economic Theory*, *25*, 367–379
Suzumura, K. (1981b). On the possibility of 'fair' collective choice rule. *International Economic Review*, *22*, 351–364
Suzumura, K. (1983). *Rational choice, collective decisions, and social welfare*. Cambridge, UK: Cambridge University Press
Suzumura, K. (1996). Interpersonal comparisons of the extended sympathy type and the possibility of social choice. In K. J. Arrow, A. K. Sen, and K. Suzumura (Eds.), *Social choice re-examined, volume 2* (pp. 202–229). London: Macmillan
Suzumura, K. (2004). An extension of Arrow's lemma with economic applications. COE/RES Discussion Paper Series No. 79, Hitotsubashi University
Tadenuma, K. (2002). Efficiency first or equity first? Two principles and rationality of social choice. *Journal of Economic Theory*, *104*, 462–472
Tadenuma, K. (2005). Egalitarian-equivalence and the Pareto principle for social preferences. *Social Choice and Welfare*, *24*, 455–473
Thomson, W. (1996). Concepts of implementation. *Japanese Economic Review*, *47*, 133–143

Characterization of the Maximin Choice Function in a Simple Dynamic Economy

Koichi Suga and Daisuke Udagawa

1 Introduction

In the literature of intergenerational equity, Rawlsian maximin principle is one of the most well-known criteria for distributive justice among generations.[1] Since this principle has an intuitive appeal to egalitarian writers, several attempts to characterize the principle have been made in welfare economics. Arrow (1973), Dasgupta (1974a, b), and Riley (1976) scrutinized the performance thereof in the context of optimal growth. Arrow shows that the utility path as well as the consumption path generated by the maximin principle has a saw-tooth shape. Dasgupta shows that it gives rise to a logical deficit such as time-inconsistency. The other line of researches has been stimulated by the axiomatic approaches of Hammond (1976, 1979) and Sen (1970, 1977). In this line, researchers extended axiomatizations of the maximin principle and applied them to intergenerational equity. The maximin path is characterized by a constant path, which emphasizes its egalitarian perspective.[2]

In a previous discussion on this topic, Suga and Udagawa (2004) addressed the question of how to characterize the maximin principle axiomatically in a simple

K. Suga
School of Political Science and Economics, Waseda University, Shinjuku-ku, Tokyo 169-8050, Japan
e-mail: ksuga@waseda.jp

D. Udagawa
Graduate School of Economics, Waseda University, Shinjuku-ku, Tokyo 169-8050, Japan
e-mail: udagawa_daisuke@suou.waseda.jp

[1] See Rawls (1971) for the description of the maximin principle. Rawls himself denies a direct application of the principle, and emphasizes that each generation has a paternalistic concern to the payoff of his immediate descendant. Since the number of generations are infinite so that the minimum utility may not exist, we usually evaluate consumption plans not by the minimum but by the infimum.

[2] See, for example, Asheim, Buchholz, and Tungodden (2001), Epstein (1986a,b) and Lauwers (1997).

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

dynamic economy, called Arrow–Dasgupta economy, where each generation has a paternalistic concern to the descendants. In Suga and Udagawa (2004), the axioms are imposed on intergenerational preference relations over the set of consumption paths. They supposed that there exists a hypothetical social planner who judges the consumption paths, and characterized the maximin principle by some axioms on the planner's intergenerational preference relation.

In this chapter, on the other hand, we examine the same question of axiomatic characterization of the maximin principle by applying axioms in a choice function framework. That is, we consider a choice-theoretic model of infinite horizon economy in which a choice function selects a consumption path from the set of feasible paths from the viewpoint of a social planner. We suppose that the social planner adopts the maximin principle as a criterion to construct an intertemporal choice function. We focus our attention to a simple dynamic economy with linear technology à la Arrow (1973) and Dasgupta (1974a,b) to characterize the maximin choice function on the set of consumption paths. We employ this choice-theoretic approach to give another look at characteristics of consumption paths derived by the maximin principle under the feasibility conditions.

With a similar motivation, Asheim, Bossert, Sprumont, and Suzumura (2006) propose a choice-theoretic model for intergenerational equity. They provide characterizations of all infinite-horizon choice functions satisfying either efficiency or time-consistency, and identify all choice functions with both properties. Their results show that the choice-theoretic approach to intergenerational resource allocation provides an interesting and viable alternative to the models based on establishing intergenerational preference relations of utility paths.

Our purpose in this chapter is to characterize the maximin principle in an infinite horizon economy. Axioms are imposed not on intergenerational preference relations but on choice functions themselves. Some of the axioms are similar to those in characterization of the maximin principle in intra-generational equity, that is, Pareto principle and extended Hammond equity. Others are conditions α and β, which are often used in choice theory to describe consistent choices.

The chapter is organized as follows. Section 2 is the description of the economy, which provides a canvas for our analysis. Axioms are stated in Sect. 3. The main theorem, the lemmas, and their proofs are contained in Sect. 4. Section 5 provides related examples. We conclude the chapter with some final remarks in Sect. 6.

2 Simple Dynamic Economy

Let $\mathcal{Z}_+$ be the set of all nonnegative integers, each element of which is used to represent a generation or time period. For simplicity, we assume that each time period consists of one generation, and each generation consists of one representative individual. There is a private good, which can either be consumed or invested to be capital that bears a return. k_t denotes the accumulated capital at the beginning of time period $t \in \mathcal{Z}_+$. In that period a fraction x_t is consumed and the remainder $k_t - x_t$ is used in production. The production technology is assumed to be linear. Then each unit used in production brings γ units of the good at the end of the period, and are transferred to the next period $t+1$. Hence

$$k_{t+1} = \gamma(k_t - x_t). \tag{1}$$

We assume that the economy is productive, so that

$$\gamma > 1. \tag{2}$$

The following feasibility condition for production is assumed. For all $t \in \mathcal{Z}_+$

$$k_t \geq 0. \tag{3}$$

A feasibility condition for consumption is also assumed. That is, any individual cannot survive without consumption. Hence, for all $t \in \mathcal{Z}_+$

$$x_t \geq 0.^3 \tag{4}$$

Now we describe our problem to find a consumption path that is selected by Rawlsian maximin principle for intergenerational justice. For the convenience of description, we adopt the following notation: let $\mathcal{L}_+^\infty = \{(x_0, x_1, \ldots, x_t, \ldots) | \ \forall t \in \mathcal{Z}_+ : x_t \geq 0\}$. Denote a consumption path by the capital letter, for example, $X = (x_0, x_1, \ldots)$. $_{\text{rep}}(x_1, \ldots, x_n)$ represents the path $(x_1, \ldots, x_n,\ x_1, \ldots, x_n, \ldots)$, which consists of $(x_1, \ldots, x_n)$ repeated infinitely many times. By the feasibility condition, consumption paths ought to be chosen from the set $\mathcal{X} = \{X \in \mathcal{L}_+^\infty | \ \forall t \in \mathcal{Z}_+ :\ 0 \leq k_{t+1} = \gamma(k_t - x_t)\}$ given $k_0 > 0$.[4] It is convenient, however, to use the following equivalent form: for any given k_0 and γ, the set of feasible consumption paths are given by

$$\mathcal{X} = \left\{ X = (x_0, x_1, \ldots) \in \mathcal{L}_+^\infty \ \middle| \ \sum_{t=0}^{\infty} \gamma^{-t} x_t \leq k_0 \right\}.$$

We denote the utility function of generation $t \in \mathcal{Z}_+$, or often called individual t, by $W_t(X)$ when the consumption path X is attained. We assume that generation t derives utility from her own consumption x_t and also from her immediate $n-1$ descendants' satisfaction, where $n \geq 2$, so that her utility function depends on the consumption stream of n generations beginning with her own. We also assume that the utility function W_t is the same for all generations $t \in \mathcal{Z}_+$, that is, $W_t = W$ for all t. Following Arrow (1973) and Dasgupta (1974a, b), we assume that W is additively separable as to generations for simplicity, that the felicity ascribed by individual t to individual $t+i$ is the same as that ascribed by individual $t+i$ to herself, that the

[3] Any path on which at least one generation survives is meaningful for the discussion of intergenerational equity. It is possible for a generation to exhaust the whole amount of the good inherited from the past. Here we impose a mild requirement on the feasibility of the consumption.

[4] $\mathcal{X}$ depends on the initial capital stock k_0. But in the following discussion, k_0 is given from the outside at the outset so that we employ the notation $\mathcal{X}$ instead of $\mathcal{X}(k_0)$.

felicity function is the same for all t, and that the felicity of the future generations is discounted in the utility of the present generation. That is,

$$W_t(X) = W(x_t, x_{t+1}, \ldots, x_{t+n-1}) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}), \tag{5}$$

where $\rho_0 = 1$ and ρ_i $(1 \leq i \leq n-1)$ are a parameters reflecting the weight each generation attach to the future generations. We assume that the weight of a farther future generation is smaller, that is, $\rho_i \geq \rho_{i+1}$ $(0 \leq i \leq n-1)$. The felicity function U is assumed to satisfy the following conditions: (a) $U : \Re_+ \to \Re$ is twice continuously differentiable; (b) $U'(x) > 0$ and $U''(x) < 0$.

We focus our concern on the case in which the optimal consumption path for the maximin principle has a saw-tooth shape.[5] Therefore, we assume

$$\gamma^i \rho_i < \gamma^j \rho_j \quad (0 \leq i < j \leq n). \tag{6}$$

This assumption requires that each generation obtains more utility if she bequeaths capital to the next generation than that if she consumes it by herself. Although the utility of the next generation is discounted by ρ, the total utility will go up if the increase in production is included.

Then, the maximin principle of justice gives a solution to the problem

$$\begin{aligned} \max_{X \in S} \min_t W_t(X) &\Longleftrightarrow \max_{X \in S} \min_t W(x_t, x_{t+1}, \cdots, x_{t+n-1}), \\ &\Longleftrightarrow \max_{X \in S} \min_t \sum_{i=0}^{n-1} \rho_i U(x_{t+i}), \end{aligned} \tag{7}$$

where S is any subset of $\mathcal{X}$.

Now we present Arrow's theorem on the maximin path. Let $\hat{x}$ be the consumption level that allows to bequeath the same amount of capital as the initial level to the next generation, that is,

$$\hat{x} = \frac{\gamma - 1}{\gamma} k_0. \tag{8}$$

Clearly, the consumption path ${}_{\text{rep}}(\hat{x})$ satisfies the feasibility condition. In other words, the constant consumption $\hat{x}$ will cause k_t to remain constant at the initial level k_0.

Let $(x_0^{\mathrm{R}}, x_1^{\mathrm{R}}, \ldots, x_n^{\mathrm{R}})$ be the solution to the problem

$$\max_{x_0, x_1, \ldots, x_{n-1}} \sum_{i=0}^{n-1} \rho_i U(x_i), \tag{9}$$

$$\text{s.t.} \quad \sum_{i=0}^{n-1} \gamma^{-i} x_i = \hat{x} \sum_{i=0}^{n-1} \gamma^{-i}. \tag{10}$$

[5] The inequality (6) is a necessary condition for the case. See Arrow (1973).

Equation (10) is equivalent to the condition $k_n = k_0$. Therefore, this problem can be interpreted as the maximization problem of generation 0's utility, subject to the restriction that generation 0 must bequeath k_0 to generation n.

Then we have the following theorem by Arrow (1973), which is the most fundamental proposition in this field.

Theorem 1. [Arrow (1973); Theorem 3] *Suppose that the utility of any generation t is given by*

$$W_t(X) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is nonincreasing in i. Then the feasible consumption path that maximizes $\min_t W_t$ can be characterized as follows. Choose x_i^* $(i = 0, \ldots, n-1)$ to maximize $W(x_0, \ldots, x_{n-1})$ subject to the constraint

$$\sum_{i=0}^{n-1} \gamma^{-i} x_i = \hat{x} \sum_{i=0}^{n-1} \gamma^{-i},$$

where $\hat{x}$ is given in (8). Then at the optimum (i) $x_{nl+i} = x_i^*$ $(0 \leq i \leq n-1)$ for any $l \in \mathcal{Z}_+$. For this path the following properties hold: (ii) $x_i^* < x_{i+1}^*$; (iii) $W_t = \min_t W_t$, if t is divisible by n; (iv) for all other t, $W_t \geq \min_t W_t$, and (v) the inequality is strict if $\rho_i > \rho_{i+1}$ for some $i < n-1$.

We define a choice function C that maps any nonempty set $S \subseteq \mathcal{X}$ of feasible consumption paths to its subset, given a utility function W. Because W is given and fixed throughout this chapter, a choice function is denoted by $C(S)$. We define the Rawlsian choice function C^{R}, which maps any feasible set S of consumption paths to the set of all maximin consumption paths X^{R} in S, given a utility function W. It is not generally true that $C^{\mathrm{R}}(S) \neq \varnothing$ for all $S \subseteq \mathcal{X}$ and W, but Arrow (1973) showed $C^{\mathrm{R}}(\mathcal{X}) \neq \varnothing$ under the above utility function W.

3 Axioms

In this section, we define several axioms for a characterization of the maximin principle in this simple dynamic economy.[6] First, we define two binary relations on $\mathcal{X}$. One is the strict Paretian relation, $\succ^{\mathrm{P}}$, which is given by: for any $X^1, X^2 \in \mathcal{X}$,

$$X^1 \succ^{\mathrm{P}} X^2 \iff \forall t : W_t(X^1) > W_t(X^2).$$

Another is the Hammond equity relation, $\succsim^{\mathrm{H}}$, which is defined as, for any $X^1, X^2 \in \mathcal{X}$

[6] See Hammond (1976, 1979), Sen (1970), and Suzumura (1983) to understand the meanings of the axioms in the classical environment. In the following definitions, we employ those in Suzumura (1983).

$$X^1 \succsim^{\mathrm{H}} X^2 \iff \exists t^1, t^2 \in \mathcal{Z}_+ : \begin{array}{l} \text{(i) } W_{t^1}(X^1) \leq W_{t^2}(X^1), \\ \text{(ii) } W_{t^1}(X^1) \geq W_{t^1}(X^2), \\ \text{(iii) } W_{t^2}(X^1) \leq W_{t^2}(X^2), \text{and} \\ \text{(iv) } W_t(X^1) = W_t(X^2) \; \forall t \neq t^1, t^2. \end{array}$$

By extending the Hammond equity principle, we introduce a new concept of equity among groups of generations. It is called the extended Hammond equity principle, which implies a fairness requirement that we should treat two groups of generations equally if they are regarded equal in utility profiles. As an auxiliary step, we follow Suzumura (1983) to introduce the lexicographic ordering R^{L} on the Euclidean n-space E^n. For every $v \in E^n$, let $i(v)$ denote the ith smallest element, ties being broken arbitrarily, so that we have

$$v_{1(v)} \leq v_{2(v)} \leq \cdots \leq v_{n(v)}.$$

We may then define three binary relations P^{L}, I^{L}, and R^{L} on E^n by

$$v^1 P^{\mathrm{L}} v^2 \iff \exists r \leq n : \begin{cases} \forall i \in \{1, 2, \ldots, r-1\} : v^1_{i(v^1)} = v^2_{i(v^2)} \\ \& \\ v^1_{r(v^1)} > v^2_{r(v^2)}, \end{cases}$$

$$v^1 I^{\mathrm{L}} v^2 \iff \forall i \in \{1, 2, \ldots, r-1\} : v^1_{i(v^1)} = v^2_{i(v^2)},$$

and

$$v^1 R^{\mathrm{L}} v^2 \iff v^1 P^{\mathrm{L}} v^2 \;\; or \;\; v^1 I^{\mathrm{L}} v^2 \;\; \text{for all } v^1, v^2 \in E^n.$$

We are now in the position of defining an axiom for extended Hammond equity. Take any two groups of generations G_1, G_2, which consist of finite number n of successive generations. For any consumption path X^1, X^2, we have two n-dimensional vectors $(W_t(X^1))_{t \in G_1}$ and $(W_t(X^2))_{t \in G_2}$. With this notation we define an extension of Hammond equity relations in the case of sympathy to $n-1$ future generations. The strict extended Hammond relation, $\succ^{\mathrm{EH}}$, is defined by: for any $X^1, X^2 \in \mathcal{X}$, $X^1 \succ^{\mathrm{EH}} X^2$ if and only if there exist t^r and t^p such that

(i) $W_t(X^2) \geq W_t(X^1)$ $\quad (t = t^r - (n-1), \ldots, t^r)$,
(ii) $W_t(X^1) \geq W_t(X^2)$ $\quad (t = t^p - (n-1), \ldots, t^p)$,
(iii) $W_t(X^1) = W_t(X^2)$ $\quad$ (otherwise), and
(iv) $(W_t(X^1))_{t = t^r - (n-1), \ldots, t^r} P^{\mathrm{L}} (W_t(X^2))_{t = t^p - (n-1), \ldots, t^p}$,

where $W_t(X^i) = W_0(X^i)$ for $t < 0$, $i = 1, 2$. The extended Hammond indifference relation, $\sim^{\mathrm{EH}}$, is defined by: for any $X^1, X^2 \in \mathcal{X}$, $X^1 \sim^{\mathrm{EH}} X^2$ if and only if $(W_t(X^1))_{t \in G_1} I^{\mathrm{L}} (W_t(X^2))_{t \in G_2}$ holds. The extended Hammond equity relation, $\succsim^{\mathrm{EH}}$, is defined by: for any $X^1, X^2 \in \mathcal{X}$

$$X^1 \succsim^{\mathrm{EH}} X^2 \iff X^1 \succ^{\mathrm{EH}} X^2 \; or \; X^1 \sim^{\mathrm{EH}} X^2.$$

This relation represents a concept of equity between two groups of successive generations, which is applied to the case where a change in the consumption of a generation causes a change in the utilities of the whole group. The reason why we need this type of requirement is that a change in the consumption of some generation under the feasibility of the economy brings increase in utility to a group of successive generations and decrease to another group that does not satisfy the conditions presupposed in the definition of the Hammond equity relation.

Now we provide five axioms. The first axiom simply requires non-emptiness of the choice set from the set of all feasible paths.

Definition 1. *A choice function C satisfies nonempty choice from $\mathcal{X}$ (NE) iff* $C(\mathcal{X}) \neq \varnothing$.

The second axiom is a requirement that if a consumption path in a feasible set S is extended Hammond superior to a path in the choice set $C(S)$, then it is also included in $C(S)$.

Definition 2. *A choice function C satisfies inclusion of extended Hammond superior paths (IEH) iff* $\forall X^1, X^2 \in \mathcal{X}\ \forall S \subseteq \mathcal{X}$*:*

$$[X^1 \succsim^{\mathrm{EH}} X^2 \ \&\ X^1 \in S \ \&\ X^2 \in C(S)] \Rightarrow X^1 \in C(S).$$

The third axiom is a requirement that a path which is Pareto inferior to another path in a feasible set S is excluded from the choice set $C(S)$.

Definition 3. *A choice function C satisfies exclusion of Pareto inferior paths (EP) iff* $\forall X^1, X^2 \in \mathcal{X}\ \forall S \subseteq \mathcal{X}$*:*

$$[X^1 \succ^{\mathrm{P}} X^2 \ \&\ X^1 \in S] \Rightarrow X^2 \notin C(S).$$

The next two axioms are conditions of consistency for the choice sets. The fourth axiom is a requirement that any path in the choice set for a larger feasible set is also included in the choice set for a smaller feasible set if the path belongs to that set.

Definition 4. *A choice function C satisfies condition α iff* $\forall S^1, S^2 \subseteq \mathcal{X}, S^1 \subseteq S^2 :$ $\forall X^1 \in S^1$*:*

$$X^1 \in C(S^2) \Rightarrow X^1 \in C(S^1).$$

The fifth axiom is a requirement that if a path in the choice set for a smaller feasible set is included in the choice set for a larger feasible set, then any other path in the choice set for the smaller feasible set is also included in the choice set for the larger feasible set.

Definition 5. *A choice function C satisfies condition β iff* $\forall S^1, S^2 \subseteq \mathcal{X}, S^1 \subseteq S^2 :$ $\forall X^1 \in S^1, X^2 \in S^2 :$

$$[X^1 \in C(S^1) \cap C(S^2) \ \&\ X^2 \in C(S^1)] \Rightarrow X^2 \in C(S^2).$$

4 Main Theorem

We are in the position to provide our main theorem about the characterization of the Rawlsian choice function.

Lemma 1. *Suppose that the utility of any generation t is given by*

$$W_t(X) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is non increasing in i. If a choice function C satisfies NE, EP, IEH, conditions α and β, then,

$$C(\mathcal{X}) = C^{\mathrm{R}}(\mathcal{X}).$$

To prove this lemma we need Lemmas 2.–4..

Lemma 2. *Suppose that the utility of any generation t is given by*

$$W_t(X) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is nonincreasing in i. If a choice function C satisfies NE, EP, IEH, α, and β, then $W_0(X) = \min_t W_t(X)$ for all $X \in C(\mathcal{X})$.

Proof. By NE, $C(\mathcal{X}) \neq \varnothing$. Suppose that $X^* \in C(\mathcal{X})$ and that $W_0(X^*) \neq \min_t W_t(X^*)$. There are two cases to be considered: (i) there exists $\min_t W_t(X^*)$ and $\min_t W_t(X^*) < W_0(X^*)$; or (ii) there does not exist $\min_t W_t(X^*)$. In both cases, we can find some generation enjoying less welfare than generation 0. Let t^m be such generation. For any $q \in (0,1)$, we can construct a feasible consumption path X^1 defined as follows:

$$\begin{cases} x_0^1 = x_0^* - \varepsilon \\ x_{t^m+n-1}^1 = x_{t^m+n-1}^* + q\varepsilon\gamma^{t^m+n-1} \\ x_t^1 = x_t^* \qquad (t \neq 0,\ t^m+n-1). \end{cases}$$

For sufficiently small $\varepsilon > 0$, we have the following:

$$\begin{cases} W_0(X^1) < W_0(X^*), \\ W_t(X^1) > W_t(X^*), \qquad (t^m \leq t \leq t^m+n-1), \\ W_0(X^1) > W_{t^m}(X^1), \\ W_t(X^1) = W_t(X^*), \qquad (0 < t < t^m \text{ or } t^m+n-1 < t). \end{cases}$$

Then, by the definition of the extended Hammond relation,

$$X^1 \succsim^{\mathrm{EH}} X^*. \tag{11}$$

In making the path X^1 from X^* there remains an amount of the consumption good $(1-q)\varepsilon$. If we increase the consumption by $\delta > 0$ for each generation by dividing the amount $(1-q)\varepsilon$, the equality

$$\delta(1+\gamma^{-1}+\gamma^{-2}+\cdots)=\delta\sum_{t=0}^{\infty}\gamma^{-t}=(1-q)\varepsilon$$

must hold. Hence, we can construct a feasible consumption path X^2 defined as follows:

$$x_t^2=x_t^1+(1-q)\frac{\varepsilon}{\sum_{t=0}^{\infty}\gamma^{-t}}$$

for all $t\geq 0$. Then

$$X^2\succ^{\mathrm{P}}X^1. \tag{12}$$

By (12) and condition EP,

$$X^1\notin C(\{X^*,X^1,X^2\}). \tag{13}$$

Now we show $X^*\notin C(\{X^*,X^1,X^2\})$. Suppose on the contrary that $X^*\in C(\{X^*,X^1,X^2\})$. Then we obtain $X^1\in C(\{X^*,X^1,X^2\})$ by (11) and condition IEH. Applying condition α to this relation,

$$X^1\in C(\{X^*,X^1\}). \tag{14}$$

Since $X^*\in C(\mathcal{X})$, we have

$$X^*\in C(\{X^*,X^1\}), \tag{15}$$

with the help of $\{X^*,X^1\}\subseteq\mathcal{X}$ and condition α. Equations (14) and (15) together imply

$$\{X^*,X^1\}=C(\{X^*,X^1\}). \tag{16}$$

Under the assumption $X^*\in C(\{X^*,X^1,X^2\})$, (16) with condition β implies

$$X^1\in C(\{X^*,X^1,X^2\}),$$

which contradicts (13). Therefore

$$X^*\notin C(\{X^*,X^1,X^2\}) \tag{17}$$

must hold.

On the other hand, $X^*\in C(\mathcal{X})$ implies that $X^*\in C(\{X^*,X^1,X^2\})$ with the help of condition α. This contradicts (17). Hence $W_0(X)=\min_t W_t(X)$ for any $X\in C(\mathcal{X})$. □

Lemma 3. *Suppose that the utility of any generation t is given by*

$$W_t(X)=\sum_{i=0}^{n-1}\rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is nonincreasing in i. If a choice function C satisfies NE, EP, IEH, conditions α, and β, then generation 0 in $X^* \in C(\mathcal{X})$ has the largest welfare among all feasible consumption paths where generation 0 has the least welfare among all the generations. That is,

$$W_0(X^*) = \max_{X \in \mathcal{D}_0} W_0(X)$$

for any $X^* \in C(\mathcal{X})$, where $\mathcal{D}_0 = \{X \in \mathcal{X} \mid W_0(X) = \min_t W_t(X)\}$.

Proof. By NE, $C(\mathcal{X}) \neq \varnothing$. Suppose, on the contrary, that $X^* \in C(\mathcal{X})$ and that there is $X^{**} \in \mathcal{D}_0$ such that $W_0(X^*) < W_0(X^{**})$. Let $X^{***} = \arg\max_{X \in \mathcal{D}_0} W_0(X)$. Then $W_0(X^*) < W_0(X^{**}) \leq W_0(X^{***})$. Hence, without loss of generality, we assume $X^{***} = X^{**}$. By the feasibility condition and assumptions of U, $(x_0^{**}, \ldots, x_{n-1}^{**})$ is the unique solution of the problem:

$$\max_{x_0,\ldots,x_{n-1}} \sum_{t=0}^{n-1} \rho_t U(x_t)$$
$$\text{s.t.} \quad \sum_{t=0}^{n-1} \gamma^{-t} x_t = \hat{x} \sum_{t=0}^{n-1} \gamma^{-t}.$$

Therefore, $(x_0^{**}, \ldots, x_{n-1}^{**})$ is characterized by

$$\mathrm{MRS}(x_i^{**}, x_j^{**}) = \frac{\rho_i U'(x_i^{**})}{\rho_j U'(x_j^{**})} = \gamma^{j-i},$$

for any $0 \leq i < j \leq n-1$.

Two cases should be distinguished.

Case 1: $x_i^* < x_i^{**}$.

The assumptions of U assure that

$$\mathrm{MRS}(x_i^*, x_j^*) = \frac{\rho_i U'(x_i^*)}{\rho_j U'(x_j^*)} > \gamma^{j-i}.$$

Consider X^1 and X^2 defined as follows:

$$\begin{cases} x_i^1 = x_i^* + q\varepsilon, \\ x_j^1 = x_j^* - \gamma^{j-i}\varepsilon, \\ x_t^1 = x_t^* \qquad (t \neq i, j), \end{cases}$$

and for all $t \geq 0$

$$x_t^2 = x_t^1 + (1-q)\frac{\varepsilon}{\sum_{t=0}^{\infty} \gamma^{-t}},$$

where $0 < q < 1$, $\varepsilon > 0$. Then

$$X^2 \succ^{\mathrm{P}} X^1. \tag{18}$$

By (18) and condition EP,

$$X^1 \notin C(\{X^*, X^1, X^2\}). \tag{19}$$

For sufficiently small ε,

$$\begin{cases} W_t(X^1) > W_t(X^*), & (t = i-(n-1), \ldots, i), \\ W_t(X^1) < W_t(X^*), & (t = j-(n-1), \ldots, j), \\ W_t(X^1) = W_t(X^*), & (otherwise) \end{cases}$$

hold so that $(W_t(X^1))_{t=j-(n-1),\ldots,j} P^{\mathrm{L}} (W_t(X^1))_{t=i-(n-1),\ldots,i}$ by the continuity of W. Hence we have $X^1 \succsim^{\mathrm{EH}} X^*$ by the extended Hammond equity.

Now we show $X^* \notin C(\{X^*, X^1, X^2\})$. Suppose on the contrary that $X^* \in C(\{X^*, X^1, X^2\})$. Then we obtain $X^1 \in C(\{X^*, X^1, X^2\})$ by (11) and condition IEH. Applying condition α to this relation,

$$X^1 \in C(\{X^*, X^1\}). \tag{20}$$

Since $X^* \in C(\mathcal{X})$, we have

$$X^* \in C(\{X^*, X^1\}) \tag{21}$$

with the help of $\{X^*, X^1\} \subseteq \mathcal{X}$ and condition α. Equations (20) and (21) together imply

$$\{X^*, X^1\} = C(\{X^*, X^1\}). \tag{22}$$

Under the assumption $X^* \in C(\{X^*, X^1, X^2\})$, (16) with condition β implies

$$X^1 \in C(\{X^*, X^1, X^2\}),$$

which contradicts (19). Therefore

$$X^* \notin C(\{X^*, X^1, X^2\}) \tag{23}$$

must hold.

On the other hand, $X^* \in C(\mathcal{X})$ implies that $X^* \in C(\{X^*, X^1, X^2\})$ with the help of condition α. This contradicts (23). Hence this case cannot be true.

Case 2: $x_i^* > x_i^{**}$.

The assumptions of U assure that

$$\mathrm{MRS}(x_i^*, x_j^*) = \frac{\rho_i U'(x_i^*)}{\rho_j U'(x_j^*)} < \gamma^{j-i}.$$

Define X^1 and X^2 as follows:

$$\begin{cases} x_i^1 = x_i^* - \varepsilon, \\ x_j^1 = x_j^* + q\gamma^{j-i}\varepsilon, \\ x_t^1 = x_t^* \qquad (t \neq i, j), \end{cases}$$

and for all $t \geq 0$

$$x_t^2 = x_t^1 + (1-q)\frac{\varepsilon}{\sum_{t=0}^{\infty} \gamma^{-t}},$$

where $0 < q < 1$, $\varepsilon > 0$. Then, by the same reasoning as *Case* 1, we come to the contradiction that $X^* \in C(\{X^*, X^1, X^2\})$ and $X^* \notin C(\{X^*, X^1, X^2\})$. Hence this case cannot be true either.

By *Cases* 1 and 2, we have a contradiction for any $0 \leq i < j \leq n-1$ if $\mathrm{MRS}(x_i^*, x_j^*) \neq \gamma^{j-i}$. Therefore, generation 0 in $X^* \in C(\mathcal{X})$ has the largest welfare among all feasible consumption paths where generation 0 has the least welfare among all the generations. □

The next lemma shows a sufficient condition for a consumption path to be infeasible. The idea of the proof is due to Lemma 1 in Arrow (1973).

Lemma 4. *Suppose that the utility of any generation t is given by*

$$W_t(X) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is nonincreasing in i. If a consumption path X satisfies that $\sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln} \geq \sum_{s=0}^{n-1} \gamma^{-s} x_s^{\mathrm{R}}$ for all $l \in \mathcal{Z}_+$ and $\sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln} > \sum_{s=0}^{n-1} \gamma^{-s} x_s^{\mathrm{R}}$ for some $l' \in \mathcal{Z}_+$, then X is infeasible.

Proof. By the feasibility condition, the relation between k_{ln} and $k_{(l+1)n}$ can be written as

$$k_{(l+1)n} = \gamma^n \left(k_{ln} - \sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln} \right). \tag{24}$$

On the other hand, a Rawlsian maximal consumption path satisfies the condition that $k_n = k_0$. Hence

$$k_0 = \gamma^n \left(k_0 - \sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln}^{\mathrm{R}} \right). \tag{25}$$

By (24) and (25),

$$k_{(l+1)n} - k_0 = \gamma^n \left[(k_{ln} - k_0) - \left(\sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln} - \sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln}^{\mathrm{R}} \right) \right].$$

For simplicity of description, define h_l and a_l as follows:

$$h_l = \sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln} - \sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln}^{\mathrm{R}},$$
$$a_l = \gamma^{-ln}(k_{ln} - k_0).$$

Then, $\gamma^{(l+1)n} a_{l+1} = \gamma^n(\gamma^{ln} a_l - h_l)$ iff $a_{l+1} = a_l - \gamma^{-ln} h_l$. Since $a_0 = 0$ by definition, $a_l = \sum_{u=0}^{l-1} \gamma^{-nu} h_u$ is true. Now, by the assumptions of lemma and definitions of h_l and a_l, the following inequality holds:

$$\limsup_{l \to \infty} \sum_{u=0}^{l-1} \gamma^{-nu} h_u > 0.$$

Then, for some $\varepsilon > 0$, there is sufficiently large $\bar{l}$ such that $a_l < -\varepsilon \quad \text{for} \quad l \geq \bar{l}$. Since $\gamma^{-ln} k_0 < \varepsilon$ for any sufficiently large l, there exists some $l' \in \mathcal{Z}_+$ such that $a_{l'} < -\gamma^{l'n} k_0$. Hence $k_{l'n} < 0$, and X is infeasible. □

Now we provide the proof of lemma 1. with these lemmas.

Proof of lemma 1.:

By NE, $C(\mathcal{X}) \neq \varnothing$. Let X^* be any consumption path in $C(\mathcal{X})$. By Lemma 3. and Theorem 1, $W_0(X^*) \geq W_0(X^{\mathrm{R}})$. Suppose that $W_0(X^*) > W_0(X^{\mathrm{R}})$. Then $W_{ln}(X^*) > W_0(X^{\mathrm{R}})$ for all $l \in \mathcal{Z}_+$ by Lemma 2.. Since $(x_0^{\mathrm{R}}, \ldots, x_{n-1}^{\mathrm{R}})$ is the unique solution of the maximization problem (9) and (10) and the consumption path X^{R} is an infinite repetition of $(x_0^{\mathrm{R}}, \ldots, x_{n-1}^{\mathrm{R}})$ by Theorem 1,

$$\sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln}^* > \sum_{s=0}^{n-1} \gamma^{-s} x_s^{\mathrm{R}}$$

holds for all $l \in \mathcal{Z}_+$. Hence, X^* is infeasible by Lemma 4., which is a contradiction. Therefore, we have $W_0(X^*) = W_0(X^{\mathrm{R}})$.

Since, for all $l \in \mathcal{Z}_+$,

$$W_{ln}(X^*) \geq \min W_t(X^*) = W_0(X^*) = W_0(X^{\mathrm{R}})$$

holds,

$$\sum_{s=0}^{n-1} \gamma^{-s} x_{s+ln}^* \geq \sum_{s=0}^{n-1} \gamma^{-s} x_s^{\mathrm{R}}, \tag{26}$$

for all $l \in \mathcal{Z}_+$. By Theorem, 1 $(x_0^{\mathrm{R}}, \ldots, x_{n-1}^{\mathrm{R}})$ is the unique maximum of the problem (9) and (10). Hence equality (26) holds only when $(x_{ln}^*, \ldots, x_{(l+1)n-1}^*) = (x_0^{\mathrm{R}}, \ldots, x_{n-1}^{\mathrm{R}})$. Therefore, $X^* = {}_{\mathrm{rep}}(x_0^{\mathrm{R}}, \ldots, x_{n-1}^{\mathrm{R}}) = X^{\mathrm{R}}$. □

We provide the converse of Lemma 1..

Lemma 5. *Suppose that the utility of any generation t is given by*

$$W_t(X) = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

$\gamma^i \rho_i$ increases with i for $i \leq n-1$, and ρ_i is nonincreasing in i. The Rawlsian choice function C^{R} satisfies NE, IEH, EP, α, and β.

Proof. NE: As noted in Sect. 2, Arrow (1973) showed $C^{\mathrm{R}}(\mathcal{X}) \neq \varnothing$.

IEH: Suppose that (i) $S \subseteq \mathcal{X}$, (ii) $X^1 \in S$, (iii) $X^2 \in C^{\mathrm{R}}(S)$, and (iv) $X^1 \succsim^{\mathrm{EH}} X^2$. By (iv), $\min_t W_t(X^1) \geq \min_t W_t(X^2)$. Then $X^1 \in C^{\mathrm{R}}(S)$, so that IEH holds.

EP: Suppose that (i) $S \subseteq \mathcal{X}$, (ii) $X^1 \in S$, and (iii) $X^1 \succ^{\mathrm{P}} X^2$. Then, by (iii), $\min_t W_t(X^1) > \min_t W_t(X^2)$. So $X^2 \notin C^{\mathrm{R}}(S)$ and EP holds.

Condition α: Suppose that (i) $S^1 \subseteq S^2 \subseteq \mathcal{X}$, (ii) $X^1 \in C^{\mathrm{R}}(S^2)$, and (iii) $X^1 \in S^1$. By (i) and (ii), $\min_t W_t(X^1) \geq \min_t W_t(X)$ for all $X \in S^1$. Hence, we obtain $X^1 \in C^{\mathrm{R}}(S^1)$, and condition α holds.

Condition β: Suppose that (i) $S^1 \subseteq S^2 \subseteq \mathcal{X}$, (ii) $X^1, X^2 \in C^{\mathrm{R}}(S^1)$. By (ii), $\min_t W_t(X^1) = \min_t W_t(X^2)$. Therefore, we have $X^1 \in C^{\mathrm{R}}(S^2) \iff X^2 \in C^{\mathrm{R}}(S^2)$. □

With Lemmas 1. and 5., we finally come to the following characterization theorem.

Theorem 1. *Suppose that the utility of any generation t is given by*

$$W_t = \sum_{i=0}^{n-1} \rho_i U(x_{t+i}),$$

that $\gamma^i \rho_i$ increases with i for $i \leq n-1$, and that ρ_i is nonincreasing in i. Then, (i) the Rawlsian choice function C^{R} satisfies NE, EP, IEH, α, and β; and (ii) if a choice function C satisfies NE, EP, IEH, α, and β, then $C(\mathcal{X}) = C^{\mathrm{R}}(\mathcal{X})$.

5 Related Examples

As for independence of the axioms we must examine in five cases whether there exists a choice function that satisfies all but one axiom. In the following, we show only three examples. Examination of other cases is an important task in our future research. The next three examples are related to NE, IEH, and EP, respectively. The first example is trivial.

Example 1. The empty choice function, $C^0(S) = \varnothing$, satisfies EP, IEH, α, and β, but it violates NE.

We show that a myopic choice function satisfies the other axioms than IEH.

Example 2. A myopic choice function, $C^{\mathrm{M}}(S) = \arg\max_{X \in S} W_0(X)$, satisfies NE, EP, α, and β, but it violates IEH.

First, we show that C^{M} satisfies *EP*. Suppose that for any X^1, X^2, and any S, $[X^1 \succ^{\mathrm{P}} X^2$ and $X^1 \in S]$ hold. Then $W_0(X^1) > W_0(X^2)$. By the definition of C^{M}, if X^1 is feasible, $C^{\mathrm{M}}(S)$ does not contain X^2. Therefore, C^{M} satisfies EP.

Second, suppose that C^{M} satisfies the hypothesis of condition α. Then, by the definition of C^{M}, generation 0 has the maximal welfare on X^1 in S^2. Hence, clearly it does so in S^1 $(\subseteq S^2)$.

Third, suppose that C^{M} satisfies the hypothesis of β. Then by the definition of C^{M}, generation 0 has the same welfare on both X^1 and X^2 and therefore the conclusion of β holds.

Now, consider two consumption paths, X^1 and X^2, such that $W(X^1) = (2,0,0,0,0,...)$ and $W(X^2) = (1,1,0,0,0,...)$. IEH requires $X^2 \in C(\{X^1, X^2\})$, but $\{X^1\} = C(\{X^1, X^2\})$ by definition. Therefore, IEH does not hold.

A trivial choice function satisfies the other axioms than EP.

Example 3. A trivial choice function, $C^{\mathrm{T}}(S) = S$, satisfies NE, IEH, α, and β, but it violates EP.

C^{T} always contains all feasible consumption paths. So the conclusions of IEH, α, and β hold for any feasible set and any utility function, respectively. Therefore, C^{T} satisfies IEH, α, and β.

On the other hand, C^{T} violates EP.

6 Concluding Remarks

This chapter has provided an axiomatic characterization of the Rawlsian choice function in the Arrow–Dasgupta economy. Properties of the maximin consumption path have been examined by Arrow (1973) and Dasgupta (1974a, b), and it was shown that the maximin principle generates a saw-tooth shaped path. We make use of the axioms of non-emptiness, the Pareto principle, extended Hammond equity, conditions α and β to characterize the Rawlsian choice function. Pareto principle and Hammond equity are familiar to the characterization of the maximin principle in an intragenerational economy. Extended Hammond equity is an extension thereof in a dynamic economy with sympathetic preferences to future generations. Our versions of these conditions are exclusion of Pareto inferior paths and inclusion of extended Hammond superior paths. Conditions α and β are also familiar to the characterization of consistent choice functions.

Our characterization of the Rawlsian choice function is partial, in that Theorem 2 does not provide a complete axiomatization. We have shown that the Rawlsian choice function satisfies the above five axioms, and that the choice function satisfying these axioms generates the same choice set as that by the Rawlsian choice function when the opportunity set is the whole set of feasible paths. A full characterization of the Rawlsian choice function is a good research agenda in the field of intergenerational equity, which is left for future study.

The other remaining problems to be solved along this line of research are as follows. First, we must classify the family of choice functions that satisfies NE, IEH, α, β, and a weaker axiom of Pareto principle than EP. Since this family contains the Rawlsian choice function, we should explore whether there exists any other eligible one than the trivial choice function. Second, we must verify whether any other choice function than the myopic one that satisfies NE, EP, α, and β. Third, we should scrutinize the possibility whether other consistency axioms characterize the Rawlsian choice function.

Acknowledgments This work was presented at the 21COE-GLOPE conference in Waseda University on March 2005, and at the annual meeting of Japanese Economic Association in Chuo University on September 2005. We are grateful to the participants, especially Prof. Tomoichi Shinotsuka, Tsukuba University, and Prof. Naoki Yoshihara, Hitotsubashi University, for their valuable comments. We also acknowledge Prof. Kotaro Suzumura for organizing the research group where the idea of this work was born, and Prof. Holger Meinhardt, Karlsruhe University, for his detailed comments on the draft. Finally, but most deeply, we thank the referee and the editors of this book for their helpful and priceless comments. Their contributions are as large as those by the authors. Financial support through a Waseda University Grant for Special Research Projects and a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Science and Technology of Japan are gratefully acknowledged.

References

Arrow, K. J. (1973). Rawls's principle of just saving. *Swedish Journal of Economics*, *75*, 323–335

Asheim, G. B., Bossert, W., Sprumont, Y., & Suzumura, K. (2006). *Infinite-horizon choice functions* (mimeo). Institute of Economic Research, Hitotsubashi University

Asheim, G. B., Buchholz, W., & Tungodden, B. (2001). Justifying sustainability. *Journal of Environmental Economics and Management*, *41*, 252–268

Dasgupta, P. (1974a). On some alternative criteria for justice between generations. *Journal of Public Economics*, *3*, 405–423

Dasgupta, P. (1974b). On some problems arising from professor Rawls' conception of distributive justice. *Theory and Decision*, *4*, 325–344

Epstein, L. G. (1986a). Intergenerational consumption rules: An axiomatization of utilitarianism and egalitarianism. *Journal Economic Theory*, *38*, 280–297

Epstein, L. G. (1986b). Intergenerational preference orderings. *Social Choice and Welfare*, *3*, 151–160

Hammond, P. J. (1976). Equity, Arrow's conditions and Rawls' difference principle. *Econometrica*, *44*, 793–804

Hammond, P. J. (1979). Equity in two person situations: Some consequences. *Econometrica*, *47*, 1127–1135

Lauwers, L. (1997). Rawlsian equity and generalized utilitarianism with an infinite population. *Economic Theory*, *9*, 143–150

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press

Riley, J. G. (1976). Further remarks on the Rawlsian principle of just saving. *Scandinavian Journal of Economics*, *78*, 16–25

Sen, A. K. (1970). *Collective choice and social welfare*. Edinburgh: Oliver and Boyd
Sen, A. K. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica*, *45*, 1539–1572
Suga, K., & Udagawa, D. (2004). *Axiomatic characterization of the maximin principle in the Arrow-Dasgupta economy* (mimeo)
Suzumura, K. (1983). *Rational choice, collective decisions and social welfare*. New York: Cambridge University Press

Part III
Rational Choice, Individual Welfare, and Games

Suzumura Consistency

Walter Bossert

1 Introduction

Binary relations are at the heart of much of economic theory, both in the context of individual choice and in multi-agent decision problems. A fundamental coherence requirement imposed on a relation is the well-known transitivity axiom. If a relation is interpreted as a goodness relation, transitivity postulates that whenever one alternative is at least as good as a second and the second alternative is, in turn, at least as good as a third, then the first alternative is at least as good as the third. However, from an empirical as well as a conceptual perspective, transitivity is frequently considered too demanding and weaker notions of coherence have been proposed in the literature. Two alternatives that have received a considerable amount of attention are quasi-transitivity and acyclicity. Quasi-transitivity demands that the asymmetric factor of a relation (the betterness relation) is transitive, whereas acyclicity rules out the presence of betterness cycles. Quasi-transitivity is implied by transitivity and implies acyclicity. The reverse implications are not valid.

Suzumura (1976b) introduced an interesting alternative weakening of transitivity and showed that it can be considered a more intuitive property than quasi-transitivity. This notion of coherence, which Suzumura introduced under the name consistency, rules out the presence of cycles with at least one instance of betterness. Thus, the axiom is stronger than acyclicity and weaker than transitivity. It is equivalent to transitivity in the presence of reflexivity and completeness but independent of quasi-transitivity. Because the term consistency is used in various other contexts in economic theory (see, for instance, Thomson (1990)), I propose to refer to the axiom as *Suzumura consistency.*

Suzumura consistency is exactly what is needed to avoid the phenomenon of a money pump. If Suzumura consistency is violated by an agent's goodness relation,

W. Bossert
Department of Economics and CIREQ, University of Montréal, Canada
e-mail: walter.bossert@umontreal.ca

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,
Studies in Choice and Welfare.

there exists a cycle with at least one instance of betterness. In this case, the agent under consideration is willing to trade an alternative for another alternative (where 'willing to trade' is interpreted as being at least as well-off after the trade as before), the second alternative for a third and so on, until an alternative is reached such that getting back the original alternative is better than retaining possession of the last alternative in the chain. Thus, at the end of such a chain of exchanges, the agent is willing to give up the last alternative and, in addition, to pay a positive amount to get back the original alternative.

An important property of Suzumura consistency is that it is necessary and sufficient for the existence of an ordering extension of a relation. Szpilrajn (1930) showed that, for any asymmetric and transitive relation, there exists an asymmetric, transitive and complete relation that contains the original relation. An analogous result applies if asymmetry is replaced with reflexivity. Suzumura (1976b) has shown that the transitivity assumption can be weakened to Suzumura consistency without changing the conclusion regarding the existence of an ordering extension. Moreover, Suzumura consistency is the weakest possible property that guarantees this existence result. Because extension theorems are of considerable importance in many applications of set theory, this is a fundamental result and illustrates the significance of the property.

The purpose of this paper is to review the uses of Suzumura consistency in a variety of applications and to provide some new observations, with the objective of further underlining the importance of this axiom. The first step is a statement of Suzumura's (1976b) extension theorem in the following section, followed by an application in the theory of rational choice due to Bossert, Sprumont, and Suzumura (2005a) in Sect. 3. The last two sections provide new observations. In Sect. 4, a variant of the welfarism theorem that assumes Suzumura consistency instead of transitivity is provided, and Sect. 5 illustrates how an impossibility result in population ethics can be turned into a possibility by weakening transitivity to Suzumura consistency.

2 Relations and Extensions

Suppose X is a non-empty set of alternatives and $R \subseteq X \times X$ is a (binary) relation on X which is interpreted as a *goodness* relation, that is, $(x,y) \in R$ means that x is considered at least as good as y by the agent (or society) under consideration. The *diagonal relation* Δ on X is defined by

$$\Delta = \{(x,x) \mid x \in X\}.$$

The *asymmetric factor* of a relation R is defined by

$$P(R) = \{(x,y) \mid (x,y) \in R \text{ and } (y,x) \notin R\}$$

and the *symmetric factor* of R is

$$I(R) = \{(x,y) \mid (x,y) \in R \text{ and } (y,x) \in R\}.$$

Given the interpretation of R as a goodness relation, $P(R)$ is the *better-than* relation corresponding to R and $I(R)$ is the *equally-good* relation associated with R.

The *transitive closure* $\mathrm{tc}(R)$ of a relation R is defined by

$$\begin{aligned}\mathrm{tc}(R) = \{(x,y) \mid \; & \text{there exist } M \in \mathbb{N} \text{ and } x^0,\ldots,x^M \in X \text{ such that} \\ & x = x^0, (x^{m-1},x^m) \in R \text{ for all } m \in \{1,\ldots,M\} \text{ and } x^M = y\}.\end{aligned}$$

As is straightforward to verify,

$$R \subseteq Q \;\Rightarrow\; \mathrm{tc}(R) \subseteq \mathrm{tc}(Q) \tag{1}$$

for any two relations R and Q.

To illustrate the transitive closure, consider the following examples. First, let $X = \{x,y,z\}$ and $R = \{(x,x),(x,y),(y,y),(y,z),(z,x),(z,z)\}$. We obtain $\mathrm{tc}(R) = X \times X$. In addition to the pairs in R, the pair (x,z) must be in the transitive closure of R because we have $(x,y) \in R$ and $(y,z) \in R$. Analogously, (y,x) must be an element of $\mathrm{tc}(R)$ because $(y,z) \in R$ and $(z,x) \in R$, and (z,y) must be in $\mathrm{tc}(R)$ because $(z,x) \in R$ and $(x,y) \in R$. Now let $X = \{x,y,z\}$ and $R = \{(x,y),(y,z)\}$. As it is straightforward to verify, we have $\mathrm{tc}(R) = \{(x,y),(y,z),(x,z)\}$.

A relation R is *reflexive* if, for all $x \in X$,

$$(x,x) \in R$$

and R is *asymmetric* if

$$R = P(R).$$

Furthermore, R is *complete* if, for all $x,y \in X$,

$$x \neq y \;\Rightarrow\; (x,y) \in R \text{ or } (y,x) \in R$$

and R is *transitive* if, for all $x,y,z \in X$,

$$(x,y) \in R \text{ and } (y,z) \in R \;\Rightarrow\; (x,z) \in R.$$

R is *Suzumura consistent* if, for all $x,y \in X$,

$$(x,y) \in \mathrm{tc}(R) \Rightarrow (y,x) \notin P(R).$$

A *quasi-ordering* is a reflexive and transitive relation and an *ordering* is a complete quasi-ordering.

The notion of Suzumura consistency is due to Suzumura (1976b) and it is equivalent to the requirement that any cycle must be such that all relations involved in this cycle are instances of equal goodness – betterness cannot occur. Clearly, this requirement implies (but is not implied by) the well-known *acyclicity* axiom which

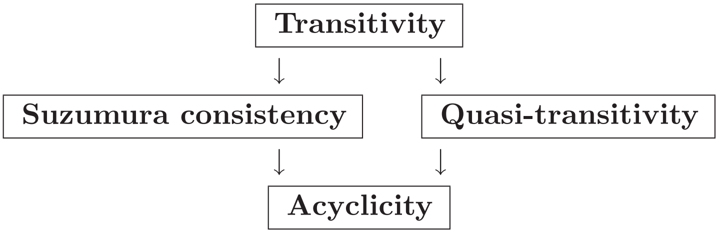


Fig. 1 Logical relationships

rules out the existence of betterness cycles (cycles where *all* relations involve the asymmetric factor of the relation). Suzumura consistency and *quasi-transitivity,* which requires that $P(R)$ is transitive, are independent. Transitivity implies Suzumura consistency but the reverse implication is not true in general. However, if R is reflexive and complete, Suzumura consistency and transitivity are equivalent. Figure 1 illustrates the relationships among transitivity and the above-mentioned weakenings of this property. Each arrow represents a direct implication, and these implications together with those resulting from chains of arrows are the only ones that are valid in the absence of further properties imposed on R.

A relation R' is an *extension* of a relation R if

$$R \subseteq R' \text{ and } P(R) \subseteq P(R').$$

If an extension R' of R is an ordering, we refer to R' as an *ordering extension* of R. One of the most fundamental results on extensions of binary relations is due to Szpilrajn (1930) who showed that any transitive and asymmetric relation has a transitive, asymmetric and complete extension. The result remains true if asymmetry is replaced with reflexivity, that is, any quasi-ordering has an ordering extension. Arrow (1963, p. 64) stated this generalization of Szpilrajn's theorem without a proof and Hansson (1968) provided a proof on the basis of Szpilrajn's original theorem.

While the property of being a quasi-ordering is sufficient for the existence of an ordering extension of a relation, this is not necessary. As shown by Suzumura (1976b), Suzumura consistency is *necessary and sufficient* for the existence of an ordering extension. This observation is stated formally in the following theorem, see Suzumura (1976b, pp. 389–390) for a proof.

Theorem 1. *A relation R has an ordering extension if and only if R is Suzumura consistent.*

Theorem 1 is an important result. It establishes that Suzumura consistency is the weakest possible property of a relation that still guarantees the existence of an ordering extension. Note that quasi-transitivity (which, as mentioned earlier, is logically independent of Suzumura consistency) has nothing to do with the possibility of extending a binary relation to an ordering.

3 Rational Choice

Suzumura consistency has recently been examined in the context of *rational choice*. Observed (or observable) choices are *rationalizable* if there exists a relation such that, for any feasible set, the set of chosen alternatives coincides with the set of greatest or maximal elements according to this relation.

Following the contributions of Hansson (1968), Richter (1966, 1971), Suzumura (1976a, 1977, Chap. 2 in 1983) and others, the approach to rational choice analyzed in this paper is capable of accommodating a wide variety of choice situations because no restrictions (other than non-emptiness) are imposed on the domain of a choice function. Letting $\mathcal{X}$ denote the power set of X excluding the empty set, a *choice function* is a mapping $C\colon \Sigma \to \mathcal{X}$ such that $C(S) \subseteq S$ for all $S \in \Sigma$, where $\Sigma \subseteq \mathcal{X}$ with $\Sigma \neq \emptyset$ is the domain of C.

The *direct revealed preference relation* $R_C \subseteq X \times X$ of a choice function C with an arbitrary domain Σ is defined as

$$R_C = \{(x,y) \mid \text{there exists } S \in \Sigma \text{ such that } x \in C(S) \text{ and } y \in S\}.$$

The *(indirect) revealed preference relation* of C is the transitive closure $\text{tc}(R_C)$ of the direct revealed preference relation R_C.

A choice function C is *greatest-element rationalizable* if there exists a relation R on X such that

$$C(S) = \{x \in S \mid (x,y) \in R \text{ for all } y \in S\}$$

for all $S \in \Sigma$. If such a relation R exists, it is called a *rationalization* of C. The most common alternative to greatest-element rationalizability is *maximal-element* rationalizability which requires the existence of a relation R such that, for all feasible sets S, $C(S)$ is equal to the set of maximal elements in S according to R, that is, no element in S is better than any element in $C(S)$. Bossert, Sprumont, and Suzumura (2005b) provide a detailed analysis of maximal-element rationalizability. Logical relationships between, and characterizations of, various notions of rationalizability, both on arbitrary domains and under more specific domain assumptions, can be found in Bossert, Sprumont, and Suzumura (2006).

To interpret a rationalization as a goodness relation, it is usually required that it satisfy additional properties such as the *richness* axioms reflexivity and completeness, or one of the *coherence* properties acyclicity, quasi-transitivity, Suzumura consistency and transitivity. The full set of rationalizability notions that can be obtained by combining one or both (or none) of the richness properties with one (or none) of the coherence properties is analyzed in Bossert and Suzumura (2008). They show that, if all these combinations are available, it is sufficient to restrict attention to greatest-element rationalizability: for each notion of maximal-element rationalizability, there exists a notion of the greatest-element rationalizability (possibly involving different richness and coherence properties) that is equivalent. Thus, restricting attention to greatest-element rationalizability does not involve any loss of generality.

Bossert, Sprumont, and Suzumura (2005a) have characterized all notions of rationalizability when the coherence property required is Suzumura consistency. As mentioned earlier, Suzumura consistency and transitivity are equivalent in the presence of reflexivity and completeness. Thus, greatest-element rationalizability by a reflexive, complete and Suzumura-consistent relation is equivalent to greatest-element rationalizability by an ordering and Richter's (1966, 1971) results apply; see Theorem 2. Moreover, greatest-element rationalizability by a complete and Suzumura-consistent relation implies greatest-element rationalizability by a reflexive, complete and Suzumura-consistent relation, and greatest-element rationalizability by a Suzumura-consistent relation implies greatest-element rationalizability by a reflexive and Suzumura-consistent relation. Analogous observations apply in the case of maximal-element rationalizability; see Bossert, Sprumont, and Suzumura (2005a, Theorem 1). As pointed out in Bossert, Sprumont, and Suzumura (2006), as soon as the coherence properties quasi-transitivity or acyclicity are imposed, reflexivity no longer is guaranteed as an additional property of a rationalization. Thus, Suzumura consistency stands out from these alternative weakenings of transitivity in this regard: as is the case for transitive greatest-element (or maximal-element) rationalizability, any notion of Suzumura-consistent greatest-element (or maximal-element) rationalizability is equivalent to the definition that is obtained if reflexivity is added as a property of a rationalization.

Richter (1971) showed that the following axiom is necessary and sufficient for greatest-element rationalizability by a transitive relation and by an ordering. Thus, the existence of a rationalizing relation that is not merely a quasi-ordering but an ordering follows from greatest-element rationalizability by a transitive relation. This observation sets transitive greatest-element rationalizability apart from other notions of greatest-element rationalizability involving weaker coherence requirements.

Transitive-closure coherence. For all $S \in \Sigma$ and for all $x \in S$,

$$(x,y) \in \mathrm{tc}(R_C) \text{ for all } y \in S \;\Rightarrow\; x \in C(S).$$

We now obtain the following result; see Bossert, Sprumont, and Suzumura (2005a).

Theorem 2. *C is greatest-element rationalizable by a (reflexive.) complete and Suzumura-consistent relation if and only if C satisfies transitive-closure coherence.*

Proof. To prove the 'only-if' part, suppose C is greatest-element rationalizable by a complete and Suzumura-consistent relation R. We prove that C is greatest-element rationalizable by a reflexive, complete and Suzumura-consistent relation. Together with the observation that Suzumura consistency and transitivity are equivalent in the presence of reflexivity and completeness and Richter's (1971) result, this establishes that transitive-closure coherence is satisfied.

Let

$$\begin{aligned} R' = {} & [R \cup \Delta \cup \{(y,x) \mid x \notin C(\Sigma) \text{ and } y \in C(\Sigma)\}] \\ & \setminus \{(x,y) \mid x \notin C(\Sigma) \text{ and } y \in C(\Sigma)\}. \end{aligned}$$

Clearly, R' is reflexive by definition.

To show that R' is complete, let $x, y \in X$ be such that $x \neq y$ and $(x,y) \notin R'$. By definition of R', this implies

$$(x,y) \notin R \text{ and } [x \notin C(\Sigma) \text{ or } y \in C(\Sigma)]$$

or

$$x \notin C(\Sigma) \text{ and } y \in C(\Sigma).$$

If the former applies, the completeness of R implies $(y,x) \in R$ and, by definition of R', we obtain $(y,x) \in R'$. If the latter is true, $(y,x) \in R'$ follows immediately from the definition of R'.

Next, we show that R' is Suzumura consistent. Let $(x,y) \in \operatorname{tc}(R')$. By definition, there exist $M \in \mathbb{N}$ and $x^0, \ldots, x^M \in X$ be such that $x = x^0$, $(x^{m-1}, x^m) \in R'$ for all $m \in \{1, \ldots, M\}$ and $x^M = y$. Clearly, we can, without loss of generality, assume that $x^{m-1} \neq x^m$ for all $m \in \{1, \ldots, M\}$. We distinguish two cases.

(i) $x^0 \notin C(\Sigma)$. In this case, it follows that $x^1 \notin C(\Sigma)$; otherwise we would have $(x^1, x^0) \in P(R')$ by definition of R', contradicting our hypothesis. Successively applying this argument to all $m \in \{1, \ldots, M\}$, we obtain $x^m \notin C(\Sigma)$ for all $m \in \{1, \ldots, M\}$. By definition of R', this implies $(x^{m-1}, x^m) \in R$ for all $m \in \{1, \ldots, M\}$. By the Suzumura consistency of R, we must have $(x^M, x^0) \notin P(R)$. Because $x^M \notin C(\Sigma)$, this implies, according to the definition of R', $(x^M, x^0) \notin P(R')$.

(ii) $x^0 \in C(\Sigma)$. If $x^M \notin C(\Sigma)$, $(x^M, x^0) \notin P(R')$ follows immediately from the definition of R'. If $x^M \in C(\Sigma)$, it follows that $x^{M-1} \in C(\Sigma)$; otherwise we would have $(x^{M-1}, x^M) \notin R'$ by definition of R', contradicting our hypothesis. Successively applying this argument to all $m \in \{1, \ldots, M\}$, we obtain $x^m \in C(\Sigma)$ for all $m \in \{1, \ldots, M\}$. By definition of R', this implies $(x^{m-1}, x^m) \in R$ for all $m \in \{1, \ldots, M\}$. By the Suzumura consistency of R, we must have $(x^M, x^0) \notin P(R)$. Because $x^0 \in C(\Sigma)$, this implies, according to the definition of R', $(x^M, x^0) \notin P(R')$.

Finally, we show that R' is a rationalization of C. Let $S \in \Sigma$ and $x \in S$.

Suppose first that $(x,y) \in R'$ for all $y \in S$. If $|S| = 1$, $x \in C(S)$ follows immediately because $C(S)$ is non-empty. If $|S| \geq 2$, we obtain $x \in C(\Sigma)$ by definition of R'. Because R is a rationalization of C, this implies $(x,x) \in R$. By definition of R', $(x,z) \in R$ for all $z \in C(S)$. Therefore, $(x,z) \in R$ for all $z \in C(S) \cup \{x\}$. Suppose, by way of contradiction, that $x \notin C(S)$. Because R is a rationalization of C, it follows that there exists $y \in S \setminus (C(S) \cup \{x\})$ such that $(x,y) \notin R$. The completeness of R implies $(y,x) \in P(R)$. Let $z \in C(S)$. It follows that $(z,y) \in R$ because R is a rationalization of C and, as established earlier, $(x,z) \in R$. This contradicts the Suzumura consistency of R.

To prove the converse implication, suppose $x \in C(S)$. Because R is a rationalization of C, we have $(x,y) \in R$ for all $y \in S$. In particular, this implies $(x,x) \in R$ and, according to the definition of R', we obtain $(x,y) \in R'$ for all $y \in S$.

The 'if' part of the theorem follows immediately from the equivalence of transitive-closure coherence and greatest-element rationalizability by a reflexive, complete and transitive rationalization established by Richter (1971) and the observation that Suzumura consistency and transitivity coincide in the presence of reflexivity and completeness. □

If completeness is dropped as a requirement imposed on a rationalization, a weaker notion of greatest-element rationalizability is obtained. In contrast to greatest-element rationalizability by a quasi-transitive or an acyclical relation which leads to much more complex necessary and sufficient conditions (see Bossert and Suzumura (2008)), requiring a rationalization to be Suzumura consistent preserves the intuitive and transparent nature of the characterization stated in Theorem 2. There is a unique minimal Suzumura-consistent relation that has to be respected by any Suzumura-consistent rationalization, namely, the *Suzumura-consistent closure* of R_C. The Suzumura-consistent closure $\mathrm{sc}(R)$ of a relation R is defined by

$$\mathrm{sc}(R) = R \cup \{(x,y) \mid (x,y) \in \mathrm{tc}(R) \text{ and } (y,x) \in R\}.$$

Clearly, $R \subseteq \mathrm{sc}(R) \subseteq \mathrm{tc}(R)$. Just as $\mathrm{tc}(R)$ is the unique smallest transitive relation containing R, $\mathrm{sc}(R)$ is the unique smallest Suzumura-consistent relation containing R; see Bossert, Sprumont, and Suzumura (2005a).

To see that this is the case, we first establish that $\mathrm{sc}(R)$ is Suzumura consistent. Suppose $M \in \mathbb{N}$ and $x^0, \ldots, x^M \in X$ are such that $(x^{m-1}, x^m) \in \mathrm{sc}(R)$ for all $m \in \{1, \ldots, M\}$. We show that $(x^M, x^0) \notin P(\mathrm{sc}(R))$. Because $\mathrm{sc}(R) \subseteq \mathrm{tc}(R)$, $(x^{m-1}, x^m) \in \mathrm{tc}(R)$ for all $m \in \{1, \ldots, M\}$, and the transitivity of $\mathrm{tc}(R)$ implies

$$(x^0, x^M) \in \mathrm{tc}(R). \tag{2}$$

If $(x^M, x^0) \notin \mathrm{sc}(R)$, we immediately obtain $(x^M, x^0) \notin P(\mathrm{sc}(R))$ and we are done. Now suppose that $(x^M, x^0) \in \mathrm{sc}(R)$. By definition of $\mathrm{sc}(R)$, we must have

$$(x^M, x^0) \in R \text{ or } [(x^M, x^0) \in \mathrm{tc}(R) \text{ and } (x^0, x^M) \in R].$$

If $(x^M, x^0) \in R$, (2) and the definition of $\mathrm{sc}(R)$ together imply $(x^0, x^M) \in \mathrm{sc}(R)$ and, thus, $(x^M, x^0) \notin P(\mathrm{sc}(R))$. If $(x^M, x^0) \in \mathrm{tc}(R)$ and $(x^0, x^M) \in R$, $(x^0, x^M) \in \mathrm{sc}(R)$ follows because $R \subseteq \mathrm{sc}(R)$. Again, this implies $(x^M, x^0) \notin P(\mathrm{sc}(R))$ and the proof that $\mathrm{sc}(R)$ is Suzumura consistent is complete.

To show that $\mathrm{sc}(R)$ is the smallest Suzumura-consistent relation containing R, suppose that Q is an arbitrary Suzumura-consistent relation containing R. To complete the proof, we establish that $\mathrm{sc}(R) \subseteq Q$. Suppose that $(x,y) \in \mathrm{sc}(R)$. By definition of $\mathrm{sc}(R)$,

$$(x,y) \in R \text{ or } [(x,y) \in \mathrm{tc}(R) \text{ and } (y,x) \in R].$$

If $(x,y) \in R$, $(x,y) \in Q$ follows because R is contained in Q by assumption. If $(x,y) \in \mathrm{tc}(R)$ and $(y,x) \in R$, (1) and the assumption $R \subseteq Q$ together imply that $(x,y) \in \mathrm{tc}(Q)$ and $(y,x) \in Q$. If $(x,y) \notin Q$, we obtain $(y,x) \in P(Q)$ in view of $(y,x) \in Q$. Since $(x,y) \in \mathrm{tc}(Q)$, this contradicts the Suzumura consistency of Q. Therefore, we must have $(x,y) \in Q$.

The property of $\mathrm{sc}(R)$ just established is crucial in obtaining a clear-cut and intuitive rationalizability result even without imposing completeness (in which case Suzumura-consistent greatest-element rationalizability is not equivalent to transitive greatest-element rationalizability). In contrast, there is no such thing as a quasi-transitive closure or an acyclical closure of a relation, which explains why rationalizability results involving these coherence properties are much more complex.

The following examples illustrate the Suzumura-consistent closure and its relation to the transitive closure. First, let $X = \{x,y,z\}$ and $R = \{(x,x),(x,y),(y,y),(y,z)$ $(z,x),(z,z)\}$. We obtain $\mathrm{sc}(R) = \mathrm{tc}(R) = X \times X$. Now let $X = \{x,y,z\}$ and $R = \{(x,y),(y,z)\}$. We have $\mathrm{sc}(R) = R$ and $\mathrm{tc}(R) = \{(x,y),(y,z),(x,z)\}$. In the first example, the Suzumura-consistent closure coincides with the transitive closure, whereas in the second, the Suzumura-consistent closure is a strict subset of the transitive closure.

Greatest-element rationalizability by means of a Suzumura-consistent (and reflexive but not necessarily complete) relation can now be characterized by employing a natural weakening of transitive-closure coherence: all that needs to be done is replacing the transitive closure of the direct revealed preference relation by its Suzumura-consistent closure.

Suzumura-consistent-closure coherence. For all $S \in \Sigma$ and for all $x \in S$,

$$(x,y) \in \mathrm{sc}(R_C) \text{ for all } y \in S \;\Rightarrow\; x \in C(S).$$

The following characterization is also due to Bossert, Sprumont, and Suzumura (2005a).

Theorem 3. *C is greatest-element rationalizable by a (reflexive and) Suzumura-consistent relation if and only if C satisfies Suzumura-consistent-closure coherence.*

Proof. To prove the 'only-if' part of the theorem, suppose R is a Suzumura-consistent rationalization of C and let $S \in \Sigma$ and $x \in S$ be such that $(x,y) \in \mathrm{sc}(R_C)$ for all $y \in S$. Consider any $y \in S$. By definition,

$$(x,y) \in R_C \text{ or } [(x,y) \in \mathrm{tc}(R_C) \text{ and } (y,x) \in R_C].$$

If $(x,y) \in R_C$, there exists $T \in \Sigma$ such that $x \in C(T)$ and $y \in T$. Because R greatest-element rationalizes C, this implies $(x,y) \in R$. If $(x,y) \in \mathrm{tc}(R_C)$ and $(y,x) \in R_C$, there exist $M \in \mathbb{N}$ and $x^0,\ldots,x^M \in X$ such that $x = x^0$, $(x^{m-1},x^m) \in R_C$ for all $m \in \{1,\ldots,M\}$ and $x^M = y$. As in the argument just used, the assumption that R greatest-element rationalizes C implies $(x^{m-1},x^m) \in R$ for all $m \in \{1,\ldots,M\}$ and, thus, $(x,y) \in \mathrm{tc}(R)$. Furthermore, $(y,x) \in R_C$ implies $(y,x) \in R$ because R is a rationalization of R. If $(x,y) \notin R$, it follows that $(y,x) \in P(R)$ in view of $(y,x) \in R$. Because $(x,y) \in \mathrm{tc}(R)$, this contradicts the Suzumura consistency of R. Therefore, $(x,y) \in R$. Because $y \in S$ has been chosen arbitrarily, this is true for all $y \in S$ and, as a consequence of the assumption that R greatest-element rationalizes C, we obtain $x \in C(S)$.

To prove the 'if' part, suppose C satisfies Suzumura-consistent-closure coherence. We first show that $\mathrm{sc}(R_C)$ is a Suzumura-consistent rationalization of C. That $\mathrm{sc}(R_C)$ is Suzumura consistent has already been established. To prove that $\mathrm{sc}(R_C)$ is a rationalization of C, suppose first that $S \in \Sigma$ and $x \in S$. Suppose $(x,y) \in \mathrm{sc}(R_C)$

for all $y \in S$. Suzumura-consistent-closure coherence implies $x \in C(S)$. Conversely, suppose $x \in C(S)$. By definition, this implies $(x,y) \in R_C$ for all $y \in S$ and, because $R_C \subseteq \mathrm{sc}(R_C)$, we obtain $(x,y) \in \mathrm{sc}(R_C)$ for all $y \in S$. The proof is completed by showing that

$$R' = (\mathrm{sc}(R_C) \cup \Delta) \setminus \{(x,y) \mid x \notin C(\Sigma) \text{ and } x \neq y\}$$

is a reflexive and Suzumura-consistent rationalization of C.

That R' is reflexive is obvious. To prove that R' is Suzumura consistent, suppose $(x,y) \in \mathrm{tc}(R')$. Thus, there exist $M \in \mathbb{N}$ and $x^0, \ldots, x^M \in X$ such that $x = x^0$, $(x^{m-1}, x^m) \in R'$ for all $m \in \{1, \ldots, M\}$ and $x^M = y$. Clearly, we can without loss of generality assume that $x^{m-1} \neq x^m$ for all $m \in \{1, \ldots, M\}$. By definition of R', $x^0 \in C(\Sigma)$. If $x^M \notin C(\Sigma)$, $(x^M, x^0) \notin P(R')$ follows immediately from the definition of R'. If $x^M \in C(\Sigma)$, it follows that $x^{M-1} \in C(\Sigma)$; otherwise we would have $(x^{M-1}, x^M) \notin R'$ by definition of R', contradicting our hypothesis. Successively applying this argument to all $m \in \{0, \ldots, M-1\}$, we obtain $x^m \in C(\Sigma)$ for all $m \in \{0, \ldots, M-1\}$. By definition of R', this implies $(x^{m-1}, x^m) \in \mathrm{sc}(R_C)$ for all $m \in \{1, \ldots, M\}$. By the Suzumura consistency of $\mathrm{sc}(R_C)$, we must have $(x^M, x^0) \notin P(\mathrm{sc}(R_C))$. Because $x^0 \in C(\Sigma)$, this implies, according to the definition of R', $(x^M, x^0) \notin P(R')$.

It remains to be shown that R' is a rationalization of C. Let $S \in \Sigma$ and $x \in S$.

First, suppose $(x,y) \in R'$ for all $y \in S$. By definition of R',

$$(x,y) \in \mathrm{sc}(R_C) \tag{3}$$

for all $y \in S \setminus \{x\}$ and $x \in C(\Sigma)$. Because $\mathrm{sc}(R_C)$ is a rationalization of C, this implies $(x,x) \in \mathrm{sc}(R_C)$. Suppose, by way of contradiction, that $x \notin C(S)$. Because $\mathrm{sc}(R)$ is a rationalization of C, it follows that there exists $y \in S \setminus \{x\}$ such that $(x,y) \notin \mathrm{sc}(R_C)$, contradicting (3).

Finally, suppose $x \in C(S)$. This implies $(x,y) \in \mathrm{sc}(R_C)$ for all $y \in S$ because $\mathrm{sc}(R_C)$ is a rationalization of C. Furthermore, because $C(S) \subseteq C(\Sigma)$, we have $x \in C(\Sigma)$. By definition of R', this implies $(x,y) \in R'$ for all $y \in S$. □

4 Welfarism

Following Arrow's (1951, 2nd ed. 1963) impossibility theorem, one route of escape from its negative consequences that has been chosen in the subsequent literature is to assume that a social ranking is established on the basis of a richer informational framework. In Arrow's setup, the individual goodness relations form the informational basis of collective choice. This approach rules out, in particular, interpersonal comparisons of well-being. An informationally richer environment is obtained if a social ranking is allowed to depend on *utility* profiles instead of profiles of goodness relations, and these utilities can be assumed to carry more than just ordinally measurable and interpersonally non-comparable information regarding the well-being of the agents. Under an implicit regularity assumption that guarantees the existence of representations of the individual goodness relations, the Arrow framework is included as a special case: it corresponds to the informational assumption of ordinal measurability and interpersonal non-comparability.

The universal set of alternatives X is assumed to contain at least three elements. There are a finite number $n \geq 2$ of agents indexed by the first n positive integers, so that the set of agents is $N = \{1,\ldots,n\}$. The set of all utility functions $U\colon X \to \mathbb{R}$ is denoted by $\mathcal{U}$ and its n-fold Cartesian product is $\mathcal{U}^n$. A *utility profile* is an n-tuple $\mathbf{U} = (U_1,\ldots,U_n) \in \mathcal{U}^n$.

A *collective choice functional* is a mapping $F\colon \mathcal{D} \to \mathcal{B}$ where $\mathcal{D} \subseteq \mathcal{U}^n$ is the domain of this functional, assumed to be non-empty, and $\mathcal{B}$ is the set of all binary relations on X. For each utility profile $\mathbf{U} \in \mathcal{D}$, $F(\mathbf{U})$ is the social preference corresponding to $\mathbf{U}$. A *reflexive and Suzumura-consistent collective choice functional* is a collective choice functional F such that $F(\mathbf{U})$ is reflexive and Suzumura consistent for all $\mathbf{U} \in \mathcal{D}$, and a *social-evaluation functional* is a collective choice functional F such that $F(\mathbf{U})$ is an ordering for all $\mathbf{U} \in \mathcal{D}$. Informational assumptions regarding the measurability and interpersonal comparability of individual utilities can be expressed by requiring the collective choice functional to be constant on sets of utility profiles that contain the same information. For example, if utilities are *cardinally measurable and fully comparable,* any utility profile $\mathbf{U}'$ that is obtained from a profile $\mathbf{U}$ by applying the same increasing affine transformation to all individual utility functions carries the same information as $\mathbf{U}$ itself. Thus, the collective choice functional must assign the same social ranking to both profiles. See Blackorby, Donaldson, and Weymark (1984) or Bossert and Weymark (2004) for discussions of information assumptions in social-choice theory.

A fundamental result in this setting is the *welfarism theorem;* see, for instance, d'Aspremont and Gevers (1977) and Hammond (1979). A social-evaluation functional F is *welfarist* if, for any utility profile $\mathbf{U}$ and for any two alternatives x and y, the social ranking of x and y according to the social ordering assigned to the profile $\mathbf{U}$ by F depends on the two utility vectors $\mathbf{U}(x) = (U_1(x),\ldots,U_n(x))$ and $\mathbf{U}(y) = (U_1(y),\ldots,U_n(y))$ only. Thus, a *single* ordering of utility vectors is sufficient to rank the alternatives for *any* profile. The welfarism theorem states that, provided that the domain of the social-evaluation functional consists of all possible utility profiles, welfarism is equivalent to the conjunction of Pareto indifference and independence of irrelevant alternatives.

In this section, it is illustrated that the welfarism theorem has an analogous formulation for reflexive and Suzumura-consistent collective choice functionals: even if every social ranking is merely required to be reflexive and Suzumura consistent rather than an ordering, the conjunction of the two axioms is (under the unlimited-domain assumption) equivalent to the existence of a single reflexive and Suzumura-consistent relation R defined on utility vectors that is sufficient to obtain the social ranking for any utility profile. This relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is referred to as a *social-evaluation relation.* The requisite axioms are the following.

Unlimited domain. $\mathcal{D} = \mathcal{U}^n$.

Pareto indifference. For all $x, y \in X$ and for all $\mathbf{U} \in \mathcal{D}$,

$$U_i(x) = U_i(y) \text{ for all } i \in N \Rightarrow (x,y) \in I(F(\mathbf{U})).$$

Independence of irrelevant alternatives. For all $x,y \in X$ and for all $\mathbf{U},\mathbf{U}' \in \mathcal{D}$ such that $U_i(x) = U_i'(x)$ and $U_i(y) = U_i'(y)$ for all $i \in N$,

$$\big[(x,y) \in F(\mathbf{U}) \Leftrightarrow (x,y) \in F(\mathbf{U}')\big] \text{ and } \big[(y,x) \in F(\mathbf{U}) \Leftrightarrow (y,x) \in F(\mathbf{U}')\big].$$

The following theorem generalizes the standard welfarism theorem by allowing social relations to be intransitive and incomplete but imposing the Suzumura-consistency requirement.

Theorem 4. *Suppose that a reflexive and Suzumura-consistent collective choice functional F satisfies unlimited domain. F satisfies Pareto indifference and independence of irrelevant alternatives if and only if there exists a reflexive and Suzumura-consistent social-evaluation relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ such that, for all $x,y \in X$ and for all $\mathbf{U} \in \mathcal{U}^n$,*

$$(x,y) \in F(\mathbf{U}) \Leftrightarrow (\mathbf{U}(x),\mathbf{U}(y)) \in R. \tag{4}$$

Proof. The 'if' part of the theorem is straightforward to verify. To prove the converse implication, suppose that F is a reflexive and Suzumura-consistent collective choice functional satisfying unlimited domain, Pareto indifference and independence of irrelevant alternatives. Define the relation $R \subseteq \mathbb{R}^n \times \mathbb{R}^n$ as follows. For all $u,v \in \mathbb{R}^n$, $(u,v) \in R$ if and only if there exist $x,y \in X$ and $\mathbf{U} \in \mathcal{U}^n$ such that $\mathbf{U}(x) = u$, $\mathbf{U}(y) = v$ and $(x,y) \in F(\mathbf{U})$. That R is well-defined follows as in the standard welfarism theorem; see, for instance, Blackorby, Donaldson, and Weymark (1984) or Bossert and Weymark (2004). Once R is well defined, (4) is immediate and, furthermore, R is reflexive because $F(\mathbf{U})$ is reflexive for all $\mathbf{U} \in \mathcal{U}^n$. The proof is completed by showing that R is Suzumura consistent. Let $u,v \in \mathbb{R}^n$ be such that $(u,v) \in \text{tc}(R)$. By definition of the transitive closure of a relation, there exist $M \in \mathbb{N}$ and $u^0,\ldots,u^M \in \mathbb{R}^n$ such that $u = u^0$, $(u^{m-1},u^m) \in R$ for all $m \in \{1,\ldots,M\}$ and $u^M = v$. By definition of R, there exist $x^0,\ldots,x^M \in X$ and $\mathbf{U}^1,\ldots,\mathbf{U}^M \in \mathcal{U}^n$ such that $\mathbf{U}^{m-1}(x^{m-1}) = u^{m-1}$, $\mathbf{U}^{m-1}(x^m) = u^m$ and $(x^{m-1},x^m) \in F(\mathbf{U}^{m-1})$ for all $m \in \{1,\ldots,M\}$. By unlimited domain, there exists $\mathbf{V} \in \mathcal{U}^n$ such that $\mathbf{V}(x^m) = u^m$ for all $m \in \{0,\ldots,M\}$. Using (4), it follows that $(x^{m-1},x^m) \in F(\mathbf{V})$ for all $m \in \{1,\ldots,M\}$. Because $F(\mathbf{V})$ is Suzumura consistent, it follows that $(x^M,x^0) \notin P(F(\mathbf{V}))$. Thus, by (4), $(v,u) = \big(\mathbf{V}(x^M),\mathbf{V}(x^0)\big) \notin P(R)$ and R is Suzumura consistent. $\square$

5 Population Ethics

The traditional social-choice framework with a fixed population is unable to capture important aspects of many public-policy choices. For example, decisions on funds devoted to prenatal care, the intergenerational allocation of resources and the design of aid packages to developing countries involve endogenous populations: depending on the selected alternative, some individuals may or may not be brought into existence. To address this issue, a social ranking must be capable of comparing alternatives with different population sizes.

The possibility of extending the welfarist approach to a variable-population environment has been examined in a variety of contributions, most notably in applied ethics; see, for instance, Parfit (1976, 1982, 1984). Impossibility results arise frequently in this area, and it is therefore of interest to examine the possibilities of escaping these negative conclusions. The purpose of this section is to illustrate that weakening transitivity to Suzumura consistency turns some of these impossibilities into possibilities. Of course, to ensure that Suzumura consistency is indeed weaker than transitivity, we cannot impose reflexivity, completeness and Suzumura consistency – as mentioned earlier, Suzumura consistency and transitivity coincide in the presence of the two richness conditions. Therefore, the question arises whether reflexivity and completeness rather than transitivity are, to a large extent, responsible for the impossibilities. This is not the case: although most of the impossibility results in this area have been established for orderings, they remain true if reflexivity and completeness are dropped.

A variable-population version of a social-evaluation relation is defined on the set of utility vectors of *any* dimension, that is, it is a relation $R \subseteq \Omega \times \Omega$, where $\Omega = \cup_{n \in \mathbb{N}} \mathbb{R}^n$. The components of a utility vector $u \in \Omega$ are interpreted as the *lifetime* utilities of those alive in the requisite alternative. For an individual who is alive, a *neutral* life is one which is as good as one without experiences. A life above neutrality is worth living, a life below neutrality is not. Following standard practice in population ethics, a lifetime-utility level of zero is assigned to neutrality.

In Blackorby, Bossert, and Donaldson (2006), it is shown that there exists no variable-population social-evaluation ordering satisfying four axioms that are common in the literature. This result can be generalized by noting that it does not make use of reflexivity or completeness – all that is needed is the transitivity of R.

The first axiom is *minimal increasingness.* It requires that, for any fixed population size, if all individuals have the same utility in two utility vectors, then the vector where everyone's utility is higher is better according to R. We use $\mathbf{1}_n$ to denote the vector of $n \in \mathbb{N}$ ones.

Minimal increasingness. For all $n \in \mathbb{N}$ and for all $\beta, \gamma \in \mathbb{R}$,

$$\beta > \gamma \;\Rightarrow\; (\beta \mathbf{1}_n, \gamma \mathbf{1}_n) \in P(R).$$

Minimal increasingness is a weak unanimity property: it only applies if everyone has the same utility in both alternatives to be compared.

Another fixed-population axiom is *weak inequality aversion.* This axiom demands that, for any fixed population size, perfect equality is at least as good as any distribution of the same total utility.

Weak inequality aversion. For all $n \in \mathbb{N}$ and for all $u \in \mathbb{R}^n$,

$$\left(\left(\frac{1}{n} \sum_{i=1}^{n} u_i \right) \mathbf{1}_n, u \right) \in R.$$

Sikora (1978) suggests a variable-population version of the Pareto principle. The axiom usually is defined as the conjunction of the strong Pareto principle and the requirement that the addition of an individual above neutrality to a utility-unaffected population is a social improvement. Because strong Pareto will be introduced as a separate axiom later on and is not needed for the impossibility result, we use the second part of the property only.

Pareto plus. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $a \in \mathbb{R}_{++}$,

$$((u,a),u) \in P(R).$$

In the axiom statement, the population common to u and (u,a) is unaffected and, thus, to defend the axiom on individual-goodness grounds, it must be argued that a level of well-being above neutrality is better than non-existence. Thus, the axiom applies the Pareto condition to situations where a person is not alive in all alternatives to be compared. While it is possible to compare alternatives with different populations from a social point of view (which is the issue addressed in population ethics), it is not clear that such a comparison can be made from the viewpoint of an individual if the person is not alive in one of the alternatives. It is therefore difficult to interpret this axiom as a Pareto condition because it appears to be based on the idea that people who do not exist have interests that should be respected. There is, therefore, an important asymmetry in the assessment of alternatives with different populations. It is perfectly reasonable to say that an individual considers life worth living if the person is alive with a positive level of lifetime well-being, but that does not justify the claim that a person who does not exist gains from being brought into existence with a lifetime utility above neutrality.

As is the case for Pareto plus, the final axiom used in our impossibility result applies to comparisons across population sizes. A variable-population social-evaluation relation leads to the *repugnant conclusion* if population size can always be substituted for well-being, no matter how close to neutrality the utilities of a large population are. That is, mass poverty may be considered superior to some alternatives in which fewer people lead very good lives. This property has been used by Parfit (1976, 1982, 1984) to argue against *classical utilitarianism,* the variable-population social-evaluation ordering that ranks utility vectors on the basis of their total utilities. If Parfit's view is accepted, R should be required to avoid the repugnant conclusion.

Avoidance of the repugnant conclusion. There exist $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$ and $\varepsilon \in (0,\xi)$ such that, for all $m > n$,

$$(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin P(R).$$

Blackorby, Bossert, and Donaldson (2006, Theorem 2) show that there exists no variable-population social-evaluation ordering satisfying the above four axioms; see Blackorby, Bossert, and Donaldson (2005), Blackorby, Bossert, Donaldson, and Fleurbaey (1998), Blackorby and Donaldson (1991), Carlson (1998), McMahan (1981), Parfit (1976, 1982, 1984) and Shinotsuka (2008) for similar observations.

The following theorem shows that reflexivity and completeness are not required – transitivity of R is sufficient to generate the impossibility.

Theorem 5. *There exists no transitive variable-population social-evaluation relation satisfying minimal increasingness, weak inequality aversion, Pareto plus and avoidance of the repugnant conclusion.*

Proof. Suppose R satisfies minimal increasingness, weak inequality aversion and Pareto plus. The proof is completed by showing that R leads to the repugnant conclusion. For any population size $n \in \mathbb{N}$, let $\xi, \varepsilon, \delta \in \mathbb{R}_{++}$ be such that $0 < \delta < \varepsilon < \xi$. Choose any integer r such that

$$r > n\frac{(\xi - \varepsilon)}{(\varepsilon - \delta)}. \tag{5}$$

Because both the numerator and denominator on the right-hand side of the inequality are positive, r is positive. By Pareto plus,

$$((\xi \mathbf{1}_n, \delta \mathbf{1}_r), \xi \mathbf{1}_n) \in P(R). \tag{6}$$

Average utility in $(\xi \mathbf{1}_n, \delta \mathbf{1}_r)$ is $(n\xi + r\delta)/(n+r)$ so, by minimal inequality aversion,

$$\left(\left(\frac{n\xi + r\delta}{n+r} \right) \mathbf{1}_{n+r}, (\xi \mathbf{1}_n, \delta \mathbf{1}_r) \right) \in R. \tag{7}$$

By (5),

$$\varepsilon > \frac{n\xi + r\delta}{n+r}$$

and, by minimal increasingness,

$$\left(\varepsilon \mathbf{1}_{n+r}, \left(\frac{n\xi + r\delta}{n+r} \right) \mathbf{1}_{n+r} \right) \in P(R). \tag{8}$$

Combining (6)–(8) and using transitivity, it follows that $(\varepsilon \mathbf{1}_{n+r}, \xi \mathbf{1}_n) \in P(R)$ and avoidance of the repugnant conclusion is violated. □

If transitivity is weakened to Suzumura consistency, the axioms in the theorem statement are compatible. Moreover, three of them can be strengthened and other properties that are commonly imposed in population ethics can be added without obtaining an impossibility.

Expressed in the current setting, the strong Pareto principle is another fixed-population axiom. If everyone alive in two fixed-population alternatives with utility vectors u and v has a utility in u that is at least as high as the utility of this person in v with at least one strict inequality, u is better than v. Clearly, this axiom is a strengthening of minimal increasingness.

Strong Pareto. For all $n \in \mathbb{N}$ and for all $u, v \in \mathbb{R}^n$,

$$u_i \geq v_i \text{ for all } i \in \{1, \ldots, n\} \text{ and } u \neq y \;\Rightarrow\; (u, v) \in P(R).$$

Continuity is a condition that prevents the social-evaluation relation R from exhibiting 'large' changes in response to 'small' changes in a utility vector. Again, the axiom imposes restrictions on fixed-population comparisons only.

Continuity. For all $n \in \mathbb{N}$ and for all $u \in \mathbb{R}^n$, the sets $\{v \in \mathbb{R}^n \mid (v,u) \in R\}$ and $\{v \in \mathbb{R}^n \mid (u,v) \in R\}$ are closed in $\mathbb{R}^n$.

Weak inequality aversion can be strengthened by requiring the restriction of R to $\mathbb{R}^n \times \mathbb{R}^n$ to be *strictly S-concave* for any population size $n \in \mathbb{N}$; see, for instance, Marshall and Olkin (1979). Strict S-concavity is equivalent to the conjunction of the *strict transfer principle* familiar from the theory of inequality measurement and *anonymity.* The strict transfer principle requires that a progressive transfer increases goodness, provided the relative rank of the individuals involved in the transfer is unchanged; see Dalton (1920) and Pigou (1912). A social-evaluation relation is anonymous if the individuals in a fixed population are treated impartially, without paying attention to their identities; see Sen (1973) for a detailed discussion. A *bistochastic $n \times n$ matrix* is a matrix whose entries are in the closed interval $[0,1]$ and all row sums and column sums are equal to one.

Strict S-concavity. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all bistochastic $n \times n$ matrices B,

(i) $(Bu,u) \in R$.
(ii) Bu is not a permutation of $u \;\Rightarrow\; (Bu,u) \in P(R)$.

Independence of the utilities of unconcerned individuals is a fixed-population separability property introduced by d'Aspremont and Gevers (1977). It requires that only the utilities of those who can possibly be affected by a choice between two fixed-population alternatives should determine their ranking.

Independence of the utilities of unconcerned individuals. For all $n,m \in \mathbb{N}$, for all $u,v \in \mathbb{R}^n$ and for all $w,s \in \mathbb{R}^m$,

$$((u,w),(v,w)) \in R \;\Leftrightarrow\; ((u,s),(v,s)) \in R.$$

We now turn to further variable-population axioms. The *negative expansion principle* is dual to Pareto plus. It requires any utility distribution to be ranked as better than one with the ceteris-paribus addition of an individual whose life is not worth living – that is, with a lifetime utility below neutrality.

Negative expansion principle. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $a \in \mathbb{R}_{--}$,

$$(u,(u,a)) \in P(R).$$

Expansion continuity applies the notion of continuity to pairs of utility vectors of different dimension, particularly pairs of vectors whose dimensions differ by one.

Expansion continuity. For all $n \in \mathbb{N}$ and for all $u \in \mathbb{R}^n$, the sets $\{t \in \mathbb{R} \mid ((u,t),u) \in R\}$ and $\{t \in \mathbb{R} \mid (u,(u,t)) \in R\}$ are closed in $\mathbb{R}$.

Note that, in the presence of Pareto plus and the negative expansion principle, expansion continuity implies *existence of critical levels,* requiring that non-trivial trade-offs between population size and well-being are possible in the sense that, for any utility vector $u \in \Omega$, there exists a utility level $c \in \mathbb{R}$ (which may depend on u) such that the ceteris-paribus addition of an individual with utility level c to an existing population with utilities u is a matter of indifference according to R.

Finally, a strengthening of avoidance of the repugnant conclusion is defined. It is obtained by replacing the existential quantifiers in the original axiom with universal quantifiers and replacing the negation of betterness in the conclusion with the negation of the at-least-as-good-as relation. This is a strong property and one might not want to endorse it; the reason why it is used to replace the weaker condition is that it makes the possibility result logically stronger.

Strong avoidance of the repugnant conclusion. For all $n \in \mathbb{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $\varepsilon \in (0,\xi)$ and for all $m > n$,

$$(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin R.$$

We do not impose *avoidance of the sadistic conclusion* or any of its variants (see Arrhenius (2000)) because it is implied by some of the properties already defined.

Theorem 6. *There exists a reflexive and Suzumura-consistent variable-population social-evaluation relation satisfying strong Pareto, continuity, strict S-concavity, independence of the utilities of unconcerned individuals, Pareto plus, the negative expansion principle, expansion continuity and strong avoidance of the repugnant conclusion.*

Proof. An example is sufficient to establish the theorem. Let $g\colon \mathbb{R} \to \mathbb{R}$ be a continuous, increasing and strictly concave function such that $g(0) = 0$ and define the relation R^* by letting, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $v \in \mathbb{R}^m$,

$$\begin{aligned}(u,v) \in R^* \Leftrightarrow & \left[n = m \text{ and } \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)\right] \\ & \text{or } \left[m = n+1 \text{ and } \exists \alpha \in \mathbb{R}_- \text{ such that } v = (u,\alpha)\right] \\ & \text{or } \left[n = m+1 \text{ and } \exists \beta \in \mathbb{R}_+ \text{ such that } u = (v,\beta)\right].\end{aligned}$$

Strong Pareto is satisfied because g is increasing, continuity is satisfied because g is continuous, strict S-concavity follows from the strict concavity of g and independence of the utilities of unconcerned individuals is satisfied because of the additively separable structure of the criterion for fixed-population comparisons. Pareto plus and the negative expansion principle follow immediately from the definition of R^*. Expansion continuity is satisfied because the comparisons involving vectors of dimensions n and $n+1$ for any $n \in \mathbb{N}$ clearly are performed in accordance with this requirement. Strong avoidance of the repugnant conclusion is satisfied because $(\varepsilon \mathbf{1}_m, \xi \mathbf{1}_n) \notin R^*$ for all $n \in \mathbb{N}$, for all $\xi \in \mathbb{R}_{++}$, for all $\varepsilon \in (0, \xi)$ and for all $m > n$. That R^* is reflexive is immediate.

It remains to show that R^* is Suzumura consistent. The first step is to prove that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$ and for all $v \in \mathbb{R}^m$,

$$(u,v) \in R^* \;\Rightarrow\; \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i) \tag{9}$$

and

$$(u,v) \in P(R^*) \;\Rightarrow\; \sum_{i=1}^{n} g(u_i) > \sum_{i=1}^{m} g(v_i). \tag{10}$$

To prove (9), suppose that $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ and $(u,v) \in R^*$. According to the definition of R^*, there are three possible cases.

Case 1. $n = m$ and $\sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$. The conclusion is immediate in this case.

Case 2. $m = n+1$ and $\exists \alpha \in \mathbb{R}_-$ such that $v = (u, \alpha)$. Thus,

$$\sum_{i=1}^{m} g(v_i) = \sum_{i=1}^{n} g(u_i) + g(\alpha) \leq \sum_{i=1}^{n} g(u_i),$$

where the inequality follows because $\alpha \leq 0$ and, by the increasingness of g and the property $g(0) = 0$, $g(\alpha) \leq 0$.

Case 3. $n = m+1$ and $\exists \beta \in \mathbb{R}_+$ such that $u = (v, \beta)$. This implies

$$\sum_{i=1}^{n} g(u_i) = \sum_{i=1}^{m} g(v_i) + g(\beta) \geq \sum_{i=1}^{m} g(v_i),$$

where the inequality follows because $\beta \geq 0$ and thus $g(\beta) \geq 0$.

To prove (10), suppose $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ are such that $(u,v) \in P(R^*)$. Again, there are three cases.

Case 1. $n = m$ and $\sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$. If $\sum_{i=1}^{m} g(v_i) \geq \sum_{i=1}^{n} g(u_i)$, we obtain $(v,u) \in R^*$ and thus a contradiction to our hypothesis $(u,v) \in P(R^*)$. Therefore, $\sum_{i=1}^{n} g(u_i) > \sum_{i=1}^{m} g(v_i)$.

Case 2. $m = n+1$ and $\exists \alpha \in \mathbb{R}_-$ such that $v = (u, \alpha)$. Thus,

$$\sum_{i=1}^{m} g(v_i) = \sum_{i=1}^{n} g(u_i) + g(\alpha) \leq \sum_{i=1}^{n} g(u_i) \tag{11}$$

as established in the proof of (9). If $\alpha = 0$, it follows that $v = (u, 0)$ which leads to $(v,u) \in R^*$, contradicting our hypothesis $(u,v) \in P(R^*)$. Thus, $\alpha < 0$ and $g(\alpha) < 0$ because $g(0) = 0$ and g is increasing. Therefore, the inequality in (11) is strict.

Case 3. $n = m+1$ and $\exists \beta \in \mathbb{R}_+$ such that $u = (v, \beta)$. This implies

$$\sum_{i=1}^{n} g(u_i) = \sum_{i=1}^{m} g(v_i) + g(\beta) \geq \sum_{i=1}^{m} g(v_i) \tag{12}$$

as established in the proof of (9). If $\beta = 0$, it follows that $u = (v, 0)$ which leads to $(v, u) \in R^*$, again contradicting the hypothesis $(u, v) \in P(R^*)$. Thus, $\beta > 0$ and the inequality in (12) is strict.

To complete the proof, suppose $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ are such that $(u, v) \in \mathrm{tc}(R^*)$. By repeated application of (9) and the transitivity of $\geq$, it follows that $\sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$. If $(v, u) \in P(R^*)$, (10) implies $\sum_{i=1}^{m} g(v_i) > \sum_{i=1}^{n} g(u_i)$, a contradiction. Thus, $(v, u) \notin P(R^*)$ and R^* is Suzumura consistent. □

Another impossibility result in population ethics is due to Broome (2004, Chap. 10). Broome suggests that existence is in itself neutral and, thus, the ceteris-paribus addition of an individual to a utility-unaffected population should lead to an equally-good alternative, at least as long as the utility of the added person (if brought into being) is within a non-degenerate interval. This intuition, which Broome calls the *principle of equal existence,* is incompatible with strong Pareto, provided that the social-evaluation relation R is transitive. The impossibility persists if transitivity is weakened to Suzumura consistency. The following axiom is a weak form of the principle of equal existence.

Principle of equal existence. There exist $u \in \Omega$ and distinct $\alpha, \beta \in \mathbb{R}$ such that

$$((u, \alpha), u) \in I(R) \text{ and } ((u, \beta), u) \in I(R). \tag{13}$$

We obtain the following impossibility result.

Theorem 7. *There exists no Suzumura-consistent variable-population social-evaluation relation satisfying strong Pareto and the principle of equal existence.*

Proof. Suppose R satisfies strong Pareto and the principle of equal existence. The proof is completed by showing that R cannot be Suzumura consistent. By the principle of equal existence, there exist $u \in \Omega$ and distinct utility levels α and β such that (13) is satisfied. Without loss of generality, suppose $\alpha > \beta$. By strong Pareto, $((u, \alpha), (u, \beta)) \in P(R)$ which, together with (13), leads to a violation of Suzumura consistency. □

Acknowledgments This paper is dedicated to Kotaro Suzumura in appreciation of his invaluable contribution to the academic community. A referee provided many helpful comments and suggestions. Financial support through a grant from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

References

Arrhenius, G. (2000). An impossibility theorem for welfarist axiologies. *Economics and Philosophy*, *16*, 247–266

Arrow, K. (1951, 2nd ed. 1963). *Social choice and individual values*. New York: Wiley

Blackorby, C., Bossert, W., & Donaldson, D. (2005). *Population issues in social choice theory, welfare economics, and ethics*. Cambridge: Cambridge University Press

Blackorby, C., Bossert, W., & Donaldson, D. (2006). Population ethics and the value of life. In M. McGillivray (Ed.), *Inequality, poverty and well-being.* Basingstoke: Palgrave Macmillan

Blackorby, C., Bossert, W., Donaldson, D., & Fleurbaey, M. (1998). Critical levels and the (reverse) repugnant conclusion. *Journal of Economics*, *67*, 1–15

Blackorby, C., & Donaldson, D. (1991). Normative population theory: A comment. *Social Choice and Welfare*, *8*, 261–267

Blackorby, C., Donaldson, D., & Weymark, J. (1984). Social choice with interpersonal utility comparisons: A diagrammatic introduction. *International Economic Review*, *25*, 327–356

Bossert, W., Sprumont, Y., & Suzumura, K. (2005a). Consistent rationalizability. *Economica*, *72*, 185–200

Bossert, W., Sprumont, Y., & Suzumura, K. (2005b). Maximal-element rationalizability. *Theory and Decision*, *58*, 325–350

Bossert, W., Sprumont, Y., & Suzumura, K. (2006). Rationalizability of choice functions on general domains without full transitivity. *Social Choice and Welfare*, *27*, 435–458

Bossert, W., & Suzumura, K. (2008). Utility in social choice. In K. Basu, & R. Kanbur (Eds.), *Welfare, development, philosophy and social science: Essays for Amartya Sen's 75th birthday, vol. i, welfare economics*. Oxford: Oxford University Press

Bossert, W., & Weymark, J. (2004). Utility in social choice. In S. Barbera, P. Hammond, & C. Seidl (Eds.), *Handbook of utility theory, vol. 2: Extensions.* Dordrecht: Kluwer

Broome, J. (2004). *Weighing lives*. Oxford: Oxford University Press

Carlson, E. (1998). Mere addition and the two trilemmas of population ethics. *Economics and Philosophy*, *14*, 283–306

Dalton, H. (1920). The measurement of the inequality of incomes. *Economic Journal*, *30*, 348–361

d'Aspremont, C., & Gevers, L. (1977). Equity and the informational basis of collective choice. *Review of Economic Studies*, *44*, 199–209

Hammond, P. (1979). Equity in two person situations: Some consequences. *Econometrica*, *47*, 1127–1135

Hansson, B. (1968). Choice structures and preference relations. *Synthese*, *18*, 443–458

Marshall, A., & Olkin, I. (1979). *Inequalities: Theory of majorization and its applications*. New York: Academic Press

McMahan, J. (1981). Problems of population theory. *Ethics*, *92*, 96–127

Parfit, D. (1976). On doing the best for our children. In M. Bayles (Ed.), *Ethics and population.* Cambridge: Schenkman

Parfit, D. (1982). Future generations, further problems. *Philosophy and Public Affairs*, *11*, 113–172

Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press

Pigou, A. (1912). *Wealth and welfare*. London: Macmillan

Richter, M. (1966). Revealed preference theory. *Econometrica*, *41*, 1075–1091

Richter, M. (1971). Rational choice. In J. Chipman, L. Hurwicz, M. Richter, & H. Sonnenschein (Eds.), *Preferences, utility, and demand.* New York: Harcourt Brace Jovanovich

Sen, A. (1973). *On economic inequality*. Oxford: Oxford University Press

Shinotsuka, T. (2008). Remarks on population ethics. In P. Pattanaik, K. Tadenuma, Y. Xu, & N. Yoshihara (Eds.), *Rational choice and social welfare.* New York: Springer

Sikora, R. (1978). Is it wrong to prevent the existence of future generations? In R. Sikora, & B. Barry (Eds.), *Obligations to future generations.* Philadelphia: Temple University Press

Suzumura, K. (1976a). Rational choice and revealed preference. *Review of Economic Studies*, *43*, 149–158

Suzumura, K. (1976b). Remarks on the theory of collective choice. *Economica*, *43*, 381–390
Suzumura, K. (1977). Houthakker's axiom in the theory of rational choice. *Journal of Economic Theory*, *14*, 284–290
Suzumura, K. (1983). *Rational choice, collective decisions and social welfare*. New York: Cambridge University Press
Szpilrajn, E. (1930). Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, *16*, 386–389
Thomson, W. (1990). The consistency principle. In T. Ichiishi, A. Neyman, & Y. Tauman (Eds.), *Game theory and applications.* New York: Academic Press

On the Microtheoretic Foundations of Cagan's Demand for Money Function

Rajat Deb, Kaushal Kishore, and Tae Kun Seo

1 Introduction

An extensive literature, both theoretical (see for instance, Bruno and Fischer (1990), Calvo and Leiderman (1992), Friedman (1971), Goldman (1974), Sargent and Wallace (1973)) and empirical, (see for instance, Aghevli and Khan (1977), Anderson, Bomberger, and Makinen (1988), Babcock and Makinen (1975), Cagan (1956), Christiano (1987), Easterly, Mauro, and Schmidt-Hebbel (1995), Engsted (1993), Metin and Maslu (1999), Michael, Nobay, and Peel (1994), Pickersgill (1968), Salemi and Sargent (1979), Taylor (1991)) has arisen around the special semi-logarithmic demand for money function introduced by Cagan (1956). Cagan's motivation behind the demand for money function was mainly in terms of transactions costs and its relationship to the consumer's ability to affect the real value of cash balances. Cagan argued that the real cost of holding cash balances fluctuates widely enough to account for the dramatic changes in the holding of cash balances observed during hyperinflation. He hypothesized that during periods of hyperinflation the demand for money is almost entirely explained by the variation in the expected rate of change in prices and that changes in expected inflation have the same effect on real balances in percentage terms regardless of the absolute amount of initial cash balances. In other words, during hyperinflations, the demand for money takes the special form: $m = k\mathrm{e}^{-\lambda\pi^{\mathrm{e}}}$, where m is the real demand for money, $\pi^{(\mathrm{e})}$ is the expected rate of inflation and k, λ are positive constants.

The theoretical papers using Cagan's functional form have been written largely in the monetarist tradition, analyzing hyperinflation and the associated problem of "inflation tax." (See, for instance, Calvo and Leiderman (1992), Friedman (1971), Sargent and Wallace (1973).) The use of Cagan's demand for money function has, however, been "ad- hoc" and no attempt has been made to rationalize the function

R. Deb, K. Kishore, and T.K. Seo
Department of Economics, Southern Methodist University, Dallas, TX 75275, USA
e-mail: rdeb@smu.edu

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

in terms of “utility maximizing” behavior. This chapter examines the possibility of providing such a rationalization, without introducing money directly into the consumer’s utility function. We assume that individuals are rational and that money is both a medium of exchange and a store of value, and that the demand for money is a result of intertemporal consumption smoothing. In this framework we try to solve the so-called “integrability problem” by asking the question as to whether Cagan’s special semi-logarithmic form of the demand for money can be generated from some underlying process of utility maximization.

We provide the answer to this question in the context of two different models. The first is a simple two-period utility maximizing model of the type used extensively in “overlapping generations” literature in macroeconomics (Diamond (1965), Samuelson (1958)). The second is a transactions cost/inventory theoretic model of the “Baumol–Tobin” type (Baumol (1952), Tobin (1956)). For the first type of model we discuss and analyze the type of utility function that gives rise to Cagan’s demand for money function. We show that while the function has the “usual” properties assumed in utility theory, no *time separable* utility function *of the type usually used* in overlapping generations models can generate Cagan’s form for the demand for money. For the second type of model, we show that in a “Baumol–Tobin” type of inventory theoretic framework, the demand for money takes Cagan’s form if and only if the transactions cost function in the model takes a specific form. Our results are shown to be valid both in static and fully dynamic versions of this model.

2 Demand for Money: Cagan’s Functional Form

Let m be the real quantity of money, M the nominal quantity of money and P the price level. The usual demand for money function used in macroeconomics posits a positive relation between the real demand for money, $m_t \equiv \frac{M_t}{P_t}$, and *real* income, y_t, and a negative relationship between the real demand for money and the nominal interest rate, i_t. A special semi-logarithmic form of this relationship may be written as, $\ln m_t = \widetilde{k} + \gamma \ln y_t - \lambda i_t$ or equivalently as:

$$\frac{M_t}{P_t} = k\mathrm{e}^{-\lambda i_t} y_t^{\gamma}, \tag{1}$$

where k, γ and λ are positive constants.

Using Fisher’s equation, $i_t = r_t + \pi_t^{\mathrm{e}}$, relating the nominal interest rate to the real interest rate, r_t, and the expected inflation rate, π_t^{e}, one can think of two types of regimes. First, we can have macroeconomic regimes with stable prices where the real interest rate is constant and is primarily determined by the marginal product of capital. Since prices are stable, nominal interest rate is constant too. In such a regime, $\frac{M_t}{P_t} = \widehat{k} y_t^{\gamma}$, where $\widehat{k} \equiv k\mathrm{e}^{-\lambda i_t} > 0$ is a positive constant. If $\gamma = 1$, one gets the classical “quantity theory” of money. A second type of scenario is that of an inflationary environment such as those studied by Cagan (1956) in which hyperinflation prevailed and real income stagnated. In this case, maintaining the assumption that the real rate of interest does not change and is approximately zero, since real income does not change as well, the demand for money takes Cagan's special form and is given by:

$$\frac{M_t}{P_t} = \bar{k}\mathrm{e}^{-\lambda \pi_t^{\mathrm{e}}}, \tag{2}$$

where $\bar{k} = ky_t^{\gamma}$. While this particular functional form for the demand for money has been extremely useful in empirical analyses of money demand during inflationary periods, the following open question remains: Can this form arise from the utility maximizing behavior of a representative agent? We will address this question in the context of two types of standard models used in macroeconomics, the Samuelson–Diamond two-period overlapping generations model and the Baumol–Tobin transactions cost/inventory theoretic model.

3 Model A: The Overlapping Generations Model

Consider a simple model with one good, where a representative agent lives for two periods. The agent's utility function is given by $u(c_1, c_2)$ where c_1 and c_2 represent the agent's consumption in periods 1 and 2, respectively. Money does not enter the utility function and thus has no intrinsic value. The agent receives (real) income, y in the first period. No income is earned in the second period. Consumption in the second period is paid from savings held in the form of money, and the demand for money is thus "derived demand" motivated by consumption smoothing. Let p_1, $p_2 > 0$ be the price of the good in the first and the second periods, respectively, and M the nominal quantity of money. Then, the agent's utility maximization problem can be written as:

$$\max_{c_1,c_2} u(c_1, c_2) \tag{3}$$

such that

$$\begin{aligned} M + p_1 c_1 &= p_1 y \equiv Y, && (4)\\ p_2 c_2 &= M, && (5)\\ c_1, c_2 &\geq 0. \end{aligned}$$

The model that is usually used in macroeconomics is in fact a special case of the above model. In the "standard" overlapping generations model it is generally assumed that the utility function u is additively time separable and that it can be written as sum of utility from consumption in period one and the discounted value of the utility from consumption in period two. Letting the discount factor be $(1+\theta)^{-1}$ with $\theta > 0$, the agent's maximization problem can, in this case, be rewritten as:

$$\max_{c_1,c_2} [u(c_1) + (1+\theta)^{-1} u(c_2)] \tag{6}$$

such that

$$
\begin{aligned}
p_2c_2 + p_1c_1 &= p_1y \equiv Y, \\
p_2c_2 &= M, \\
c_1, c_2 &\geq 0.
\end{aligned}
$$

Under the standard assumption that the agent's expectation is "rational" (Lucas, 1972) and that the expected price in period 2 is the same as that predicted by the agent, the inflation and expected inflation rates, π and π^{e} are identical and $\pi^{\mathrm{e}} = \pi \equiv \frac{p_2}{p_1} - 1$. Using the general form of Cagan's demand for money function (1), assuming that the real interest rate is zero, and letting $k_0 = ky^{\gamma-1}$ we have:[1, 2]

$$
M = k_0 \mathrm{e}^{-\lambda \pi^{\mathrm{e}}} Y, \;\; k_0 > 0, \tag{9}
$$

Solving for c_1 and c_2 from (4) and (5) we get,

$$
c_1 = \frac{Y}{p_1}\left[1 - k_0 \mathrm{e}^{-\lambda \pi^{\mathrm{e}}}\right] \tag{10}
$$

$$
c_2 = \frac{Y}{p_2}\left(k_0 \mathrm{e}^{-\lambda \pi^{\mathrm{e}}}\right) \tag{11}
$$

We will resolve two issues of rationalizability. First, we will demonstrate that there exists a "well-behaved" utility function that generate c_1 and c_2 as described in (10) and (11) as interior solutions for the utility maximization problem described in (3)–(5). Second, we will prove that the utility function that rationalizes Cagan's demand for money function *has no (differentiable) monotonic transformation* that is additively separable. This will establish that Cagan's form cannot be generated as the solution to the utility maximization problem of the "standard" model described in (4)–(6).

To describe a utility function which can rationalize Cagan's form we will introduce two additional functions: g and its inverse h.

Define a function $g : (\max\{0, 1 + \lambda^{-1} \ln k_0\}, \infty) \to (0, \infty)$ as:

$$
g(\xi) = \xi \left[k_0^{-1} \mathrm{e}^{\lambda(\xi - 1)} - 1\right] \tag{12}
$$

---

[1] Notice that k_0 is a constant because in this model y, k and γ are all constants.

[2] Since $\frac{M_t}{P_t} = k\mathrm{e}^{-\lambda i_t} y_t^{\gamma}$ we can write

$$
\begin{aligned}
M_t &= k\mathrm{e}^{-\lambda i_t} P_t y_t y_t^{\gamma-1} && (7) \\
&= k_0 \mathrm{e}^{-\lambda(r_t + \pi_t^{\mathrm{e}})} Y_t && (8)
\end{aligned}
$$

where $k_0 = ky_t^{\gamma-1} = ky^{\gamma-1}$.

It is easy to check[3] that both g and its derivative g' are positive and hence $h \equiv g^{-1}$, the inverse of g, is a well-defined continuous function from $(0,\infty)$ to $(0,\infty)$.

Now, let consumption be given by (10) and (11). Then, we have

$$\frac{c_1}{c_2} = g\left(\frac{p_2}{p_1}\right) = \frac{p_2}{p_1}\left[k_0^{-1}\mathrm{e}^{\lambda\left(\frac{p_2}{p_1}-1\right)} - 1\right],$$

Normalizing the price of the first period consumption to be 1, we can write, $y = \frac{Y}{p_1}$ and $p = \frac{p_2}{p_1}$. Using the budget constraint, we get:

$$c_2 = k_0 \frac{y}{p}\mathrm{e}^{-\lambda(p-1)}.$$

Invoking the integrability condition, we have:

$$\frac{\mathrm{d}\mu}{\mathrm{d}p} = k_0 \frac{\mu}{p}\mathrm{e}^{-\lambda(p-1)} \tag{13}$$

where for the indirect utility function v, μ is the expenditure function, $\mu = \mu(p; v(q,Y))$. Thus, μ gives us the minimum expenditure needed when the price vector is p to obtain the maximum utility when the income is y and the price vector is q. Therefore, from (13):

$$\ln\mu = \int_q^p \frac{k_0}{t}\mathrm{e}^{-\lambda(t-1)}\mathrm{d}t + A$$

Note that $\mu = y$ if $p = q$ and that the constant of integration $A = \ln y$. Hence,

$$\mu = y\exp\left\{\int_q^p \frac{k_0}{t}\mathrm{e}^{-\lambda(t-1)\mathrm{d}t}\right\}$$

where $p = p_2/p_1$ and $y = Y/p_1$. Now, fixing $p = p_2/p_1 = \beta > 1$ and normalizing income to be 1, $\overline{q}_i = \frac{q_i}{Y}$ for $i = 1,2$ gives us

$$\mu = \frac{1}{\overline{q}_1}\exp\left\{\int_{\overline{q}_2/\overline{q}_1}^{\beta} \frac{k_0}{t}\mathrm{e}^{-\lambda(t-1)}\mathrm{d}t\right\},$$

Noting that $\sum \overline{q}_i c_i = 1$, we have $\frac{1}{\overline{q}_1} = c_1 + \frac{\overline{q}_1}{\overline{q}_2}c_2$. Now, substituting $h\left(\frac{c_1}{c_2}\right)$ for $\frac{\overline{q}_1}{\overline{q}_2}$ we get:

$$\widetilde{u}(c_1,c_2) = \left[c_1 + c_2 h\left(\frac{c_1}{c_2}\right)\right]\exp\left\{\int_{h(c_1/c_2)}^{\beta} k_0\left(\frac{1}{s}\right)\mathrm{e}^{-\lambda(s-1)}\mathrm{d}s\right\} \tag{14}$$

[3] Note that $g'(\xi) = k_0^{-1}\mathrm{e}^{\lambda(\xi-1)} + \xi\lambda k_0^{-1}\mathrm{e}^{\lambda(\xi-1)} - 1$. Since ξ, λ and k_0 are all strictly positive, we see that $k_0^{-1}\mathrm{e}^{\lambda(\xi-1)} - 1 > 0$ implies that $g > 0$ and $g'(\xi) > 0$.

Thus, $\xi > 1 + \lambda^{-1}\ln k_0$ is sufficient to ensure that $g > 0$ and $g' > 0$. That $\xi > 1 + \lambda^{-1}\ln k_0$ is implied by the domain of g.

$$= c_2\left[\frac{c_1}{c_2}+h\left(\frac{c_1}{c_2}\right)\right]\exp\left\{\int_{h(c_1/c_2)}^{\beta} k_0\left(\frac{1}{s}\right)\mathrm{e}^{-\lambda(s-1)}\mathrm{d}s\right\}$$
$$= c_2\left[\frac{c_1}{c_2}+h\left(\frac{c_1}{c_2}\right)\right]\exp\left\{\sigma\left(\frac{c_1}{c_2}\right)\right\}, \tag{15}$$

where

$$\sigma\left(\frac{c_1}{c_2}\right) \equiv \int_{h(c_1/c_2)}^{\beta} k_0\left(\frac{1}{s}\right)\mathrm{e}^{-\lambda(s-1)}\mathrm{d}s$$

for some constant $\beta > \max\{0, 1+\lambda^{-1}\ln k_0\}$.

Proposition 1. *The utility function $\widetilde{u}$ in (14) is homogeneous of degree 1 in (c_1, c_2) and is strictly quasi-concave with marginal utilities being positive for both c_1 and c_2. If for this utility function an interior solution to the utility maximization problem (3) exists, then the demand for c_1 and c_2 are given by (10) and (11) and hence the demand for money is given by Cagan's demand for money function (9).*

Proof. Let

$$\phi\left(\frac{c_1}{c_2}\right) \equiv \left[\frac{c_1}{c_2}+h\left(\frac{c_1}{c_2}\right)\right]\exp\left\{\int_{h(c_1/c_2)}^{\beta} k_0\left(\frac{1}{s}\right)\mathrm{e}^{-\lambda(s-1)}\mathrm{d}s\right\}$$

and note that our utility function (14) can be written as:

$$\widetilde{u}(c_1,c_2) = c_2\phi\left(\frac{c_1}{c_2}\right). \tag{16}$$

From (16) it is obvious that $\widetilde{u}$ is homogeneous of degree 1 in (c_1, c_2). Denoting c_1/c_2 by x and using (14), the marginal utilities of $c_1$and c_2 are given by:

$$\widetilde{u}_1 = \phi'(x) \text{ and } \widetilde{u}_2 = \phi(x) - x\phi'(x). \tag{17}$$

By our definitions of g and h, $\xi = h(x)$ if and only if $x = g(\xi) = \xi\left[k_0^{-1}\mathrm{e}^{\lambda(\xi-1)}-1\right]$. Hence, $x = h(x)\left[k_0^{-1}\mathrm{e}^{\lambda(h(x)-1)}-1\right]$. This gives us:

$$x+h(x) = k_0^{-1}h(x)\mathrm{e}^{\lambda(h(x)-1)}. \tag{18}$$

Using (18), observe[4] that $\phi(x)-\phi'(x)x = \phi'(x)h(x)$. Hence, (17) gives us:

4

$$\begin{aligned}\phi'(x) &= [1+h'(x)]\frac{\phi(x)}{x+h(x)} - \phi(x)\frac{k_0\mathrm{e}^{-\lambda(h(x)-1)}h'(x)}{h(x)}\\ &= \phi(x)\left[\frac{1+h'(x)}{x+h(x)} - \frac{h'(x)}{x+h(x)}\right] \text{ by (18)}\\ &= \frac{\phi(x)}{x+h(x)}.\end{aligned}$$

$$\widetilde{u}_2 = \widetilde{u}_1 h\left(\frac{c_1}{c_2}\right) \tag{19}$$

Assume to the contrary that $\widetilde{u}_1 \leq 0$. Now, by (17), if $\widetilde{u}_1 \leq 0$, $\widetilde{u}_2 > 0$. Since, $h > 0$, this contradicts (19). Thus, $\widetilde{u}_1 > 0$ and (using (19)) $\widetilde{u}_2 > 0$.

Furthermore, by (19) along any indifference curve, $\frac{\mathrm{d}c_1}{\mathrm{d}c_2} = -h\left(\frac{c_1}{c_2}\right)$. Hence, we get

$$\frac{\mathrm{d}^2 c_1}{\mathrm{d}c_2^2} = -h'\left(\frac{c_1}{c_2}\right)\left[\frac{c_2\frac{\mathrm{d}c_1}{\mathrm{d}c_2} - c_1}{c_2^2}\right] = -h'\left(\frac{c_1}{c_2}\right)\left[\frac{-h\left(\frac{c_1}{c_2}\right)c_2 - c_1}{c_2^2}\right] > 0.$$

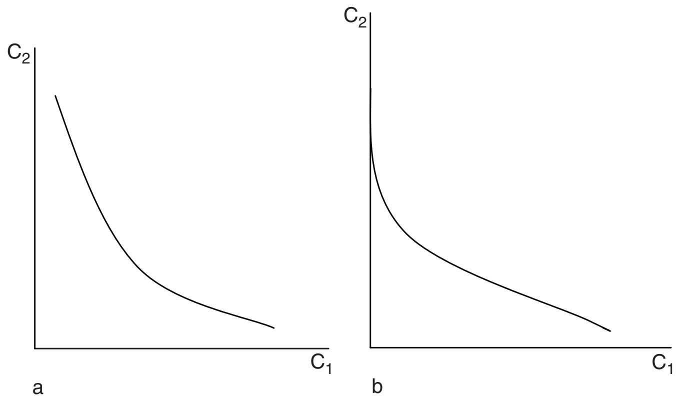
This establishes that the utility function is strictly quasi-concave.

Finally, using (19) and writing down the first-order condition for an interior solution, we have: $h\left(\frac{c_1}{c_2}\right) = \frac{p_2}{p_1}$. In other words,

$$\frac{c_1}{c_2} = g\left(\frac{p_2}{p_1}\right) \equiv \frac{p_2}{p_1}\left[k_0^{-1}\mathrm{e}^{\lambda\left(\frac{p_2}{p_1}-1\right)} - 1\right]\frac{p_2}{p_1}.$$

It is easy to verify, that (10) and (11) satisfy this condition. Hence, using strict quasi-concavity, the unique interior solution to the utility maximizing problem will yield a demand for money function having Cagan's form. □

To understand when an interior solution to our utility maximizing problem will exist note that as $\frac{c_1}{c_2} \longrightarrow 0$, $\frac{\widetilde{u}_2}{\widetilde{u}_1} \longrightarrow \max\left[0, 1+\lambda^{-1}\ln k_0\right]$ and as $\frac{c_1}{c_2} \longrightarrow \infty$, $\frac{\widetilde{u}_2}{\widetilde{u}_1} \longrightarrow \infty$. This implies that two types of indifference curves are possible. Case (a) $1+\lambda^{-1}\ln k_0 \leq 0$: In this case, $\frac{\widetilde{u}_2}{\widetilde{u}_1} \longrightarrow 0$ as $\frac{c_1}{c_2} \longrightarrow 0$, and the indifference curves do not intersect either axis. Case (b) $1+\lambda^{-1}\ln k_0 > 0$: In this case, $\frac{\widetilde{u}_2}{\widetilde{u}_1} \longrightarrow 1+\lambda^{-1}\ln k_0$ as $\frac{c_1}{c_2} \longrightarrow 0$, thus, while indifference curves do not cut the c_1-axis, they do intersect the c_2-axis. In particular, this implies that an interior solution exists in this case if and only if $\frac{p_2}{p_1} > 1+\lambda^{-1}\ln k_0$. The two cases are illustrated in the figure below:



In Case (a) an interior solution will exist for all positive values of p_1 and p_2. Case (b) on the other hand implies that an interior solution exists (and hence, Cagan's form of the money demand function is appropriate) if and only if $\frac{p_2}{p_1} - 1 > \lambda^{-1} \ln k_0$ (i.e., *if and only if the rate of inflation is high enough*). Which of these cases prevails depends on empirical values of the parameters λ and k_0. It is interesting to note that estimates in empirical studies suggest that $1 + \lambda^{-1} \ln k_0 > 0$ and that Case (b) is the more plausible of the two cases. (see for instance, Aghevli and Khan (1977), Anderson et al. (1988), Babcock and Makinen (1975), Cagan (1956), Christiano (1987), Easterly et al. (1995), Engsted (1993), Metin and Maslu (1999), Michael et al. (1994), Pickersgill (1968), Salemi and Sargent (1979), Taylor (1991).)

We have provided an example of an utility function, $\widetilde{u}$, that rationalizes Cagan's form of the demand for money function. Clearly, a necessary and sufficient condition for any utility function to generate this demand function is that it should be a strictly monotonic transformation of $\widetilde{u}$. Now, using this property, we turn to the question of rationalizing Cagan's demand function in the "standard" version of the overlapping generations model with a time separable utility function.

Proposition 2. *There does not exist a function $u : \mathcal{R}_+ \to \mathcal{R}$ and a differentiable strictly monotonic transformation v of $\widetilde{u}$ defined by (14) such that (i) $v(c_1, c_2) = u(c_1) + (1+\theta)^{-1} u(c_2)$ and (ii) v gives rise to Cagan's form of the demand for money function.*

Proof. Assume to the contrary that v is such a monotonic transformation. Then,

$$\ln \frac{v_1}{v_2} = \ln \frac{u_1}{(1+\theta)^{-1} u_2} = \ln u_1(c_1) - \ln (1+\theta)^{-1} u_2(c_2).$$

This implies that $\frac{\partial}{\partial c_1}[\frac{\partial}{\partial c_2} \ln \frac{v_1}{v_2}] \equiv 0$.[5] Since v is a monotone transformation, we would have $\frac{\partial}{\partial c_1}[\frac{\partial}{\partial c_2} \ln \frac{\widetilde{u}_1}{\widetilde{u}_2}] \equiv 0$.

But, from (19), $\ln \frac{\widetilde{u}_1}{\widetilde{u}_2} = \frac{1}{h} = -\ln h$. Hence, we have:

$$\frac{\partial}{\partial c_2} \ln \frac{\widetilde{u}_1}{\widetilde{u}_2} = \frac{h'}{h} \frac{c_1}{c_2^2}.$$

Thus, for $\frac{\partial}{\partial c_1}[\frac{\partial}{\partial c_2} \ln \frac{\widetilde{u}_1}{\widetilde{u}_2}] \equiv 0$, it must be the case that $\frac{h'}{h} c_1$ is a function of c_2 alone, say, $\frac{h'}{h} c_1 \equiv \psi(c_2) c_2^2$. The left-hand side of this equation is homogeneous of degree 1 in (c_1, c_2). This implies that the right-hand side is homogenous of degree 1 in c_2. That is $\psi(c_2) c_2^2 \equiv a c_2$ for some *positive* constant a.[6] Thus, we can write:

$$\frac{h'(x)}{h(x)} = \frac{a}{x}.$$

[5] See Sono's (1961) classic analysis of separability.

[6] Note that a is not equal to zero since this will imply that h is a constant.

Integrating both sides, we have $h(x) \equiv bx^a$, where b is a constant of integration. Thus, $g \equiv h^{-1}(x) = (1/b)^a \xi^{1/a}$. But, comparing this with $\xi \left[k_0^{-1} e^{\lambda(\xi-1)} - 1\right]$ from (12), and letting $\xi \to \infty$, we see that $(\frac{1}{b})^a \xi^{1/a} \equiv \xi \left[k_0^{-1} e^{\lambda(\xi-1)} - 1\right]$ is impossible. □

4 Model B: The Transactions Cost Model

In this model, a representative agent faces two costs: a "transactions cost" and an "opportunity cost of holding money", (see Tobin (1956)). If money is held for transaction purposes, then it cannot be invested and the interest, i, that could have been earned is foregone. Thus, the opportunity cost of holding the stock of money m is im. The transactions cost is some function of the amount of money being held and the level of transactions. The cost of transactions decreases with the amount of money held, but decreases at a decreasing rate. The real income y is a proxy for the volume of transactions and the transactions cost increases with y. Thus, we will assume, $\alpha = \alpha(m,y)$ is the transactions cost and in the neighborhood of the equilibrium holding of money $\alpha_m < 0$, $\alpha_{mm} > 0$ and $\alpha_y > 0$, where the real quantity of money is $m \equiv \frac{M}{P}$.

To hold an optimal inventory of money the agent minimizes the sum of the "transactions cost" and the "opportunity cost of holding money".

$$\min_m \left[im + \alpha(m,y)\right].$$

The first-order condition for an interior minimum is given by:

$$-\alpha_m(m,y) = i. \tag{20}$$

Since $\alpha_{mm} > 0$, in the neighborhood of the equilibrium, the first-order conditions is sufficient for a minimum, and using the implicit function theorem, we can solve (locally) for the optimal level of money demand $\hat{m}$ as a function of i and y. Thus, we have:

$$\hat{m} = \hat{m}(i,y) \text{ with } \hat{m}_i < 0\text{; moreover, } \alpha_{my} < 0 \text{ iff } \hat{m}_y > 0.$$

Thus, in equilibrium, the amount of money held will be a decreasing function of real interest rate, and if $\alpha_{my} < 0$, it will be an increasing function of total income of the agent.

Now, consider the following special transactions cost function (where γ is a positive constant)

$$\alpha(m,y) \equiv \lambda^{-1}\left[m\ln m - m - \gamma m \ln y - m \ln k\right] + k\lambda^{-1} y^{\gamma} + \tau(y), \tag{21}$$

where τ is an arbitrary real valued function of y such that $\tau'(y) \geq 0$.

Replacing $\alpha(m,y)$ with this specific form in (20), the optimum amount of money holding can be calculated. This gives us a generalized version of Cagan's demand for money function as follows:[7]

$$-\alpha_m \equiv -\lambda^{-1}[\ln m - \gamma \ln y - \ln k] = i. \tag{22}$$
$$\text{Or, } \hat{m} = k\mathrm{e}^{-\lambda i} y^{\gamma}.$$

It is easy to check that *at the equilibrium* described by (22), $\alpha_m < 0$, $\alpha_{mm} > 0$ and $\alpha_y > 0$.

Conversely, assume that we have (22). Then, in the neighborhood of the equilibrium, we can define a function $\rho(y)$ such that:

$$\alpha(m,y) \equiv \lambda^{-1}[m\ln m - m - \gamma m \ln y - m \ln k] + \rho(y).$$

Since at the equilibrium, $\alpha_y > 0$, we have $\rho'(y) > \gamma m(\lambda y)^{-1}$. Using (22), we get $\rho'(y) > \gamma k \lambda^{-1} \mathrm{e}^{-\lambda \cdot i} y^{\gamma-1}$. Since the last inequality holds for all $i > 0$, we must have $\rho'(y) \geq \gamma k \lambda^{-1} y^{\gamma-1}$. Thus, defining $\rho(y) \equiv k\lambda^{-1}y^{\gamma} + \tau(y)$, we see that $\tau'(y) \geq 0$. This gives us the specific transactions cost function (21) above. Thus, we have the generalized form $\hat{m} = k\mathrm{e}^{-\lambda \cdot i} y^{\gamma}$ of Cagan's money demand function if and only if the transactions cost has the form (21) in the neighborhood of the equilibrium m.

Note, however, that from (21) we get:

$$\alpha_y = -\gamma m(\lambda y)^{-1} + k\lambda^{-1}y^{\gamma} + \tau'(y)$$

and

$$\alpha_m = \lambda^{-1}[\ln m - \gamma \ln y - \ln k]$$

and that $\alpha_y > 0$ and $\alpha_m < 0$ only if m is "small enough". Thus, for our specific transactions cost function, money holdings in equilibrium need to be small enough in this sense. This is particularly interesting because during periods when the inflation rate is high, households typically minimize their holdings of real balances by moving into non-money assets.

To see that a similar conclusion can be derived from a fully dynamic infinite horizon representative agent model, consider a model where the agent is maximizing his or her lifetime utility given a budget constraint. The quantity of labor is fixed as is the wage rate, and only consumption enters into the instantaneous utility function, $u_t(\cdot)$. A higher consumption level in any period is associated with higher utility. Here, once again, money does not (directly) enter the utility function. However, money does affect the consumption level indirectly. Total wealth of an agent at any particular time period is given by the wage earned along with accumulated savings from previous time periods. Savings in any period can be held either as money (which earns no interest) or in the form of a financial asset f earning a real interest, r. Increasing or decreasing money holdings is costly entailing the cost of switching

[7] It is worth noting, that if one used $\alpha(m,y) = \frac{\alpha_0 y}{2m}$ with $\alpha_0 > 0$ then (20) would give us Baumol's well-known "square root" formula with $\hat{m} = \left(\frac{\alpha_0 y}{2i}\right)^{1/2}$.

between financial assets and money. We assume that this cost has a quadratic structure and is given by $\frac{\eta}{2}[\dot{m}]^2$, for some $\eta > 0$. As in our static model, $\alpha(m(t), y)$ represents the transactions cost at time t. If θ is the subjective discount rate, then the agent's intertemporal utility maximization problem can then be written as:

$$\max \int_0^\infty u_t(c(t))\,\mathrm{e}^{-\theta t}\mathrm{d}t$$

subject to

$$rf(t) + wl = \dot{f} + c(t) + \alpha(m(t), y) + \frac{\eta}{2}[\dot{m}]^2 + \frac{\dot{M}}{P}.$$

The left-hand of the budget constraint represents the agents income: interest earned $(rf(t))$ and wage income (wl). This income can be used either for current consumption $(c(t))$, paying for transactions costs $(\alpha(m(t), y))$, for accumulating (decumulating) interest earning assets $(\dot{f})$ or for increasing (decreasing) real balances $(\frac{\dot{M}}{P})$ and paying the costs associated with changing real balances.

Since $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{M}{P}\right) = \dot{m} = \frac{\dot{M}}{P} - \frac{M}{P}\left(\frac{\dot{P}}{P}\right)$, we can write $\frac{\dot{M}}{P} = \dot{m} + m\pi$. Thus, under perfect foresight, with $\pi = \pi^{\mathrm{e}}$, the budget constraint becomes

$$rf(t) + wl = \dot{f} + c(t) + \alpha(m(t), y) + \frac{\eta}{2}[\dot{m}]^2 + m(t)\pi^{\mathrm{e}} + \dot{m}.$$

Substituting for $c(t)$ in the utility function and using this budget constraint, we get:

$$\max \int_0^\infty u_t\left(rf(t) + wl - \dot{f} - \alpha(m(t), y) - \frac{\eta}{2}[\dot{m}]^2 - m(t)\pi^{\mathrm{e}} - \dot{m}\right)\mathrm{e}^{-\theta t}\mathrm{d}t.$$

The Euler equations for this problem are given by

$$\frac{\partial u_t}{\partial m} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial u_t}{\partial \dot{m}}\right), \tag{23}$$

$$\frac{\partial u_t}{\partial f} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial u_t}{\partial \dot{f}}\right). \tag{24}$$

Substituting (24) into (23) we get:

$$\eta\ddot{m} - r\eta\dot{m} - (\alpha_m + r + \pi^{\mathrm{e}}) = 0. \tag{25}$$

The characteristic equation of the system is given by

$$x^2 - rx - \eta^{-1}\alpha_{mm} = 0.$$

Under our assumption $\alpha_{mm} > 0$, one of the eigenvalues of the system will be positive and the other negative indicating the existence of a saddle point trajectory converging to the equilibrium given by:

$$(\alpha_m + r + \pi^{\mathrm{e}}) = 0,$$

which is the same as (20).

Proposition 3. *The demand for money in equilibrium has a generalized version of Cagan's functional form if and only if in the neighborhood of the equilibrium the transactions cost function $\alpha(m,y)$is given by*

$$\alpha(m,y) \equiv \lambda^{-1}[m\ln m - m - \gamma m \ln y - m\ln k] + k\lambda^{-1}y^{\gamma} + \tau(y),$$

where $\tau'(y) \geq 0$.

5 Conclusion

We have examined the possibility of rationalizing Cagan's functional form for the demand for money. We have shown that in a two period overlapping generations model this demand for money can be derived from an utility function satisfying the usual properties of differentiability, strict quasi-concavity and positivity of marginal utilities. Empirical estimates of the parameters of the model suggest that the functional form arises if and only if the inflation rate is high enough. However, under the usual assumptions that the utility function is time separable, Cagan's form would not arise in this type of model. An alternative way for rationalizing Cagan's money demand function is by using a dynamic inventory cost theoretic model of the Baumol–Tobin type. We have identified the specific transactions cost structure which would give rise to a generalized form of Cagan's demand for money function as a saddle point trajectory of such a model. Once again, we find that the model would be valid only in periods of significant inflation when households prefer other assets and reduce their holdings of money balances.

Acknowledgements We are grateful to Prof. Satya Das, the participants at the conference honoring Prof. Kotaro Suzumura held at Hitotsubashi University in March 2006 and to an anonymous referee for their comments. We remain responsible for any remaining errors.

References

Aghevli, B. B., & Khan, M. S. (1977). Inflationary finance and the dynamics of inflation: Indonesia, 1951–1972. *American Economic Review*, *67*, 390–403

Anderson, R. B., Bomberger, W. A., & Makinen, G. (1988). The demand for money, the "reform effect," and the money supply process in hyperinflation: The evidence from Greece and Hungary reexamined. *Journal of Money, Credit and Banking*, *20*, 653–672

Babcock, J., & Makinen, G. (1975). The Chinese hyperinflation reexamined. *Journal of Political Economy*, *83*, 1259–1267

Baumol, W. (1952). The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics*, *66*, 545–556

Bruno, M., & Fischer, S. (1990). Seigniorage, operating rules, and the high inflation trap. *Quarterly Journal of Economics*, *105*, 353–374

Cagan, P. (1956). *The monetary dynamics of hyperinflation.* In M. Friedman (Ed.)., *Studies in the quantity theory of money*. Chicago, IL: The University of Chicago Press****

Calvo, G. A., & Leiderman, L. (1992). Optimal inflation tax under precommitment: Theory and evidence. *American Economic Review*, *82*, 179–194
Christiano, L. J. (1987). Cagan's model of hyperinflation under rational expectations. *International Economic review*, *28*, 33–49
Diamond, P. (1965). National debt in a neoclassical growth model. *American Economic Review*, *55*, 1126–1150
Easterly, W. R., Mauro, P., & Schmidt-Hebbel, K. (1995). Money demand and seigniorage-maximizing inflation. *Journal of Money, Credit, and Banking*, *27*, 584–603
Engsted, T. (1993). Cointegration and Cagan's model of hyperinflation under rational expectation. *Journal of Money, Credit, and Banking*, *25*, 350–360
Friedman, M. (1971). Government revenue from inflation. *Journal of Political Economy*, *79*, 846–856
Goldman, S. (1974). Hyperinflation and the rate of growth in the money supply. *Journal of Economic Theory*, *5*, 250–257
Lucas, R. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, *4*, 103–124
Metin, K., & Maslu, I. (1999). Money demand, the Cagan model, testing rational expectations vs adaptive expectations: The case of Turkey. *Empirical Economics*, *24*, 415–426
Michael, P., Nobay, A., & Peel, D. A. (1994). The German hyperinflation and the demand for money revisited. *International Economic Review*, *35*, 1–22
Pickersgill, J. (1968). Hyperinflation and monetary reform in the Soviet Union, 1921–26. *Journal of Political Economy*, *76*, 1037–1048
Salemi, M., & Sargent, T. (1979). Demand for money during hyperinflation under rational expectation. *International Economic Review*, *20*, 741–758
Samuelson, P. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, *66*, 467–482
Sargent, T., & Wallace, N. (1973). Rational expectations and the dynamics of hyperinflation. *International Economic Review*, *14*, 328–350
Sono, M. (1961). The effect of price changes on the demand and supply of seperable goods. *International Economic review*, *2*, 239–275
Taylor, M. P. (1991). The hyperinflation model of money demand revisited. *Journal of Money, Credit, and Banking*, *23*, 327–351
Tobin, J. (1956). The interest elasticity of the transactions demand for cash. *Review of Economics and Statistics*, *38*, 241–247

Hicksian Surplus Measures of Individual Welfare Change When There is Price and Income Uncertainty

Charles Blackorby, David Donaldson, and John A. Weymark

1 Introduction

When there is no uncertainty, it is well known that the Hicksian compensating and equivalent variations are exact measures of individual welfare change. That is, the sign of either of these measures of Hicksian consumer's surplus correctly identifies whether a change in prices and income makes an individual consumer better or worse off.[1] It is also well known that Marshallian consumer's surplus is not an exact measure of individual welfare change except under restrictive assumptions.[2]

The use of the expected value of a Hicksian or Marshallian measure of consumer's surplus to evaluate the welfare consequences of price changes in uncertain environments can be traced back to the seminal analysis of Waugh (1944), who showed that under standard assumptions about individual demand, expected

C. Blackorby
Department of Economics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
and
GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France
e-mail: c.blackorby@warwick.ac.uk

D. Donaldson
Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, B.C, V6T 1Z1, Canada
e-mail: dvdd@telus.net

J.A. Weymark
Department of Economics, Vanderbilt University, VU Station B, Box 1819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA
e-mail: john.weymark@vanderbilt.edu

[1] Although both the compensating and equivalent variations are exact measures of individual welfare change, they are defined using different reference prices and, hence, do not, in general, assign the same numerical value to a given change in prices and income.

[2] See, for example, Boadway and Bruce (1984, Chap. 7) and Chipman and Moore (1976, 1980).

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,

Marshallian consumer's surplus and expected compensating variation are both negative if a stochastic price is stabilized at its arithmetic mean. For an individual whose preferences satisfy the expected utility hypothesis, the use of an expected surplus measure, whether it be Hicksian or Marshallian, is a valid measure of individual welfare change under uncertainty if and only if its sign correctly determines whether his expected utility increases or decreases as a result of a change in the distribution of prices and incomes across states. Anderson and Riley (1976) have argued that these expected surplus measures do not correctly track individual preferences when a stochastic price is stabilized unless the marginal utility of income (in the utility representation of preferences used to compute expected utilities) is independent of both the level of income and the value of this price.

Rogerson (1980) and Turnovsky, Shalit, and Schmitz (1980) have identified restrictions on preferences for which expected Marshallian surplus is a valid indicator of individual welfare change when prices and, in the case of Rogerson, incomes are stochastic. For the case in which only one price is uncertain, Helms (1984, 1985) has characterized the restrictions on preferences for which expected compensating variation is a valid measure of individual welfare change, both when the amount of price variability after the change in the distribution of this price is unrestricted and when the stochastic price is stabilized at its mean value. In each case, these restrictions are quite stringent.

In the models considered by Helms, the consumer allocates a certain income over one or more commodities whose prices are certain and one commodity whose price is uncertain. However, whether uncertainty is generated by, for example, trade shocks (Anderson and Riley, 1976) or by factors that affect the volatility of commodity prices (Newbery and Stiglitz, 1981), it is often the case that the incomes of consumers are also uncertain and one or more prices are state dependent. Furthermore, incomes may be directly affected by random events such as health and the timing of a worker's entry into the labor market, in which case the design of social insurance programmes needs to be evaluated. See, for example, Varian (1980).

In this chapter, we extend Helms's analyses by identifying the circumstances under which a consumer's surplus criterion based on a Hicksian surplus measure in each state is a valid measure of individual welfare change when income and some or all of the prices vary across states. For concreteness, we use compensating variations in our analysis, but our theorems are also valid for equivalent variations. Although the mechanism that generates a change in the state distribution of prices and incomes can take many forms, for concreteness, we suppose that it is a government project. To evaluate the welfare consequences of a project for an individual consumer whose preferences satisfy the expected utility hypothesis, we employ a surplus evaluation function that aggregates the ex post compensating variations in each state into an overall surplus measure. Such a surplus evaluation function is a consistent measure of individual welfare change if it is positive valued if and only if the project makes the consumer better off ex ante.

For the kinds of state-dependent prices and incomes that we consider, we show that a consistent measure of individual welfare change based on the ex post compensating variations must regard a project as being welfare improving if and only if its expected compensating variation is positive. Furthermore, the indirect utility function that the consumer uses to evaluate prices and income in each state and that is used to compute expected utilities must be affine in income, with the origin term a constant and the weight on income independent of those prices that are uncertain. These restrictions imply that preferences are homothetic. If all prices are uncertain, these conditions are inconsistent with the homogeneity properties of an indirect utility function and, hence, we obtain an impossibility result.

In Section 2, we describe our state-contingent alternatives model of uncertainty and formally define the compensating variations obtained in each state. We introduce our consistency criterion and the domains that we consider in Section 3. In Section 4, we adapt a result due to Blackorby and Donaldson (1985) in order to provide a partial characterization of the restrictions implied by consistency. A complete characterization of the restrictions implied by consistency on our domains is established in Section 5. In Section 6, we discuss our theorems and relate them to results on consistent measures of welfare change that have been obtained in a variety of different contexts. We provide some concluding remarks in Section 7.

2 Compensating Variations for the State-Contingent Alternatives Model of Uncertainty

We employ the state-contingent alternatives model of uncertainty with a finite number of states due to Arrow (1953, 1964). For discussions of the expected utility theorem for this model, see Arrow (1965), Blackorby, Davidson, and Donaldson (1977), and Diewert (1993).

We assume that there are M states ($M \geq 2$) and let $\mathcal{M} = \{1,\ldots,M\}$ denote the set of states. In each state, there are N commodities ($N \geq 2$). The set of commodities is $\mathcal{N} = \{1,\ldots,N\}$. In state m, the prices of the commodities are $p^m = (p_1^m,\ldots,p_N^m) \in \mathbb{R}^N_{++}$ and the consumer has income $y^m \in \mathbb{R}_+$.[3] Ex ante, the consumer faces the *state-contingent price–income vector* $(p,y) \in \mathbb{R}^{MN}_{++} \times \mathbb{R}^M_+$, where $p = (p^1,\ldots,p^M)$ and $y = (y^1,\ldots,y^M)$.

Ex post consumption in state m is $c^m = (c_1^m,\ldots,c_N^m) \in \mathbb{R}^N_+$. The consumer's ex ante *state-contingent consumption vector* is $c = (c^1,\ldots,c^M) \in \mathbb{R}^{MN}_+$. The probability that state m occurs is $\pi_m > 0$, where $\sum_m \pi_m = 1$. These probabilities can be either subjective or objective, but are fixed throughout our analysis.

We assume that the consumer's preferences over state-contingent commodity vectors in $\mathbb{R}^{MN}_+$ are continuous, strictly monotonic, convex, and satisfy the expected utility hypothesis. Hence, these preferences can be represented by a utility function $U : \mathbb{R}^{MN}_+ \to \mathbb{R}$ for which

$$U(c) = \sum_m \pi_m u(c^m) \tag{1}$$

[3] $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the set of nonnegative and positive numbers, respectively.

for all $c \in \mathbb{R}^{MN}_{+}$, where the function $u\colon \mathbb{R}^{N}_{+} \to \mathbb{R}$ is continuous, increasing in each of its arguments, concave, and state independent. Following Arrow (1965), u is called a *Bernoulli utility function*.[4] Note that u represents preferences over ex post consumption bundles. Any increasing transform of u also represents these preferences. However, only increasing affine transforms of u are Bernoulli utility functions; that is, only increasing affine transforms of u can be used to represent the ex ante preferences in the expected utility form given in (1).

The *Bernoulli indirect utility function* $v\colon \mathbb{R}^{N}_{++} \times \mathbb{R}_{+} \to \mathbb{R}$ is defined by setting

$$v(p^m, y^m) = \max_{c^m \in \mathbb{R}^N_+} \{u(c^m) \mid p^m c^m \leq y^m\} \tag{2}$$

for all $(p^m, y^m) \in \mathbb{R}^{N}_{++} \times \mathbb{R}_{+}$. Hence, the consumer preferences for state-contingent price–income vectors can be represented by the *indirect expected utility function* $V\colon \mathbb{R}^{MN}_{++} \times \mathbb{R}^{M}_{+} \to \mathbb{R}$ defined by setting

$$V(p, y) = \sum_m \pi_m v(p^m, y^m) \tag{3}$$

for all $(p, y) \in \mathbb{R}^{MN}_{++} \times \mathbb{R}^{M}_{+}$. It follows from our assumptions on u that the function v is continuous, decreasing, and convex in prices, increasing in income, and homogeneous of degree zero in prices and income.[5]

When there is no price or income uncertainty, with $p = (p^0, \ldots, p^0)$ and $y = (y^0, \ldots, y^0)$ say, then

$$V(p, y) = v(p^0, y^0). \tag{4}$$

Thus, v represents preferences over certain price–income vectors. As is the case with u, any increasing transform of v represents these preferences over certain outcomes, but only increasing affine transforms of v can be used to compute expected utilities as in (3).

Suppose that the price–income pair in state m is initially $(\bar{p}^m, \bar{y}^m)$ and, therefore, the consumer has utility $\bar{u}^m = v(\bar{p}^m, \bar{y}^m)$ in state m. Now consider changing this price–income pair to $(\hat{p}^m, \hat{y}^m)$. The consumer then has utility $\hat{u}^m = v(\hat{p}^m, \hat{y}^m)$ in this state. The *compensating variation* associated with this change,

$$s^m = S^m(\bar{p}^m, \bar{y}^m, \hat{p}^m, \hat{y}^m), \tag{5}$$

is the maximum amount the consumer would pay for the change. It is defined implicitly by

$$v(\hat{p}^m, \hat{y}^m - s^m) = v(\bar{p}^m, \bar{y}^m) = \bar{u}^m. \tag{6}$$

Note that

[4] A Bernoulli utility function in the state-contingent alternatives model of uncertainty is the analogue of a von Neumann and Morgenstern (1944) utility function in the lottery model of uncertainty.

[5] The function v is decreasing in prices if the value of v decreases when the price of every good is increased.

$$u^m = v(p^m, y^m) \leftrightarrow y^m = e(p^m, u^m), \tag{7}$$

where e is the expenditure function dual to u. Thus, the compensating variation in state m can be written as

$$\begin{aligned} s^m &= e(\hat{p}^m, \hat{u}^m) - e(\hat{p}^m, \bar{u}^m) \\ &= \hat{y}^m - e(\hat{p}^m, \bar{u}^m) \\ &= [\hat{y}^m - \bar{y}^m] + [e(\bar{p}^m, \bar{u}^m) - e(\hat{p}^m, \bar{u}^m)]. \end{aligned} \tag{8}$$

Because the expenditure function is increasing in its last argument,

$$s^m \geq 0 \leftrightarrow \hat{u}^m \geq \bar{u}^m \quad \text{for all} \quad m \in \mathcal{M}. \tag{9}$$

Therefore, this state-specific measure of willingness-to-pay is nonnegative if and only if the consumer is no worse off in state m as a result of the change from $(\bar{p}^m, \bar{y}^m)$ to $(\hat{p}^m, \hat{y}^m)$. Hence, the compensating variation correctly identifies whether a change in prices and income in a given state makes the consumer better off or not. This observation is simply a reflection of the well-known fact that the compensating variation is a valid indicator of individual welfare change when there is no uncertainty.

3 Consistency

A project affects the consumer by changing the vector of state-contingent prices and incomes. Let $(\bar{p}, \bar{y})$ (respectively, $(\hat{p}, \hat{y})$) denote the pre-project (respectively, post-project) prices and incomes. This project changes the consumer's indirect expected utility from $V(\bar{p}, \bar{y})$ to $V(\hat{p}, \hat{y})$. We assume that the same set D of vectors of state-contingent prices and incomes are possible both before and after the implementation of a project.

The question is whether the vector of state-contingent compensating variations $s = (s^1, \ldots, s^M)$ defined in (8) can be used to measure the change in the well-being of the consumer for a project that changes $(\bar{p}, \bar{y})$ to $(\hat{p}, \hat{y})$. More precisely, for the domain D, we ask if there exists some real-valued function of the state-contingent compensating variations that is positive valued for projects that improve the well-being of the consumer and that is nonpositive for those that do not. This *surplus evaluation function* is a function $\Gamma \colon S(D) \to \mathbb{R}$, where $S(D) \subseteq \mathbb{R}^M$ is the set of vectors of state-contingent compensating variations that are achievable when the domain is D. We assume that Γ is continuous and increasing.

Consistency of the surplus evaluation function with consumer well-being on the domain D is defined as follows.

Consistency. (Γ, V) is consistent on $D \subseteq \mathbb{R}^{MN}_{++} \times \mathbb{R}^M_+$ if and only if

$$\Gamma(s^1, \ldots, s^M) \geq 0 \leftrightarrow V(\hat{p}, \hat{y}) \geq V(\bar{p}, \bar{y}) \tag{10}$$

for all $(\bar{p},\bar{y}),(\hat{p},\hat{y}) \in D$.

Let

$$D_y = \mathbb{R}^M_+ \tag{11}$$

and, for all $K \subseteq \mathcal{N}$, let

$$D^K_p = \{p \in \mathbb{R}^{MN}_{++} \mid \forall j \in K, \forall m,m' \in \mathcal{M},\ p^m_j = p^{m'}_j\} \tag{12}$$

and

$$D^K = D^K_p \times D_y. \tag{13}$$

The sets D^K, $K \subseteq \mathcal{N}$, are the domains that we consider for our project evaluations.

For the domain D^K, the pair (Γ,V) has to be consistent for all nonnegative incomes and for all positive prices for which the prices of the goods with indices in the set K are the same in each state. In this domain, the prices with indices in $\mathcal{N} \setminus K$ are permitted to differ across states. Clearly, the more prices that are permitted to differ across states, the more restrictions that Γ and V must satisfy.

4 A Useful Lemma

By interpreting $\mathcal{M}$ as a set of individuals, instead of a set of states, Blackorby and Donaldson (1985) have defined an *indirect Bergson–Samuelson social welfare function* $V^{\mathrm{BS}} \colon \mathbb{R}^{MN}_{++} \times \mathbb{R}^M_+ \to \mathbb{R}$ by setting

$$V^{\mathrm{BS}}(p,y) = W(v^1(p^1,y^1),\ldots,v^M(p^M,y^M)) \tag{14}$$

for all $(p,y) \in \mathbb{R}^{MN}_{++} \times \mathbb{R}^M_+$, where $W \colon \mathbb{R}^M \to \mathbb{R}$ is a continuous, increasing Bergson–Samuelson social welfare function and, for all $m \in \mathcal{M}$, p^m are the prices person m faces, y^m is his income, and $v^m \colon \mathbb{R}^N_{++} \times \mathbb{R}_+ \to \mathbb{R}$ is his indirect utility function. Note that individuals may face different prices in (14). As above, we can compute a compensating variation for each individual $m \in \mathcal{M}$ (using the function v^m instead of v) and define consistency (using V^{BS} instead of V) as in (10).

Blackorby and Donaldson (1985) have shown that for the domains $D^{\varnothing}$ (all prices can be person specific) and $D^{\mathcal{N}}$ (no price can be person specific), consistency implies that the vector of individual incomes must be separable from the prices in the indirect Bergson–Samuelson social welfare function and that for every vector of compensating variations in the domain of Γ, the sign of the surplus evaluation function Γ must be the same as the sign of a linear function of the individual surplus measures.[6] Their proofs apply equally well to any domain D^K with $K \subseteq \mathcal{N}$.

Because our indirect expected utility function is formally a special case of Blackorby and Donaldson's indirect Bergson–Samuelson social welfare function

[6] The separability result for the domain $D^{\mathcal{N}}$ was first established by Roberts (1980, Proposition 1).

(with $v^m = \pi_m v$ and $V^{\text{BS}} = V$), their results also hold for our model. Hence, consistency implies that the state-contingent incomes must be separable from the state-contingent prices in the indirect expected utility function of the consumer and that for every vector of state-contingent compensating variations in $S(D)$, the surplus evaluation function Γ must have the same sign as a linear function of these compensating variations.

Lemma 1. *For all $K \subseteq \mathcal{N}$, if (Γ, V) is consistent on D^K, then* (i) *for every $(p,y) \in D^K$, V can be written as*

$$V(p,y) = \bar{V}(p, \phi(y)), \tag{15}$$

where $\bar{V}$ is continuous, increasing in $\phi(y)$, and homogeneous of degree zero in (p,y) and (ii) *there exist $a_m > 0$ for all $m \in \mathcal{M}$ such that ϕ can be written as*

$$\phi(y) = \sum_m a_m y^m \tag{16}$$

for all $y \in \mathbb{R}^N_+$. Furthermore,

$$\Gamma(s^1, \ldots, s^M) \geq 0 \leftrightarrow \sum_m a_m s^m \geq 0 \tag{17}$$

for all $(s^1, \ldots, s^M) \in S(D^K)$.

For a formal proof of Lemma 1, see Blackorby and Donaldson (1985, Lemma 1, Theorem 1, and Corollary 1.2). The proof strategy is as follows. By considering a project in which only the state-contingent incomes change, (8) implies that the compensating variation in any state is simply the difference between the new and the old income. Because the left side of (10) is independent of prices for such a project, so is the right side, from which the separability result in (15) follows. The homogeneity of degree zero of V implies that ϕ can be chosen to be homogeneous of degree one. Using this homogeneity property, it can be shown that ϕ satisfies an additive Cauchy equation, whose solution is given by (16).[7] Because prices have not been changed, the sign of the change in indirect expected utility is the same as the sign of the change in the value of ϕ, from which (17) follows.

As we have seen, when a project changes only incomes but not prices, the compensating variation in a state is equal to the difference between the pre- and post-project incomes in that state. Thus, the surplus evaluation function must ignore information about income levels. What Lemma 1 demonstrates is that in order for this function to be sensitive only to income differences and at the same time respect the homogeneity properties of the indirect expected utility function, it must assign each state a weight and then use these weights to compute a weighted sum of compensating variations.

[7] See Aczél (1969, Chap. 2) or Eichhorn (1978, Chap. 1) for an introduction to Cauchy equations.

5 The Theorems

Lemma 1 has identified some restrictions that must be satisfied by the indirect expected utility function V and by the surplus evaluation function Γ if they are to be consistent with each other. However, we have yet to identify the restrictions implied by consistency on the Bernoulli indirect utility function v or on the choice of the weights that are used to aggregate the state-contingent compensating variations, other than that these weights are positive. In this section, we characterize these restrictions.

The proof of Lemma 1 does not exploit the assumption that $V(p,y)$ is the expected value of the Bernoulli utilities $v(p^m,y^m)$ obtained in each state. The conclusions in this lemma are also valid if the ex ante utility $V(p,y)$ is any continuous, increasing function of the ex post utilities $v(p^m,y^m)$. We now show that Lemma 1 and the assumption that the consumer's preferences satisfy the expected utility hypothesis imply (i) that the Bernoulli utility function v must be affine in income, with the origin term a constant and the weight on income possibly price dependent, and (ii) for all vectors of state-contingent compensating variations in $S(D)$, the sign of the surplus evaluation function must be the same as the sign of the expected value of these compensating variations. The first of these conditions implies that the Bernoulli direct utility function u is homothetic.

Theorem 1. *For all $K \in \mathcal{N}$, if (Γ, V) is consistent on D^K, then there exists a function $\alpha\colon \mathbb{R}^N_{++} \to \mathbb{R}_{++}$ and a scalar β for which*

$$v(p^0,y^0) = \alpha(p^0)y^0 + \beta \tag{18}$$

for all $(p^0,y^0) \in \mathbb{R}^N_{++} \times \mathbb{R}_+$, where α is continuous, decreasing, convex, and homogeneous of degree minus one. Furthermore,

$$\Gamma(s^1,\ldots,s^M) \geq 0 \leftrightarrow \sum_m \pi_m s^m \geq 0 \tag{19}$$

for all $(s^1,\ldots,s^M) \in S(D^K)$.

Proof. From (3), (15), and (16), we obtain

$$\bar{V}\left(p, \sum_m a_m y^m\right) = \sum_m \pi_m v(p^m,y^m). \tag{20}$$

Consider any $p^0 \in \mathbb{R}^N_{++}$ and let $p^* = (p^0,\ldots,p^0)$. That is, there is no price uncertainty. Define

$$z^m := a_m y^m \quad \text{for all} \quad m \in \mathcal{M}, \tag{21}$$

$$\hat{v}^m(z^m) := \pi_m v(p^0,y^m), \tag{22}$$

and

$$\hat{V}\left(\sum_m z^m\right) := \bar{V}\left(p^*, \sum_m z^m\right). \tag{23}$$

Substituting (21)–(23) into (20) yields

$$\hat{V}\left(\sum_m z^m\right) = \sum_m \hat{v}^m(z^m). \tag{24}$$

Equation (24) is a Pexider equation whose solution is

$$\hat{v}^m(z^m) = \bar{\alpha}(p^0)z^m + \bar{\beta}^m(p^0) \quad \text{for all} \quad m \in \mathcal{M}, \tag{25}$$

where $\bar{\alpha}(p^0) > 0$ because $\hat{v}^m$ is increasing in z^m.[8]

Note that

$$\bar{\beta}^m(p^0) = \hat{v}^m(0) = \pi_m v(p^0, 0) \quad \text{for all} \quad m \in \mathcal{M}. \tag{26}$$

Define

$$\beta(p^0) := v(p^0, 0). \tag{27}$$

From (26) and (27), we obtain

$$\bar{\beta}^m(p^0) = \pi_m \beta(p^0). \tag{28}$$

Substituting (25) and (28) into (22) and using (21) yields

$$\pi_m v(p^0, y^m) = \bar{\alpha}(p^0) a_m y^m + \pi_m \beta(p^0) \tag{29}$$

or, equivalently,

$$v(p^0, y^m) = \bar{\alpha}(p^0)\left[\frac{a_m}{\pi_m}\right] y^m + \beta(p^0). \tag{30}$$

Because v is state independent and $D_y = \mathbb{R}^M_+$, (30) implies that

$$a_m = \kappa \pi_m \quad \text{for all} \quad m \in \mathcal{M}, \tag{31}$$

where $\kappa > 0$ because v is increasing in y^m. Defining

$$\alpha(p^0) := \kappa \bar{\alpha}(p^0) \tag{32}$$

yields

$$v(p^0, y^0) = \alpha(p^0) y^m + \beta(p^0). \tag{33}$$

For v to be homogenous of degree zero in prices and income, α must be homogenous of degree minus one and β must be homogenous of degree zero.

Preferences that are representable by an indirect utility function with the functional form given in (33) are called quasi-homothetic. That is, they exhibit the Gorman (1961) polar form. For the demands generated by these preferences to be nonnegative for all prices and incomes, β must be independent of prices. See

---

[8] See Aczél (1969, Chap. 3) or Eichhorn (1978, Sect. 3.1).

Blackorby, Boyce, and Russell (1978).[9] Thus, v must satisfy (18). The other properties of α follow straightforwardly from the corresponding properties of v.

Using (31), (17) yields (19). □

As indicated in the proof of Theorem 1, if v satisfies (30), then it cannot be an indirect utility function unless the function α satisfies the restrictions stated in Theorem 1 and β is a constant. However, if we choose a utility level $\bar{u}$ for ex post utility such that the consumption of any good is positive if this utility level is achieved, then it is not necessary for β to be a constant if (Γ, V) is only required to be consistent on the subset of D^K for which ex post utilities are at least $\bar{u}$. Instead, β only needs to be continuous, nonincreasing, convex, and homogeneous of degree zero in prices.

Some intuition for Theorem 1 can be obtained by considering the special case in which the Bernoulli indirect utility function v is differentiable. Suppose that there is no price uncertainty and that a project only changes the state-contingent incomes marginally. Let p^0 be the price vector in each state both before and after the project is implemented, y be the initial state-contingent income vector, and $\mathrm{d}y = (\mathrm{d}y^1, \ldots, \mathrm{d}y^M)$ be the vector of income changes that result from this project. Note that $\mathrm{d}y$ is also the vector of compensating variations associated with this project. By Lemma 1, consistency requires that

$$\sum_m a_m \,\mathrm{d}y^m \geq 0 \leftrightarrow \sum_m \pi_m v_y(p^0, y^m)\,\mathrm{d}y^m \geq 0. \tag{34}$$

Because the left side of (34) does not depend on the level of y, in order for this equivalence to hold for all $y \in \mathbb{R}^M_+$ and all $\mathrm{d}y$ in a neighborhood of the origin for which $y + \mathrm{d}y \in \mathbb{R}^M_+$, the marginal utility of income function v_y must be positive and cannot depend on income. Hence, v is an increasing affine function of income. That is, v satisfies (33). Using (33), (34) simplifies to

$$\sum_m a_m \,\mathrm{d}y^m \geq 0 \leftrightarrow \alpha(p^0) \sum_m \pi_m \,\mathrm{d}y^m \geq 0, \tag{35}$$

which can only hold for all $\mathrm{d}y$ in a neighborhood of the origin if the weights $a = (a_1, \ldots, a_M)$ are proportional to the probabilities $\pi = (\pi_1, \ldots, \pi_M)$.

A striking feature of Theorem 1 is that consistency requires that projects be evaluated in terms of expected compensating variation. That is, a project is welfare improving for an individual if and only if the expected compensating variation of the project is positive. As we have noted, previous studies of cost-benefit analysis under uncertainty for a single individual simply assume that the surplus evaluation function is the expected value of some measure of consumer's surplus. For the domains we are considering, we have shown that this must be the case, at least when

---

[9] This result can be shown quite easily for the case in which v is differentiable. By Euler's Theorem, $\sum_{i=1}^n p_i^0 \partial\beta(p^0)/\partial p_i^0 = 0$. Hence, if β is not independent of prices, there must exist some price vector $\bar{p}^0$ and good j for which $\partial\beta(\bar{p}^0)/\partial p_j^0 > 0$. Using Roy's Identity, the demand for good j at $(\bar{p}^0, y^m)$ is $c_j(\bar{p}^0, y^m) = -[y^m \partial\alpha(\bar{p}^0)/\partial p_j^0 + \partial\beta(\bar{p}^0)/\partial p_j^0]/\alpha(\bar{p}^0)$, which is negative when y^m is sufficiently close to 0.

surplus is measured using the compensating variation. In particular, it is not possible to require the surplus evaluation function to exhibit inequality aversion in the distribution of compensating variations across states.

In Theorem 1, K is the set of goods for which prices are certain across states. If $K = \mathcal{N}$, then there is no price uncertainty, whereas if $K = \varnothing$, then all prices can vary across states. Note that any $p \in D_p^K$ can be written as

$$p = \left(p_K^0, p_{-K}^1, \ldots, p_K^0, p_{-K}^M\right), \tag{36}$$

where p_K^0 are the prices of the goods that are certain across states.

Theorem 2 shows that the weight on income in the Bernoulli indirect utility function (18) can only depend on the prices of the goods that are certain. Furthermore, requiring the Bernoulli indirect utility function v to satisfy this restriction and requiring the surplus evaluation function to identify a project as being welfare improving if and only if the expected compensating variation is positive are jointly necessary and sufficient for (Γ, V) to be consistent on any of the domains we are considering, except for the domain in which all prices are uncertain.

Theorem 2. *For all $K \in \mathcal{N} \backslash \varnothing$, (Γ, V) is consistent on D^K if and only if* (18) *and* (19) *hold and there exists a function $\alpha_K : \mathbb{R}_{++}^{|K|} \to \mathbb{R}_{++}$ for which*

$$\alpha(p_K^0, p_{-K}^0) = \alpha_K(p_K^0) \tag{37}$$

for all $(p_K^0, p_{-K}^0) \in \mathbb{R}_{++}^N$, where α_K is continuous, decreasing, convex, and homogeneous of degree minus one.

Proof. Suppose that (Γ, V) is consistent on D^K. From (20) and Theorem 1, we have

$$\bar{V}\left(p, \sum_m \pi_m y^m\right) = \sum_m \pi_m \left[\alpha(p^m) y^m + \beta\right] \tag{38}$$

for all $(p, y) \in D^K$.

Consider any $j \notin K$ and, contrary to the theorem, suppose that there exist distinct $p^{m'}, p^{m''} \in \mathbb{R}_{++}^N$ for which $p_i^{m'} = p_i^{m''}$ for all $i \neq j$ and $\alpha(p^{m'}) \neq \alpha(p^{m''})$. Consider any $\bar{p} \in D_p^K$ for which $\bar{p}^{m'} = p^{m'}$ and $\bar{p}^{m''} = p^{m''}$. Next, consider any distinct $\bar{y}, \hat{y} \in D_y$ for which $\bar{y}^m = \hat{y}^m$ for all $m \neq m', m''$ and

$$\pi_{m'} \bar{y}^{m'} + \pi_{m''} \bar{y}^{m''} = \pi_{m'} \hat{y}^{m'} + \pi_{m''} \hat{y}^{m''}. \tag{39}$$

By construction, the value of the left side of (38) is the same when evaluated at $(\bar{p}, \bar{y})$ and $(\bar{p}, \hat{y})$. Thus, (38) implies that

$$\begin{aligned} &\pi_{m'} \alpha(p^{m'}) \bar{y}^{m'} + \pi_{m''} \alpha(p^{m''}) \bar{y}^{m''} \\ &\qquad = \pi_{m'} \alpha(p^{m'}) \hat{y}^{m'} + \pi_{m''} \alpha(p^{m''}) \hat{y}^{m''}. \end{aligned} \tag{40}$$

Because (40) must hold for any nonnegative $\bar{y}^{m'}$, $\bar{y}^{m''}$, $\hat{y}^{m'}$, and $\hat{y}^{m''}$ that satisfy (39), it follows that $\alpha(p^{m'}) = \alpha(p^{m''})$, a contradiction. Thus, (37) is satisfied.

The necessity part of the argument is completed by noting that the properties of α_K in the theorem statement follow immediately from the properties of α in Theorem 1.

Now, suppose that (18), (19), and (37) are satisfied. Consider any $(\bar{p},\bar{y})$, $(\hat{p},\hat{y}) \in D^K$, where $(\bar{p},\bar{y}) = \left(\bar{p}^0_K,\bar{p}^1_{-K},\ldots,\bar{p}^0_K,\bar{p}^M_{-K},\bar{y}^1,\ldots,\bar{y}^M\right)$ and $(\hat{p},\hat{y}) = \left(\hat{p}^0_K,\hat{p}^1_{-K},\ldots,\hat{p}^0_K,\hat{p}^M_{-K},\hat{y}^1,\ldots,\hat{y}^M\right)$. Then,

$$\begin{aligned}\bar{V}(\hat{p},\hat{y}) - \bar{V}(\bar{p},\bar{y}) &= \sum_m \pi_m v(\hat{p}^0_K,\hat{p}^m_{-K},\hat{y}^m) - \sum_m \pi_m v(\bar{p}^0_K,\bar{p}^m_{-K},\bar{y}^m) \\ &= \sum_m \pi_m \left[\alpha_K(\hat{p}^0_K)\hat{y}^m - \alpha_K(\bar{p}^0_K)\bar{y}^m\right]. \end{aligned} \tag{41}$$

Hence,

$$\bar{V}(\hat{p},\hat{y}) - \bar{V}(\bar{p},\bar{y}) \geq 0 \leftrightarrow \sum_m \pi_m \left[\alpha_K(\hat{p}^0_K)\hat{y}^m - \alpha_K(\bar{p}^0_K)\bar{y}^m\right] \geq 0. \tag{42}$$

From (6), the compensating variation s^m in state m is defined implicitly by

$$\alpha_K(\hat{p}^0_K)\left[\hat{y}^m - s^m\right] + \beta = \alpha_K(\bar{p}^0_K)\bar{y}^m + \beta. \tag{43}$$

Thus,

$$s^m = \frac{1}{\alpha_K(\hat{p}^0_K)}\left[\alpha_K(\hat{p}^0_K)\hat{y}^m - \alpha_K(\bar{p}^0_K)\bar{y}^m\right]. \tag{44}$$

It follows from (42) and (44) that

$$V(\hat{p},\hat{y}) - V(\bar{p},\bar{y}) \geq 0 \leftrightarrow \sum_m \pi_m s_m \geq 0, \tag{45}$$

which completes the sufficiency argument. □

In Theorem 2, we have assumed that there is at least one price that is certain. If all prices and incomes can be stochastic, then there is no surplus evaluation function Γ that can provide a consistent cost-benefit test on $D^{\varnothing}$ for an individual whose preferences satisfy the expected utility hypothesis.

Theorem 3. *There is no function $\Gamma\colon S(D^{\varnothing}) \to \mathbb{R}$ such that (Γ,V) is consistent on $D^{\varnothing}$.*

Proof. The necessity part of the proof of Theorem 2 applies equally well when $K = \varnothing$. As a consequence, α must be independent of all prices. That is, for all $(p^0,y^0) \in \mathbb{R}^N_{++} \times \mathbb{R}_+$,

$$v(p^0,y^0) = \xi y^0 + \beta \quad \text{for some} \quad \xi > 0. \tag{46}$$

However, if v has this functional form, then it cannot be homogeneous of degree zero and, hence, consistency is impossible. $\square$

It is also possible to use Theorem 1 to prove Theorem 3 without relying on Theorem 2. Suppose that (Γ, V) is consistent on $D^{\varnothing}$. Project 1 changes the prices and incomes from $(\bar{p}, \bar{y})$ to $(\hat{p}, \hat{y})$. These prices and incomes can be chosen so that the expected compensating variation $\sum_m \pi_m s^m$ of this project is negative and the compensating variation s^M in state M is positive for the individual under consideration. Because (Γ, V) is consistent, Theorem 1 implies that the change in expected utility is negative and, hence, this individual is worse off as a result of Project 1. Now, let $(\tilde{p}, \tilde{y})$ be the vector of prices and incomes for which $(\tilde{p}^m, \tilde{y}^m) = (\hat{p}^m, \hat{y}^m)$ for all $m \neq M$ and $(\tilde{p}^M, \tilde{y}^M) = \lambda(\hat{p}^M, \hat{y}^M)$, where $0 < \lambda \neq 1$. In Project 2, prices and incomes are changed from $(\bar{p}, \bar{y})$ to $(\tilde{p}, \tilde{y})$. Because the Bernoulli indirect utility function v is homogenous of degree zero in prices and income, the indirect expected utility is the same with $(\tilde{p}, \tilde{y})$ as it is with $(\hat{p}, \hat{y})$. Therefore, Project 2 makes the individual worse off. Let $\tilde{s}^M$ be the compensating variation in state M for this project. Because the expenditure function is homogeneous of degree one in prices, it follows from (8) that $\tilde{s}^M = \lambda s^M$. The compensating variations for the other states are the same with both projects. Because $\pi_M > 0$, by choosing λ to be sufficiently large, the expected compensating variation for Project 2 is positive, violating consistency.

Because the Bernoulli indirect utility function v is homogeneous of degree zero in prices and income, expected utility is unaffected if the prices and income in each state are divided by the price of good one. With this price normalization, the price of good one is certain and always equal to one. It might seem then that Theorem 3 contradicts the special case of Theorem 2 in which $K = \{1\}$. However, this is not the case. While such a price normalization does not change the expected utility either before or after a project is implemented, it does change the value of the compensating variation in any state for which the post-project price of good one is not initially equal to one. In other words, normalizing by setting the price of good one so that it is always equal to one is innocuous from the perspective of calculating expected utility, but it is not innocuous from the perspective of calculating expected compensated variation. As we have seen, it is for precisely this reason that an impossibility result is obtained when all prices and income can vary across states because we can scale the prices and income in any state without changing the prices and income in any other state. This independent scaling is not possible if any price must have the same value in every state.[10]

[10] Also note that the domain obtained from $D^{\varnothing}$ by normalizing the price of good one is not the same as $D^{\{1\}}$. For the domain $D^{\{1\}}$, the price of good one is certain in any price vector $p \in D_p^{\{1\}}$, but it need not be the same as the price of good one in some other price vector $q \in D_p^{\{1\}}$. However, with the price normalization, not only is the price of good one constant across states in a given state-contingent price vector, it has the same value in every state-contingent price vector.

6 Discussion

When the Bernoulli indirect utility function v is differentiable, we can measure the consumer's risk aversion with respect to income with the Arrow (1965)–Pratt (1964) coefficient of relative risk aversion:

$$\rho_y(p^0, y^0) = -\frac{v_{yy}(p^0, y^0)}{v_y(p^0, y^0)} y^0 \tag{47}$$

for all $(p^0, y^0) \in \mathbb{R}^N_{++} \times \mathbb{R}_+$. By Theorem 1, this coefficient must be identically zero. In other words, the consumer must be risk neutral towards income uncertainty.

Analogous to the Arrow–Pratt coefficient of relative risk aversion ρ_y for income, Turnovsky et al. (1980) have defined a coefficient of relative risk aversion ρ_{p_i} for the price of good i by taking derivatives with respect to p_i instead of with respect to y in (47) and then multiplying the resulting fraction by p_i^0 instead of by y^0. If it is assumed that the consumer's indirect utility function has the expected utility form given in (3), then the consumer's attitudes towards income and price uncertainty can be measured using the coefficients ρ_y and ρ_{p_i}, $i \in \mathcal{N}$. These measures are only invariant to increasing affine transforms of the Bernoulli indirect utility function v, which are also the transforms that do not affect the consumer's ex ante preferences over state-contingent prices and incomes. However, when computing the compensating variation in each state, only the ordinal properties of v are used. As a consequence, if v' is any increasing transform of v, then the compensating variation associated with a project in any state is the same with v' as it is with v, even if the risk attitudes associated with v' differ from those associated with v. Therefore, when the surplus evaluation function in (19) is used to determine whether a project is worthwhile or not, it makes the same recommendations for a consumer whose preferences are characterized by the function v as it does for a consumer whose preferences are characterized by v'. For this reason, as we have seen, restrictions must be placed on v in order for this cost-benefit test to be consistent.[11]

The restrictions on the Bernoulli indirect utility function v that we have identified for consistency imply that the consumer is risk neutral towards income and that the marginal utility of income does not depend on any price that can vary across states. If there is no price uncertainty and a project only changes incomes, then our cost-benefit test declares a project to be worthwhile if it increases the expected value of income. For a consumer who is risk neutral towards income, this is all that he cares about. However, if the consumer is not risk neutral, then he cares about the distribution of incomes, not just its expected value, and consistency would be lost. If a project also changes prices, by using the compensating variation to measure the surplus in each state, price changes are converted into an equivalent income change using the post-project prices. In order for expected compensating variation to provide a consistent cost-benefit test when some of the prices are stochastic, the marginal utility of income must be constant across states.[12] Because any distribution

[11] See Helms (1985, p. 609) for similar observations about the use of expected compensating variation as a test for whether stabilizing a single stochastic price is beneficial for an individual.

[12] In (44), $\alpha_K(\hat{p}_K^0)$ is the marginal utility of income at the post-project prices. If this value depends on any price that is not certain, then it could not be factored out in going from (44) to (45).

of incomes across states is possible, this requires that the marginal utility of income be independent of any price that can be state dependent.

In a model with a continuum of states, Helms (1984) has investigated when expected compensating variation is a consistent measure of individual welfare change when the only source of uncertainty is in the price of one good and there are no restrictions on the stochastic variability that this price might exhibit. Helms has shown that risk neutrality towards income and independence of the marginal utility of income with respect to this price are necessary and sufficient for consistency provided that the demand for this good is positive. That is, the Bernoulli indirect utility function must satisfy (33) with α independent of the price that is stochastic.[13]

Our theorems are closely related to results about the consistency of cost-benefit tests based on compensating or equivalent variations established by Blackorby and Donaldson (1985, 1986) and Blackorby, Donaldson, and Moloney (1984) in a variety of contexts.

Blackorby and Donaldson (1985) have shown (i) that no continuous, increasing surplus evaluation function defined on individual compensating or equivalent variations can be consistent with an indirect Bergson–Samuelson social welfare function when all prices and incomes can be person specific and (ii) that when everyone faces the same prices, consistency requires individual preferences to be quasi-homothetic with everyone having the same price-dependent weight on income in their indirect utility functions.[14] The latter condition is the necessary and sufficient condition identified by Gorman (1953) for the existence of community indifference curves.[15]

In Blackorby et al. (1984), a single consumer, whose utility is a continuous, increasing function of the instantaneous utilities obtained from his consumption in each of a finite number of periods, chooses these consumptions to maximize lifetime utility, given the prices of the goods and his wealth in a perfect capital market. They have shown (i) that no discounted sum of the compensating or equivalent variations in each period can serve as a consistent measure of welfare change for such a consumer if prices are free to vary across periods and (ii) that the instantaneous preferences must be quasi-homothetic with a common price-dependent weight on income in the corresponding indirect utility functions if all prices are constant across periods.[16]

[13] Neither Helms (1984) nor the other articles considered in the rest of this section take account of the restrictions required to ensure that demands are nonnegative for all admissible prices and incomes. For this reason, they only show that preferences must be quasi-homothetic, rather than being fully homothetic.

[14] Related results may be found in Hammond (1977, 1980) and Roberts (1980).

[15] Blackorby and Donaldson (1999) have established similar results about the consistency of the sum of individual Marshallian consumers' surpluses with an indirect Bergson–Samuelson indirect social welfare function. For discussions of the use of expected Marshallian consumer's surplus as a measure of individual welfare change, see Rogerson (1980), Stennek (1999), and Turnovsky et al. (1980).

[16] The analogue of a perfect capital market in our model is a perfect insurance market that permits an individual to transfer wealth across states.

In Blackorby and Donaldson (1986), there is a single period and each individual consumes an amount of a single commodity should he live, which occurs with positive probability. For the case in which each person's preferences satisfy the expected utility hypothesis and everyone has some level of consumption that makes life just worth living, they have shown (i) that the sum of the individual compensating or equivalent variations is not consistent with the ranking of alternative distributions of survival probabilities and of consumptions obtained with any continuous, increasing Bergson–Samuelson social welfare function if both the probabilities of survival and the consumptions can be person specific and (ii) that when everyone has the same survival probability, then consistency requires each individual's preferences to be representable by a utility function that is affine in consumption with a common weight on consumption that depends only on the survival probability. As Blackorby and Donaldson (1986, Sect. III) have noted, the probabilities in this model correspond to prices in the riskless multi-good model.

There is clearly a close family resemblance between these results and those obtained here. This is not surprising. Although our model and those described above differ in some important respects, the overall measure of individual or social welfare in each case is a continuous, increasing function of the utility functions that are used to compute the individual or state-contingent or period-contingent compensating or equivalent varitions. Furthermore, the surplus evaluation function is, in each case, a continuous, increasing function of these surpluses. It is these common structural features of these models that accounts for the similarity of the results about the consistency of welfare evaluations based on Hicksian measures of consumer's surplus that are obtained with them.

7 Concluding Remarks

The restrictions on preferences that Helms (1984, 1985) has shown are required for expected compensated variation to be a consistent measure of individual welfare change are much less restrictive when a single stochastic price is stabilized at its mean value compared with the case in which all distributions can be stochastic. However, they are still quite stringent and are unlikely to be satisfied in practice. For the prices that are allowed to vary across states and for income, we have placed no restrictions on the pre- and post-project distributions. We could instead restrict our domains by, for example, considering projects that stabilize income or some of the prices. On such domains, the conditions required for consistency would be weaker than those obtained here. However, Helms's theorems suggest that they will nevertheless be quite restrictive, so considering more specialized domains does not appear to be a promising direction in which to seek more positive results.

In view of the rather stringent conditions required for a surplus evaluation function based on the ex post compensating variations to be a consistent measure of individual welfare change, it is natural to ask if there is any measure of consumer's surplus that applies more generally when prices or incomes are uncertain. An affirmative answer is provided by the ex ante compensating and equivalent variations introduced by Schmalensee (1972).[17]

The *ex ante compensating variation* s_c for a project that changes the state-contingent prices and income from $(\bar{p},\bar{y})$ to $(\hat{p},\hat{y})$ is defined implicitly by

$$\sum_m \pi_m v(\hat{p}^m,\hat{y}^m - s_c) = V(\bar{p},\bar{y}). \tag{48}$$

That is, s_c is the amount by which an individual's income can be reduced in each state in order for the post-project situation to give him the same ex ante expected utility as is achieved before the project is implemented. Because v is increasing in income, s_c is positive if and only if the project makes the consumer better off ex ante. Thus, s_c can serve as an exact measure of welfare change for any individual whose preferences satisfy the expected utility hypothesis. Similarly, the *ex ante equivalent variation* s_e is the amount of income that needs to be provided to an individual in each state in the pre-project situation in order to give him the same ex ante expected utility as is achieved after the project is implemented. It too is an exact measure of individual welfare change.

Schmalensee (1972) did not advocate the use of these ex ante measures because he thought that they are non-operational. However, assuming that the appropriate coefficients of risk aversion can be determined from analyzing behavior under uncertainty, Anderson (1979) has argued that these measures are operational, and so has endorsed their use, as has Helms (1985).[18] Given that information about risk attitudes is needed in order to determine if expected compensating (or equivalent) variation is a consistent measure of individual welfare change, it therefore seems that there is little reason to use the expected value of some consumer's surplus measure to evaluate projects that involve price and income uncertainty instead of the ex ante compensating or equivalent variation.

Acknowledgments This article was presented at the International Conference on Rational Choice, Individual Rights and Non-Welfaristic Normative Economics in Honor of Kotaro Suzumura at Hitotsubashi University, the Workshop on Advances in Collective Choice at the University of Maastricht, the Workshop on Economic Decisions at the Public University of the Navarre, the Workshop on Advances in Microeconomics and Social Choice Theory at the University of Verona, the Eighth International Meeting of the Society for Social Choice and Welfare in Istanbul, and at a seminar at GREQAM. We are grateful for the comments that we have received on these occasions. We are particularly grateful to Peter Hammond for his comments at the Hitotsubashi conference.

[17] Schmalensee refers to these measures as compensating and equivalent option prices. An alternative proposal for measuring individual welfare change under uncertainty may be found in Boadway and Bruce (1984, Chap. 7).

[18] The relationship between these ex ante measures and the expected value of a Hicksian or Marshallian measure of ex post consumer's surplus has been explored in Choi and Johnson (1987) and Stennek (1999).

References

Aczél, J. (1969). *Lectures on functional equations and their applications*. New York: Academic Press

Anderson, J. E. (1979). On the measurement of welfare cost under uncertainty. *Southern Economic Journal*, *45*, 1160–1171

Anderson, J. E., and Riley, J. G. (1976). International trade with fluctuating prices. *International Economic Review*, *17*, 76–97

Arrow, K. J. (1953). Le rôle des valeurs boursières pour la répartition la meillure des risques. In *Econométrie* (pp. 41–47). Paris: Centre National de la Recherche Scientifique. (Translated as Arrow (1964))

Arrow, K. J. (1965). *Aspects of the theory of risk-bearing*. Helsinki: Yrjö Jahnssonin säätiö

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies*, *31*, 91–96

Blackorby, C., Boyce, R., and Russell, R. R. (1978). Estimation of demand systems generated by the Gorman polar form: A generalization of the S-branch utility tree. *Econometrica*, *46*, 345–363

Blackorby, C., Davidson, R., and Donaldson, D. (1977). A homiletic exposition of the expected utility hypothesis. *Economica*, *44*, 351–358

Blackorby, C., and Donaldson, D. (1985). Consumers' surpluses and consistent cost-benefit tests. *Social Choice and Welfare*, *1*, 251–262

Blackorby, C., and Donaldson, D. (1986). Can risk-benefit analysis provide consistent policy evaluations of projects involving loss of life? *Economic Journal*, *96*, 758–773

Blackorby, C., and Donaldson, D. (1999). Market demand curves and Dupuit–Marshall consumers' surpluses: A general equilibrium analysis. *Mathematical Social Sciences*, *37*, 139–163

Blackorby, C., Donaldson, D., and Moloney, D. (1984). Consumer's surplus and welfare change in a simple dynamic model. *Review of Economic Studies*, *51*, 171–176

Boadway, R. W., and Bruce, N. (1984). *Welfare economics*. Oxford: Basil Blackwell

Chipman, J. S., and Moore, J. C. (1976). The scope of consumer's surplus arguments. In A. M. Tang, F. M. Westfield, and J. S. Worley (Eds.), *Evolution, welfare, and time in economics: Essays in honor of Nicholas Georgescu-Roegen* (pp. 69–123). Lexington, MA: Lexington Books

Chipman, J. S., and Moore, J. C. (1980). Compensating variation, consumer's surplus, and welfare. *American Economic Review*, *70*, 933–949

Choi, E. K., and Johnson, S. R. (1987). Consumer's surplus and price uncertainty. *International Economic Review*, *28*, 407–411

Diewert, W. E. (1993). Symmetric means and choice under uncertainty. In W. E. Diewert and A. O. Nakamura (Eds.), *Essays in index number theory, volume 1* (pp. 355–433). Amsterdam: North-Holland

Eichhorn, W. (1978). *Functional equations in economics*. Reading, MA: Addison-Wesley

Gorman, W. M. (1953). Community preference fields. *Econometrica*, *21*, 63–80

Gorman, W. M. (1961). On a class of preference fields. *Metroeconomica*, *13*, 53–56

Hammond, P. J. (1977). Dual interpersonal comparisons of utility and the welfare economics of income distribution. *Journal of Public Economics*, *7*, 51–71

Hammond, P. J. (1980). Dual interpersonal comparisons of utility and the welfare economics of income distribution: A corrigendum. *Journal of Public Economics*, *14*, 105–106

Helms, L. J. (1984). Comparing stochastic price regimes: The limitations of expected surplus measures. *Economics Letters*, *14*, 173–178

Helms, L. J. (1985). Expected consumer's surplus and the welfare effects of price stabilization. *International Economic Review*, *26*, 603–617

Newbery, D. M., and Stiglitz, J. E. (1981). *The theory of commodity price stabilization: A study in the economics of risk*. Oxford: Clarendon Press

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*, 122–136

Roberts, K. (1980). Price-independent welfare prescriptions. *Journal of Public Economics*, *13*, 277–297

Rogerson, W. P. (1980). Aggregate expected consumer surplus as a welfare index with an application to price stabilization. *Econometrica*, *48*, 423–436

Schmalensee, R. (1972). Option demand and consumer's surplus: Valuing price changes under uncertainty. *American Economic Review*, *62*, 813–824

Stennek, J. (1999). The expected consumer's surplus as a welfare measure. *Journal of Public Economics*, *73*, 265–288

Turnovsky, S. J., Shalit, H., and Schmitz, A. (1980). Consumer's surplus, price instability, and consumer welfare. *Econometrica*, *48*, 135–152

Varian, H. R. (1980). Redistributive taxation as social insurance. *Journal of Public Economics*, *14*, 49–68

von Neumann, J., and Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press

Waugh, F. V. (1944). Does the consumer benefit from price instability? *Quarterly Journal of Economics*, *58*, 602–614

Beyond Normal Form Invariance: First Mover Advantage in Two-Stage Games with or without Predictable Cheap Talk

Peter J. Hammond

1 Motivation and Introduction

1.1 Von Neumann's Standard Paradigm

Following Zermelo's (1912) pioneering analysis of chess and similar games, von Neumann (1928) devised a *standard paradigm*, according to which multiperson decision problems in modern economic analysis and other social science are nearly always modeled as noncooperative games in strategic form. This paradigm relies on two key assumptions, of which the first can be stated as follows:

Assumption 1. *A multiperson decision problem is fully described by a game in extensive form, whose structure is commonly known to all players in the game.*

Von Neumann's (1928) own extensive form description was later incorporated in *The Theory of Games and Economic Behavior*. Kuhn (1953) pointed out the implicit assumption that the order of different players' information sets was commonly known to all players at all stages of the game, and extended the von Neumann description to relax this assumption. Much more generally, we can now envisage an extensive form of game as a stochastic process subject to the control of different players, with each player's information at each time described by a filtration. One key assumption, however, is that this stochastic process fits within Kolmogorov's (1933) framework of one overall probability space that includes everything random. As argued in Hammond (2007), this fails to allow for the possibility of having events that no player can foresee, and which may indeed even be impossible for any ideal observer to foresee.

P.J. Hammond
Department of Economics, University of Warwick, Coventry, CV4 7AL, UK
e-mail: p.j.hammond@warwick.ac.uk

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*,
Studies in Choice and Welfare.

1.2 Normal Form Invariance

The second assumption, which seems to have originated in von Neumann (1928), can be stated as follows:

Assumption 2. *It loses no generality to reduce the game in extensive form to the corresponding game in strategic or normal form, where each player makes a single strategic plan that covers all eventualities in the extensive form.*

It is perhaps worth going back all the way to von Neumann's original article, as adapted in von Neumann and Morgenstern (1943, 1953), to see how he justified normalizing the extensive form. First, normal form strategies are described on p. 79:

> Imagine now that each player ..., instead of making each decision as the necessity for it arises, makes up his mind in advance for all possible contingencies; i.e., that the player ... begins to play with a complete plan: a plan which specifies what choices he will make in every possible situation, for every possible actual information which he may possess at that moment in conformity with the pattern of information which the rules of the game provide for him for that case. We call such a plan a *strategy*.

Then pages 79–84 proceed to simplify the description of an extensive form game to arrive at the normal form of the game in which each player makes just one move, and all moves are chosen simultaneously. In fact (p. 84):

> Each player must make his choice [of strategy] in absolute ignorance of the choices of the others. After all choices have been made, they are submitted to an umpire who determines ... the outcome of the play for [each] player.
>
> Observe that in this scheme no space is left for any kind of further 'strategy.' Each player has one move, and one move only; and he must make it in absolute ignorance of everything else.

Normalizing an extensive form game in this way is an extremely powerful device. And if the players of a game really do simultaneously submit their choices of a strategy to an umpire, who then sees that the players never deviate from their announced choices, then von Neumann and Morgenstern's claim on p. 85 seems entirely justified:

> ... we obtained an all-inclusive formal characterization of the general game of *n* persons .... We followed up by developing an exact concept of strategy which permitted us to replace the rather complicated general scheme of a game by a much more simple special one, which was nevertheless shown to be fully equivalent to the former .... In the discussion which follows it will sometimes be more convenient to use one form, sometimes the other. It is therefore desirable to give them specific technical names. We will accordingly call them the *extensive* and the *normalized* form of the game, respectively.
>
> Since these two forms are strictly equivalent, it is entirely within our province to use in each particular case whichever is technically more convenient at that moment. We propose, indeed, to make full use of this possibility, and must therefore re-emphasize that this does not in the least affect the absolute general validity of all our considerations.

It is this simplification that gives such power to familiar "normal form" concepts like Nash equilibrium, as well as to less familiar ones like trembling-hand perfect equilibrium (Selten, 1975), proper equilibrium (Myerson, 1978), correlated equilibrium (Aumann, 1987), rationalizable strategies (Berhmeim, 1984 and Pearce, 1984).

Also, Mailath, Samuelson, and Swinkels (1993) show how even ostensibly extensive form ideas such as Selten's (1965) concept of subgame perfect equilibrium, or Kreps and Wilson's (1982) concept of sequential equilibrium, have their (reduced) normal form counterparts.

Game theorists do relax normal form invariance somewhat by using extensive form solution concepts. For example, requiring players to respond credibly when other players deviate from expected behavior was the original motivation for subgame perfection. See also Amershi, Sadanand, and Sadanand (1985, 1989a, 1989b, 1992); Hammond (1993); Sadanand and Sadanand (1995); Battigalli (1997); Battigalli and Siniscalchi (1999, 2002); and Asheim and Dufwenberg (2003), among other works that cast doubt on the normal form invariance hypothesis.

1.3 Outline of Chapter

The purpose of this chapter is to present a theoretical argument supporting the view that normal form invariance may be unduly restrictive. To do so, Section 2 considers a simple "Battle of the Sexes" game, where experimental evidence suggests that the first move does confer an advantage. It sets out the claim that this may be due to what would happen in the unique credible equilibrium of an associated game where cheap talk is possible after the first move, but before the second.

Section 3 begins to analyze a general two-stage game where one player moves first, and the only other player moves second, but without knowing the first player's move. It then allows simultaneous cheap talk by both players at an intermediate stage, between their two moves.

Because we are looking for an equilibrium that the players can infer, we require player 1's cheap talk to be "predictable" in the sense that it results from a pure strategy, which is independent of her (hidden) action. Hence, we consider a game where player 1 combines a mixed act with a pure message strategy. Afterwards, player 2 first sends a message without knowing what 1 has done, then forms his conditional beliefs, given 1's message and chooses an optimal mixed act accordingly.

Not surprisingly, any perfect Bayesian equilibrium (PBE) in the game with predictable cheap talk must induce a Nash equilibrium in the corresponding game without cheap talk. On the other hand, any Nash equilibrium without cheap talk can be extended into a PBE by making the second player "inattentive" to all cheap talk when forming his beliefs and choosing his strategy. Thus, cheap talk alone fails to refine the set of PBEs.

To facilitate such a refinement, Section 4 invokes a particular version of the revelation principle in the form due to Myerson (1982), as amended by Kumar (1985). First, this will allow player 2's message to be ignored, since anything he says could affect only his own actions. Second, the revelation principle will allow general predictable cheap talk by player 1 to be replaced by "direct" cheap talk in the form of two suggestions for player 2, at his only information set: (i) the conditional probabilities that should be attached to player 1's earlier moves; (ii) player 2's choice of mixed act. Moreover, as argued in Section 4, we can limit attention to "straightforward" PBEs, where player 2 accepts both player 1's suggestions.

Section 5 finally introduces a credibility refinement. This requires a straightforward PBE to survive even when player 2 is "Nash attentive," that is, when he accepts any suggestion by player 1 for choosing a Nash equilibrium of the game without cheap talk. The resulting "credible" equilibrium with cheap talk leads to an optimal Nash equilibrium for player 1 in the original game without cheap talk. When this optimal Nash equilibrium is unique, "sophistication" allows this cheap talk to remain implicit, and so unnecessary. While these results may be hardly surprising, they do show how tacit communication can explain first-mover advantage in games like Battle of the Sexes.

Section 6 considers "virtual observability." This occurs when, as in Battle of the Sexes, sophistication effectively converts the game into one of the perfect information, with the second player knowing the first move. Three examples show that virtual observability is rather special.

The concluding Section 7 discusses possible extensions and suggestions for future work that relaxes normal form invariance in other ways.

Except where it is standard, most notation will be explained wherever it is first used. Given any finite set F, however, let $\Delta(F)$ denote the set of probability distributions over F. Also, if F' is a proper subset of F, let $\Delta(F') \subset \Delta(F)$ denote those distributions that attach probability one to F'. Finally, if X and Y are arbitrary sets, let $X^Y := \prod_{y \in Y} X_y$ denote the set of all mappings $y \mapsto x_y$ from Y to X.

2 Battle of the Sexes

2.1 Two Different Extensive Forms

The two games in Figs. 1 and 2 are different extensive form versions of the familiar "Battle of the Sexes" game, whose normal form is given in Fig. 3. As is well known, there are two Nash equilibria in pure strategies, namely (B,b) and (S,s). There is also one mixed strategy Nash equilibrium where player 1 chooses B with probability $\frac{2}{3}$, and player 2 chooses b with probability $\frac{1}{3}$.

Nevertheless, experiments strongly suggest that the player who moves first enjoys an advantage, in so far as (B,b) is played more often than (S,s) in Fig. 1, but less often in Fig. 2.[1] These results have usually been ascribed to "positional order"

[1] A "preliminary" experiment along these lines is described by Amershi, Sadanand, and Sadanand (1989a). Kreps (1990, p. 100) writes about "casual experiences playing this game with students." Later formal experiments yielding similar results were reported in Cooper, Dejong, Forsythe, and Ross (1989, 1993). See also Schotter, Weigelt, and Wilson (1994); Rapoport (1997); Güth, Huck, and Rapoport (1998); Muller and Sadanand (2003); and Weber, Camerer, and Knez (2004). The work by Güth, Huck, and Rapoport (1998) even includes an experiment in which a form of cheap talk is explicitly allowed. The experimental design, however, includes the wording

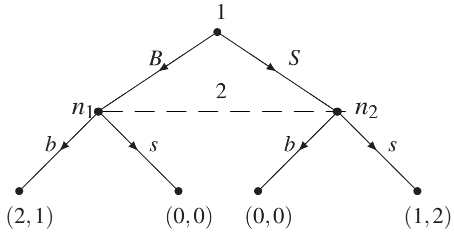


Fig. 1 Battle of the sexes where player 1 moves first

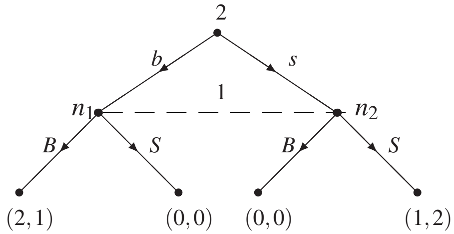


Fig. 2 Battle of the sexes where player 2 moves first

Fig. 3 Battle of the sexes in normal form

	b	s
B	2, 1	0, 0
S	0, 0	1, 2

or "presentation" effects that are seen as psychological or behavioral rather than fully rational responses to a change in the extensive form of the game.

2.2 *Direct Cheap Talk in Battle of the Sexes*

Consider the extensive form of Fig. 1, where player 1 moves first, and this is common knowledge. Suppose that, during an intermediate stage that succeeds player 1's move but precedes player 2's, the two players are allowed to communicate and indulge in unrestricted and mutually comprehensible "cheap talk."

As argued in Sect. 4, however, an extended version of the revelation principle implies that, in perfect Bayesian equilibrium (PBE), only player 1's cheap talk is relevant; it is already too late for player 2 to influence any action choice except his own. Moreover, we need only consider *direct* cheap talk where player 1's message m is a pair suggesting conditional probabilities $\rho(\cdot) \in \Delta(\{B,S\})$ and a mixed strategy $\sigma(\cdot) \in \Delta(\{b,s\})$ for player 2 at his only information set. Finally, the same principle allows us to limit attention to a "straightforward" PBE, where player 2 accepts 1's suggestions.

"B learns about A's decision" in the instructions. This may bias the results by offering the subjects too little encouragement to recognize the possibility of sending or receiving a deceptive message.

Now, any straightforward PBE would seem to involve just one of three possible direct messages that player 1 might send, corresponding to the three different Nash equilibria of the normal form:

1. Corresponding to the equilibrium (B,b), a message with $\rho(B) = \sigma(b) = 1$ that yields the two players' expected payoffs of $(2,1)$
2. Corresponding to the equilibrium (S,s), a message with $\rho(S) = \sigma(s) = 1$ that yields the two players' expected payoffs of $(1,2)$
3. Corresponding to the mixed strategy equilibrium, a message with

$$\rho(B) = \sigma(s) = \tfrac{2}{3} \quad \text{and} \quad \rho(S) = \sigma(b) = \tfrac{1}{3},$$

which yields the two players' expected payoffs of $(\frac{2}{3}, \frac{2}{3})$.

2.3 *One Credible Equilibrium with Cheap Talk*

In this Battle of the Sexes game with cheap talk, suppose all three "straightforward" messages could be regarded as credible. Then player 1 would expect player 2 to respond appropriately to whichever straightforward message she sends. So she would definitely choose the first of the three. But then, if player 2 hears any direct message except "I have played B and recommend that you play b", he should wonder whether player 1 has really not played B, or whether player 1 has somehow misspoken after playing B. Thus, player 2's best response to any other direct message actually becomes unclear. In the case of Battle of the Sexes, however, all that matters is that player 2 does choose b when player 1 suggests he should. This leaves us with just one possible outcome of any *credible* perfect Bayesian equilibrium (PBE).

Finally, if predictable direct cheap talk would produce a unique credible equilibrium message, we assume that both players are sufficiently "sophisticated" to reason what it will be. But this removes any need for cheap talk. Player 2 can work out the unique equilibrium message that he would receive in any credible PBE of the game with predictable direct cheap talk, and player 1 should know this also. By tacitly inferring what would happen if cheap talk were actually permitted, they reach the same unique outcome as in any credible PBE with predictable cheap talk.

3 General Two-Stage Games

3.1 *The Basic Extensive Game*

Instead of the specific Battle of the Sexes game discussed in Sect. 2, consider a general two-stage game Γ_0 with two players 1 and 2, for whom all the following facts are common knowledge. Player 1 begins the game by choosing an action a_1 from the finite set A_1. Then player 2 at his only information set, without seeing a_1, finishes the game by choosing an action a_2 from the finite set A_2. Each player i's payoff is denoted by $u_i(a_1, a_2)$ (for $i = 1, 2$). Allowing for mixed strategies $\alpha_i \in \Delta(A_i)$, the normal form of Γ_0 can be written as

$$G_0 = \langle \{1,2\}, \Delta(A_1), \Delta(A_2), v_1, v_2 \rangle, \tag{1}$$

with (expected) payoffs $v_i : \Delta(A_1) \times \Delta(A_2) \to \mathbb{R}$ for $i = 1, 2$ given by

$$v_i(\alpha_1, \alpha_2) := \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \alpha_1(a_1)\alpha_2(a_2)u_i(a_1, a_2). \tag{2}$$

Next, given their respective beliefs $\pi_1 \in \Delta(A_2)$ and $\pi_2 \in \Delta(A_1)$, define the two players' mixed strategy best response sets

$$B_1(\pi_1) := \arg\max_{\alpha_1 \in \Delta(A_1)} v_1(\alpha_1, \pi_1) \tag{3}$$

$$\text{and} \quad B_2(\pi_2) := \arg\max_{\alpha_2 \in \Delta(A_2)} v_2(\pi_2, \alpha_2). \tag{4}$$

Finally, we denote the set of mixed strategy Nash equilibra of G_0 by

$$E_0 := \{ (\alpha_1, \alpha_2) \in \Delta(A_1) \times \Delta(A_2) \mid \alpha_1 \in B_1(\alpha_2),\ \alpha_2 \in B_2(\alpha_1) \}. \tag{5}$$

These are also the Nash (and perfect Bayesian) equilibra of Γ_0.

3.2 Predictable Cheap Talk

Cheap talk is introduced by allowing the two players to choose simultaneous message strategies $m_i \in M_i$ (for $i = 1, 2$) after player 1 has chosen a_1, but before player 2 chooses a_2. Often it will be convenient to let $m \in M := M_1 \times M_2$ denote the typical message pair (m_1, m_2). Of course, the main claim of this chapter is precisely that it really *is* restrictive to reduce complex interactions to single strategy choices by each player.[2] Nevertheless, such restrictions seem not to detract from the force of the main argument.

Also, we look eventually for a predictable unique equilibrium of the game with cheap talk. Note, however, that no mixed message strategies could work this way; player 2 could not predict what messages result from such randomization. Nor can player 1 make her message depend on the action that results from a mixed action

[2] Moreover, this rules out the kind of "long" cheap talk considered by Aumann and Hart (2003). Their model, however, involves messages that are sent by choosing one among only a finite set of "keystrokes." Also, the only example they provide of an equilibrium involving long cheap talk is presented in their Section 2.8. In a particular signaling game, it amounts to finding a mixed message strategy with infinite support. The formulation used here would allow any such message to be sent in only one stage.

strategy. So we consider only "predictable" cheap talk that results in one fixed message strategy for each player, independent of player 1's earlier action.

3.3 An Extensive Form Game

An obvious two-person extensive game of perfect recall with predictable cheap talk proceeds in three successive stages as follows:

First action stage: Player 1 has one initial information set where she chooses a mixed action strategy $\alpha_1 \in \Delta(A_1)$.

Intermediate message stage: Both players simultaneously choose predictable messages $m_1 \in M_1$ and $m_2 \in M_2$. Though player 1 knows α_1 and even a_1, predictability rules out using this information. Hence, both players communicate as though they have a single information set at this stage.

Second action stage: Player 2 has an information set $H_2(m)$ for each possible message pair $m \in M$. This enables him to choose a function $\alpha_2(\cdot|\cdot) \in [\Delta(A_2)]^M$ mapping each $m \in M$ to a mixed action strategy $\alpha_2(\cdot|m) \in \Delta(A_2)$.

Let Γ denote this extensive game. Its normal form can be written as

$$G = \langle \{1,2\}, S_1, S_2, w_1, w_2 \rangle, \tag{6}$$

where the two players' permitted (mixed) strategy sets have typical members denoted by

$$(\alpha_1, m_1) \in S_1 := \Delta(A_1) \times M_1 \tag{7}$$

$$\text{and} \quad (m_2, \alpha_2(\cdot|\cdot)) \in S_2 := M_2 \times [\Delta(A_2)]^M. \tag{8}$$

Also, definition (2) allows the two players' expected final payoffs $w_i : S_1 \times S_2 \to \mathbb{R}$ to be written as

$$w_i(\alpha_1, m_1, m_2, \alpha_2(\cdot|\cdot,\cdot)) := v_i(\alpha_1, \alpha_2(\cdot|m_1, m_2)). \tag{9}$$

3.4 Characterizing Perfect Bayesian Equilibrium

In a general extensive form game, a *perfect Bayesian equilibria* (PBE) is a *strategy–belief profile* which, for each player i and for each information set H where i has the move, combines: (i) a behavioral strategy specifying what (mixed) move i makes at H; (ii) a belief system specifying what subjective probabilities player i attaches to the different nodes of H. Moreover, this combination must satisfy the following two requirements:

Consistent beliefs: Player i's beliefs at H are derived by Bayesian updating, provided the conditional probabilities are well defined, given equilibrium moves at previous information sets;

Sequential rationality: Player i's move at H should maximize i's conditional expected payoff, given the players' behavior strategies at all other information sets, and given player i's beliefs at H.

For the game Γ, accordingly, any strategy–belief profile involves player 2's conditional beliefs at each information set $H_2(m)$, after observing the message pair $m = (m_1, m_2) \in M$. We regard any such *belief system* as a mapping $m \mapsto \pi(\cdot|m)$ from M to $\Delta(A_1)$, denoted by

$$\pi(\cdot|\cdot) \in [\Delta(A_1)]^M. \tag{10}$$

We now give conditions for a particular strategy–belief profile

$$(\alpha_1^*, m^*, \alpha_2^*(\cdot|\cdot), \pi^*(\cdot|\cdot)) \in \Delta(A_1) \times M \times [\Delta(A_2)]^M \times [\Delta(A_1)]^M \tag{11}$$

in Γ to be a PBE.

At each last information set $H_2(m)$ of Γ, following the observed message pair $m \in M$, player 2's equilibrium belief system $\pi^*(\cdot|\cdot)$ determines his best response set $B_2(\pi^*(\cdot|m))$. Sequential rationality therefore requires player 2's behavior strategy at $H_2(m)$ to satisfy

$$\alpha_2^*(\cdot|m) \in B_2(\pi^*(\cdot|m)) \quad \text{for each } m \in M. \tag{12}$$

Earlier, anticipating player 2's equilibrium message m_2^* and sequentially rational response to each pair (m_1, m_2^*), player 1 chooses the pair

$$(\alpha_1^*, m_1^*) \in \underset{(\alpha_1, m_1) \in \Delta(A_1) \times M_1}{\arg\max} v_1(\alpha_1, \alpha_2^*(\cdot|m_1, m_2^*)). \tag{13}$$

This implies in particular that in the first action stage 1, anticipating both the equilibrium message pair $m^* \in M$ and player 2's induced response $\alpha_2^*(\cdot|m^*)$, player 1 chooses a mixed action strategy satisfying

$$\alpha_1^* \in B_1(\alpha_2^*(\cdot|m^*)). \tag{14}$$

During the intermediate message stage, player 2 anticipates player 1's choice of (α_1^*, m_1^*) and his own sequentially rational response to each pair $m \in M$. Hence player 2's equilibrium message m_2^* satisfies

$$m_2^* \in \underset{m_2 \in M_2}{\arg\max}\, v_2(\alpha_1^*, \alpha_2^*(\cdot|m_1^*, m_2)). \tag{15}$$

Finally, consistency of beliefs on the equilibrium path implies that

$$\pi^*(\cdot|m^*) = \alpha_1^*. \tag{16}$$

Then (12) implies that player 2 chooses a mixed strategy satisfying

$$\alpha_2^*(\cdot|m^*) \in B_2(\alpha_1^*). \tag{17}$$

3.5 Perfect Bayesian and Nash Equilibria

The following simple result establishes that, because any PBE of Γ induces Nash equilibrium strategies along an equilibrium path, it induces Nash equilibrium action strategies in the game G_0 without cheap talk.

Lemma 1. *Suppose the strategy–belief profile $(\alpha_1^*, m^*, \alpha_2^*(\cdot|\cdot), \pi^*(\cdot|\cdot))$ is a PBE in the game Γ with predictable cheap talk. Then the mixed action strategy profile $(\alpha_1^*, \alpha_2^*(\cdot|m^*))$ in $\Delta(A_1) \times \Delta(A_2)$ induced along the equilibrium path must be a Nash equilibrium in the game Γ_0 without cheap talk.*

Proof. Given the equilibrium message pair m^*, conditions (14) and (17) imply that the induced mixed strategies α_1^* and $\alpha_2^*(\cdot|m^*)$ are mutual best responses. So the strategy pair belongs to the set E_0 of Nash equilibria of the game Γ_0 without cheap talk, as defined in (5). □

The next result shows that cheap talk alone excludes none of the Nash equilibria in the game Γ_0. In particular, all three Nash equilibria in the Battle of the Sexes example of Sect. 2 can be extended to PBEs with appropriate cheap talk.

Definition 1. In the game Γ with predictable cheap talk, player 2's strategy–belief system $(\alpha_2(\cdot|\cdot), \pi(\cdot|\cdot)) \in [\Delta(A_2) \times \Delta(A_1)]^M$ is *inattentive* if both $\alpha_2(\cdot|m)$ and $\pi(\cdot|m)$ are constant, independent of m, for all message pairs $m \in M$. A PBE $(\alpha_1^*, m^*, \alpha_2^*(\cdot|\cdot), \pi^*(\cdot|\cdot))$ in Γ is *inattentive* if player 2's equilibrium strategy–belief system is inattentive.

Lemma 2. *Let $(\bar{\alpha}_1, \bar{\alpha}_2) \in E_0$ be any Nash equilibrium in the game Γ_0 without cheap talk. Let M be any message space for player 1. Then the corresponding game Γ with predictable cheap talk in M has an inattentive PBE, which induces $(\bar{\alpha}_1, \bar{\alpha}_2)$ along the equilibrium path.*

Proof. Consider the strategy–belief profile in Γ where

1. player 1 combines $\alpha_1^* = \bar{\alpha}_1$ with an arbitrary message $m_1^* \in M_1$
2. player 2 sends an arbitrary message $m_2^* \in M_2$
3. player 2's strategy–belief system is inattentive, with

$$\alpha_2^*(\cdot|m) = \bar{\alpha}_2 \quad \text{and} \quad \pi^*(\cdot|m) = \bar{\alpha}_1 \quad \text{for all } m \in M. \tag{18}$$

It is easy to see that $(\alpha_1^*, m^*, \alpha_2^*(\cdot|\cdot), \pi^*(\cdot|\cdot))$ must be a PBE. □

4 An Extended Revelation Principle

4.1 Direct Cheap Talk

The revelation principle will involve a new game $\hat{\Gamma}$, which is like Γ, except the following:

1. Player 2's message space M_2 becomes a singleton $\{\bar{m}_2\}$, so he can only send a constant message $\bar{m}_2$. This makes 2's message irrelevant, of course, so we ignore it from now on.
2. Player 1's general messages $m_1 \in M_1$ are replaced by *direct* messages

$$\hat{m} = (\rho, \sigma) \in \hat{M} := \Delta(A_1) \times \Delta(A_2). \tag{19}$$

Here, following Kumar's (1985) extension of the revelation principle, the first component $\rho \in \Delta(A_1)$ of each direct message that player 1 might send can be interpreted as beliefs about player 1's strategy that 1 suggests to 2. Following Myerson (1982), the second component $\sigma \in \Delta(A_2)$ can be interpreted as the mixed strategy that 1 suggests to 2.[3]

The typical strategy–belief profile in the game $\hat{\Gamma}$ with direct cheap talk will be denoted by

$$(\hat{\alpha}_1, \hat{m}, \hat{\alpha}_2(\cdot|\cdot), \hat{\pi}(\cdot|\cdot)) \in \Delta(A_1) \times \hat{M} \times [\Delta(A_2)]^{\hat{M}} \times [\Delta(A_1)]^{\hat{M}}. \tag{20}$$

4.2 Equivalent Straightforward Equilibria

Definition 2. In the game $\hat{\Gamma}$ with direct cheap talk, the strategy–belief profile $(\hat{\alpha}_1, \hat{m}, \hat{\alpha}_2(\cdot|\cdot), \hat{\pi}(\cdot|\cdot))$ with $\hat{m} = (\rho, \sigma)$ is *straightforward* if

$$\hat{\pi}(\cdot|\hat{m}) = \rho = \hat{\alpha}_1 \quad \text{and} \quad \hat{\alpha}_2(\cdot|\hat{m}) = \sigma. \tag{21}$$

A strategy–belief profile that is straightforward and also a PBE is a *straightforward PBE*.

That is, a strategy–belief profile is straightforward if player 1 suggests beliefs that match her mixed action and if player 2 accepts both suggestions that make up player 1's direct message.

The following result extends to our setting the versions of the revelation principle due to Myerson (1982) and Kumar (1985).

Theorem 1. *Let $(\alpha_1^*, m^*, \alpha_2^*(\cdot|\cdot), \pi^*(\cdot|\cdot))$ be any PBE strategy–belief profile in the game Γ with general predictable cheap talk. Then in the associated game $\hat{\Gamma}$ with direct cheap talk there is an* equivalent *PBE*

[3] Following Forges (1986), many later writers describe direct messages as "canonical."

$$(\hat{\alpha}_1^*, \hat{m}^*, \hat{\alpha}_2^*(\cdot|\cdot), \hat{\pi}^*(\cdot|\cdot)) \tag{22}$$

that is inattentive, straightforward, and generates the same equilibrium action strategy pair

$$(\hat{\alpha}_1^*, \hat{\alpha}_2^*(\cdot|\hat{m}^*)) = (\alpha_1^*, \alpha_2^*(\cdot|m^*)). \tag{23}$$

Proof. By Lemma 1, the mixed action strategy pair $(\alpha_1^*, \alpha_2^*(\cdot|m^*))$ generated by the PBE of Γ must be a Nash equilibrium of the game Γ_0 without cheap talk. To construct the equivalent PBE strategy–belief profile (22), first choose $\hat{\alpha}_1^* = \alpha_1^*$. Next, define the equivalent direct message $\hat{m}^* \in \hat{M}$ in the game $\hat{\Gamma}$ as the Nash equilibrium pair $(\alpha_1^*, \alpha_2^*(\cdot|m^*))$ itself. Finally, define an inattentive strategy–belief system for player 2 by choosing $\hat{\pi}^*(\cdot|\hat{m}) := \alpha_1^*$ and $\hat{\alpha}_2^*(\cdot|\hat{m}) := \alpha_2^*(\cdot|m^*)$ for each direct message $\hat{m} \in \hat{M} = \Delta(A_1) \times \Delta(A_2)$.

Evidently the constructed strategy–belief profile (22) is both inattentive and straightforward. As in Lemma 2, it is also a PBE of $\hat{\Gamma}$. □

The extended revelation principle is especially useful in allowing any PBE in the game Γ with predictable cheap talk to be converted to an inattentive straightforward PBE in the associated game $\hat{\Gamma}$ with direct cheap talk. Nevertheless, Lemma 2 applies even in $\hat{\Gamma}$. For this reason, an extra consideration is needed to refine the set of Nash equilibria.

5 Credible Equilibria with Direct Cheap Talk

5.1 Nash Attentiveness

The following definition requires player 2 to accept player 1's direct message in $\hat{\Gamma}$ whenever it suggests a specific Nash equilibrium of the game Γ_0 without cheap talk.

Definition 3. In the game $\hat{\Gamma}$ with direct cheap talk, player 2's strategy–belief system $(\hat{\alpha}_2(\cdot|\cdot), \hat{\pi}(\cdot|\cdot)) \in [\Delta(A_2) \times \Delta(A_1)]^{\hat{M}}$ is *Nash attentive* if it satisfies $(\hat{\alpha}_2(\cdot|\hat{m}), \hat{\pi}(\cdot|\hat{m})) = \hat{m}$ whenever the direct message $\hat{m} = (\rho, \sigma) \in \hat{M} = \Delta(A_1) \times \Delta(A_2)$, viewed as a pair of mixed strategies, constitutes a Nash equilibrium of the game Γ_0 without cheap talk. A PBE strategy–belief profile is *Nash attentive* if player 2's strategy–belief system is Nash attentive.

5.2 First-Mover Advantage with Cheap Talk

We now show that the PBEs of $\hat{\Gamma}$ with Nash attentive beliefs generate Nash equilibria in Γ_0 that are optimal for the first mover.

Definition 4. In the game Γ_0 without cheap talk, the Nash equilibrium mixed strategy pair $(\alpha_1^*, \alpha_2^*) \in \Delta(A_1) \times \Delta(A_2)$ is *optimal for player 1* if $v_1(\alpha_1^*, \alpha_2^*) \geq v_1(\alpha_1, \alpha_2)$

for all (α_1, α_2) in the set E_0 of mixed strategy Nash equilibria in Γ_0. The same pair is *uniquely optimal for player 1* if $v_1(\alpha_1^*, \alpha_2^*) > v_1(\alpha_1, \alpha_2)$ for all alternative Nash equilibria $(\alpha_1, \alpha_2) \in E_0 \setminus \{(\alpha_1^*, \alpha_2^*)\}$.

Theorem 2. *Let* $(\hat{\alpha}_1^*, \hat{m}^*, \hat{\alpha}_2^*(\cdot|\cdot), \hat{\pi}^*(\cdot|\cdot))$ *be any straightforward Nash attentive PBE strategy–belief profile in the game* $\hat{\Gamma}$ *with predictable direct cheap talk. Then the action profile* $(\alpha_1^*, \alpha_2^*) := (\hat{\alpha}_1^*, \hat{\alpha}_2^*(\cdot|\hat{m}^*))$ *induced on the equilibrium path is an optimal Nash equilibrium for player 1 in the game* Γ_0 *without cheap talk.*

Proof. Applying equilibrium condition (13) to $\hat{\Gamma}$ instead of Γ gives

$$(\hat{\alpha}_1^*, \hat{m}^*) \in \arg\max_{(\alpha_1, \hat{m}) \in \Delta(A_1) \times \hat{M}} v_1(\alpha_1, \hat{\alpha}_2^*(\cdot|\hat{m}). \tag{24}$$

Let $(\bar{\alpha}_1, \bar{\alpha}_2) \in E_0$ be any Nash equilibrium in Γ_0. Because player 2's strategy $\hat{\alpha}_2^*(\cdot|\hat{m})$ is Nash attentive in the game $\hat{\Gamma}$, player 1's expected payoff from choosing $(\alpha_1, \hat{m})$ with $\alpha_1 = \bar{\alpha}_1$ and $\hat{m} = (\bar{\alpha}_1, \bar{\alpha}_2)$ will be

$$v_1(\bar{\alpha}_1, \hat{\alpha}_2^*(\cdot|\hat{m})) = v_1(\bar{\alpha}_1, \bar{\alpha}_2). \tag{25}$$

Now (24) implies that $v_1(\hat{\alpha}_1^*, \hat{\alpha}_2^*(\cdot|\hat{m}^*)) \geq v_1(\bar{\alpha}_1, \hat{\alpha}_2^*(\cdot|\hat{m}))$, and so $v_1(\alpha_1^*, \alpha_2^*) \geq v_1(\bar{\alpha}_1, \bar{\alpha}_2)$ by (25). This holds for every $(\bar{\alpha}_1, \bar{\alpha}_2) \in E_0$. But Lemma 1 implies that $(\alpha_1^*, \alpha_2^*) \in E_0$, so it must be an optimal Nash equilibrium for player 1. □

The next definition considers what happens when player 2 may not be fully Nash attentive, but is nevertheless attentive at least to messages that suggest following a Nash attentive straightforward PBE.

Definition 5. A straightforward PBE strategy–belief profile in the game $\hat{G}$ with direct cheap talk is *credible* if it is identical to a Nash attentive straightforward PBE along the equilibrium path.

Obviously, by Theorem 2, any such credible PBE must also induce an optimal Nash equilibrium outcome for player 1.

5.3 First-Mover Advantage without Cheap Talk

Suppose the game $\hat{\Gamma}$ with predictable direct cheap talk has a unique credible PBE. Then the two players can reasonably expect each other to infer what this direct cheap talk would be, even in the game Γ_0 without cheap talk. The following definition singles out the corresponding Nash equilibrium of this game.

Definition 6. A Nash equilibrium of the game Γ_0 without cheap talk is *sophisticated* if it is induced by a credible straightforward PBE of the corresponding game $\hat{\Gamma}$ with predictable direct cheap talk, and moreover this credible PBE is unique.

Fig. 4 A game with no sophisticated equilibrium

	l	r
L	1, 1	0, 0
R	0, 0	1, 1

Theorem 3. *Suppose (α_1^*, α_2^*) is a uniquely optimal Nash equilibrium for player 1 in Γ_0. Then (α_1^*, α_2^*) is the unique sophisticated equilibrium.*

Proof. Theorem 2 implies that there is a unique credible PBE of $\hat{\Gamma}$, and that this equilibrium induces (α_1^*, α_2^*). □

Figure 4 specifies an example of a normal form game G_0 in which, if player 1 moves first in the associated extensive form Γ_0, there is no sophisticated equilibrium. Not surprisingly, cheap talk plays a key role here in enabling coordination on one of the two Nash equilibria that are equally good for player 1. But if the two players' payoffs after (L, ℓ) were $(1+\varepsilon, \delta)$ instead, for any $\varepsilon > 0$ and any $\delta > 0$, then (L, ℓ) would be the unique sophisticated equilibrium.

6 The Special Case of Virtual Observability

6.1 Definition

Corresponding to our basic game Γ_0 without cheap talk, there is an associated extensive form game

$$\Gamma_1 := \langle \{1,2\}, \Delta(A_1), [\Delta(A_2)]^{A_1}, v_1, v_2 \rangle \tag{26}$$

of perfect information, where player 2 is informed of 1's move and so can make his mixed strategy $\alpha_2 \in \Delta(A_2)$ a function of player 1's action a_1. Now the Battle of Sexes example of Fig. 1 has a unique sophisticated equilibrium where both players effectively act as though player 1's move could indeed be observed. It is a case where the same pure strategy profile $(a_1, a_2) \in A_1 \times A_2$ in the game G_0 happens to be both the unique outcome of any credible PBE in $\hat{\Gamma}$ and of any subgame perfect equilibrium in Γ_1. Weber et al. (2004) call this "virtual observability." The next three examples remind us that it is really a very special property.

6.2 Duopoly: Cournot vs. Stackelberg

Consider a duopoly where firm 1 is able to choose its quantity before firm 2. Also, suppose both firms know this and that firm 2 can observe 1's output. Then it is fairly obvious that any sophisticated equilibrium must be a subgame perfect equilibrium where firm 1 acts as a Stackelberg leader and firm 2 as a follower. If firm 1's output remains hidden, however, the normal form of the game corresponds to one in which the duopolists choose their quantities simultaneously. Then a sophisticated equilibrium is Cournot.

For example, suppose each firm $i \in \{1,2\}$ has the profit function

$$\Pi_i(q_i, q_j) = \beta_i q_i - \gamma q_i q_j - \tfrac{1}{2} q_i^2,$$

which is quadratic in its own quantity q_i and also depends on the other's quantity q_j. Suppose too that each firm is risk neutral and so maximizes expected profit. Finally, suppose that the three parameters β_1, β_2, and γ are positive and satisfy the restrictions $\beta_1 > \gamma\beta_2$, $\beta_2 > \gamma\beta_1$, and $\gamma < 1/\sqrt{2}$. Even if the first firm pursues a mixed strategy, the second firm's optimal choice satisfies $q_2 = \beta_2 - \gamma \mathbb{E}\, q_1$, where $\mathbb{E}$ denotes the mathematical expectation. Thus, the first firm's expected profit is

$$\mathbb{E}\Pi_1 = (\beta_1 - \gamma q_2)\mathbb{E}q_1 + \gamma^2(\mathbb{E}q_1)^2 - \tfrac{1}{2}\mathbb{E}q_1^2.$$

This is maximized by choosing the Stackelberg leader's pure strategy $q_1^{\mathrm{S}} := (\beta_1 - \gamma\beta_2)/(1-2\gamma^2)$, which exceeds the unique Cournot equilibrium quantity $q_1^{\mathrm{C}} := (\beta_1 - \gamma\beta_2)/(1-\gamma^2)$. It follows that virtual observability fails, even though there is a unique Nash equilibrium and it uses pure strategies.

6.3 Mixed Strategies

Consider the simple and familiar example of matching pennies, whose normal form is shown in Fig. 5. There is a unique Nash equilibrium, associated with a unique straightforward PBE strategy–belief profile in the corresponding game of predictable direct cheap talk. The only direct message $\hat{m} = (\rho, \sigma) \in \Delta(\{H,T\}) \times \Delta(\{h,t\})$ that is sent in this unique equilibrium has $\rho(H) = \rho(T) = \sigma(h) = \sigma(t) = \frac{1}{2}$. Obviously, the need for mixed action strategies in Nash equilibrium implies that virtual observability cannot hold.

6.4 Multiple Nash Equilibria

The game in Fig. 6 is matching pennies played for a stake of \$4 supplied by a third party. The game is also extended by allowing each (steady handed) player to choose "edge" as well as heads or tails. If just one player chooses edge, the stake is withdrawn, and neither wins anything. But if both choose edge the third party pays each \$1 for being imaginative.

	h	t
H	$1,-1$	$-1,1$
T	$-1,1$	$1,-1$

Fig. 5 Matching pennies

Fig. 6 Extended matching pennies

	h	t	e
H	4, 0	0, 4	0, 0
T	0, 4	4, 0	0, 0
E	0, 0	0, 0	1, 1

In the corresponding extensive form game Γ_1 with perfect information where player 1 moves first, player 2 would choose t in response to H; h in response to T; and e in response to E. So Γ_1 has (E,e) as a unique subgame perfect equilibrium. This is not induced by a credible straightforward PBE of $\hat{G}$; however, because a better Nash equilibrium of G_0 for player 1 is the familiar mixed strategy equilibrium with $\alpha_1(H) = \alpha_1(T) = \alpha_2(h) = \alpha_2(t) = \frac{1}{2}$, since player 1's expected payoff is 2 rather than 1. Once again, virtual observability fails, and in this case it does so even though the unique subgame perfect equilibrium is a Nash equilibrium in pure strategies.

6.5 Implications of Virtual Unobservability

When virtual observability fails, the extensive game Γ_0 is fundamentally different from Γ_1 where player 2 is informed of player 1's earlier move. Sometimes, as in Figs. 5 and 6, this is because player 1 gains by keeping her initial move concealed. Sometimes, however, as in Sect. 6.2, player 1 could gain from having her initial move revealed. In that example, the first duopolist would earn more profit from being a Stackelberg leader. It would also like to report having chosen the Stackelberg leader's optimal quantity q_1^S, expecting the second firm to choose its best response $q_2^S := \beta_2 - \gamma q_1^S$. However, that report is not credible because, if it were believed, the first firm does even better by choosing its best response $q_1 = \beta_1 - \gamma q_2^S \neq q_1^S$. So requiring the follower to be attentive only to the Nash equilibrium message q_1^C in any Nash attentive straightforward PBE imposes a binding constraint on the leader's strategy choice.

7 Concluding Remarks

7.1 Beyond Experimental Anomalies

Experimental economists have recognized that there is a first-mover advantage in Battle of the Sexes and similar games. They typically ascribe this advantage, however, to "positional" or "presentational" effects, suggesting the need to look beyond orthodox rationality concepts in order to explain their experimental results.

This chapter, by contrast, introduces a "sophisticated" refinement of Nash equilibrium that can explain first-mover advantage using only a minor variation of standard rationality and equilibrium concepts. This refinement, like the "manipulated Nash equilibrium" concept explored in Amershi, Sadanand, and Sadanand (1985, 1989b, 1989a, 1992) and in Sadanand and Sadanand (1995), depends on the extensive form of the game. So it violates von Neumann's hypothesis of normal form invariance. Unlike manipulated Nash equilibrium, however, the tacit communication that underlies forward induction arguments is explicitly modeled through a corresponding game with cheap talk. This cheap talk is required to be predictable so that it can remain tacit.[4]

Nevertheless, the precise relationship between sophisticated and manipulated Nash equilibrium deserves further exploration. The ideas presented here should also be applied to a much broader class of games, starting with the "recursive games" considered in Hammond (1982).

7.2 Beyond Orthodox Game Theory

Much of orthodox game theory is built on two assumptions of what one may call the "ZNK paradigm" – due to Zermelo (1912), von Neumann (1928), and Kolmogorov (1933). This chapter has criticized normal form invariance, the second of these. But the first, claiming that games can be modeled with a single extensive form, is also questionable, as discussed in Hammond (2007). So, of course, is a third key assumption, namely that all players are fully rational, and so will always find the optimal action at each information set.

Indeed, following Zermelo (1912), orthodox game theory predicts that any two-person zero-sum game of perfect information such as Go should be played perfectly, and so perfectly predictably. Yet we find the following in a prominent novel by an author who won the Nobel Prize for Literature in 1968.

> 'This is what war must be like,' said Iwamoto gravely.
>
> He meant of course that in actual battle the unforeseeable occurs and fates are sealed in an instant. Such were the implications of White 130. All the plans and studies of the players, all the predictions of us amateurs and of the professionals as well had been sent flying.
>
> As an amateur, I did not immediately see that White 130 assured the defeat of the 'invincible Master.'
>
> Yasunari Kawabata (1954) *The Master of Go*, translated from the author's own shortened version by Edward G. Seidensticker (New York: Alfred A. Knopf, 1972); end of Chapter 37.

Such considerations remind us how far the three standard assumptions take us from reality. To conclude, it seems that the systematic study of games and economic behavior has barely progressed beyond a promising but possibly misleading beginning.

---

[4] A conjecture is that relaxing predictability in the game with cheap talk would allow player 1 to achieve her optimal correlated equilibrium. Where this is better than her optimal Nash equilibrium, cheap talk is essential as a correlation device. Without it, player 2 cannot infer what correlated equilibrium strategy to choose.

Acknowledgements A key idea used here appeared in Hammond (1982). This was an extensive revision of notes originally prepared for a seminar at the Mathematical Economics Summer Workshop of the Institute for Mathematical Studies in the Social Sciences, Stanford University, in July 1981. Eric van Damme, Elon Kohlberg, David Kreps, and Robert Wilson aroused and then revived my interest in this topic, while Hervé Moulin, Richard Pitbladdo, Kevin Roberts, Stephen Turnbull, and other seminar participants made helpful comments, though their views may not be at all well represented here. Research support from the National Science Foundation at that time is gratefully acknowledged.

The earlier chapter had a serious flaw, however, because its “sophisticated” equilibria could fail to be Nash. Much later, Luis Rayo and other members of my graduate game theory course at Stanford made me aware that there was some experimental corroboration of the ideas presented here, thus re-awakening my interest. Most recently, Geir Asheim and especially Ilya Segal have made suggestions that have led to significant improvements.

My gratitude to all those named above, while recognizing that the usual disclaimer absolving them of responsibility applies even more than usual. Some ideas in the first part of this later version were included in my presentation to the conference honouring Kotaro Suzumura at Hitotsubashi University in March 2006. The remainder of this presentation appears elsewhere as the basis of Hammond (2007).

References

Amershi, A., Sadanand, A., & Sadanand, V. (1985). *Manipulated Nash equilibria I: Forward induction and thought process dynamics in extensive form* (Economics Working Paper 928). University of British Columbia

Amershi, A., Sadanand, A., & Sadanand, V. (1989a). *Manipulated Nash equilibria III: Some applications and a preliminary experiment* (Economics Working Paper 1989-6). University of British Columbia

Amershi, A., Sadanand, A., & Sadanand, V. (1989b). *Manipulated Nash equilibria II: Some properties* (Economics Working Paper 1989-5). University of British Columbia

Amershi, A., Sadanand, A., & Sadanand, V. (1992). Player importance and forward induction. *Economics Letters*, *38*, 291–297

Asheim, G., & Dufwenberg, M. (2003). Deductive reasoning in extensive games. *Economic Journal*, *113*, 305–325

Aumann, R. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, *55*, 1–18

Aumann, R., & Hart, S. (2003). Long cheap talk. *Econometrica*, *71*, 1619–1660

Battigalli, P. (1997). On rationalizability in extensive games. *Journal of Economic Theory*, *74*, 40–60

Battigalli, P., & Siniscalchi, D. (1999). Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory*, *88*, 188–230

Battigalli, P., & Siniscalchi, D. (2002). Strong belief and forward-induction reasoning. *Journal of Economic Theory*, *106*, 356–391

Berhmeim, B. (1984). Rationalizable strategic behavior. *Econometrica*, *52*, 1007–1028

Cooper, R., Dejong, D., Forsythe, R., & Ross, T. (1989). Communication in the battle-of-the-sexes games: Some experimental evidence. *Rand Journal of Economics*, *20*, 568–587

Cooper, R., Dejong, D., Forsythe, R., & Ross, T. (1993). Forward induction in the battle-of-the-sexes games. *American Economic Review*, *83*, 1303–1315

Forges, F. (1986). An approach to communication equilibria. *Econometrica*, *54*, 1375–1385

Güth, W., Huck, S., & Rapoport, A. (1998). The limitations of the positional order effect: Can it support silent threats and non-equilibrium behavior? *Journal of Economic Behavior and Organization*, *34*, 313–325

Hammond, P. (1982). *Sophisticated dynamic equilibria for extensive games* (Economics Technical Report). Institute for Mathematical Studies in the Social Sciences, Stanford University

Hammond, P. (1993). Aspects of rationalizable behavior. In K. Binmore, A. Kirman, & P. Tani (Eds.), *Frontiers of game theory* (pp. 277–305). Cambridge, MA: M.I.T Press

Hammond, P. (2007). Schumpeterian innovation in modelling decisions, games, and economic behaviour. *History of Economic Ideas*, *15*, 179–195

Kolmogorov, A. (1933). *Grundbegriffe der wahrscheinlichkeitsrechnung*. Berlin; translated as *Foundations of Probability*, Chelsea, New York: Springer, 1956

Kreps, D. (1990). *Game theory and economic modelling*. Oxford: Oxford University Press

Kreps, D., & Wilson, R. (1982). Sequential equilibrium. *Econometrica*, *50*, 863–894

Kuhn, H. (1953). Extensive games and the problem of information. In H.W. Kuhn (Ed.), *In contributions to the theory of games, II* (pp. 193–216); reprinted in *Classics in game theory.* Princeton: Princeton University Press, 1997

Kumar, P. (1985). *Consistent mechanism design and the noisy revelation principle*. Unpublished doctoral dissertation, Department of Economics, Stanford University

Mailath, G., Samuelson, L., & Swinkels, J. (1993). Extensive form reasoning in normal form games. *Econometrica*, *61*, 273–302

Muller, R., & Sadanand, A. (2003). Order of play, forward induction, and presentation effects in two-person games. *Experimental Economics*, *6*, 5–25

Myerson, R. (1978). Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, *7*, 73–80

Myerson, R. (1982). Optimal coordination mechanisms in generalized principal–agent problems. *Journal of Mathematical Economics*, *10*, 67–81

Neumann, J. von, & Morgenstern, O. (1943, 1953). *Theory of games and economic behavior, 3rd. ed.* Princeton: Princeton University Press

Pearce, D. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, *52*, 1029–1050

Rapoport, A. (1997). Order of play in strategically equivalent games in extensive form. *International Journal of Game Theory*, *26*, 113–136

Sadanand, A., & Sadanand, V. (1995). Equilibria in non-cooperative games II: Deviations based refinements of Nash equilibrium. *Bulletin of Economic Research*, *47*, 93–113

Schotter, A., Weigelt, K., & Wilson, C. (1994). A laboratory investigation of multiperson rationality and presentation effects. *Games and Economic Behavior*, *6*, 445–468

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*, *121*, 301–324; 667–689

Selten, R. (1975). Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, *4*, 25–55

von Neumann, J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, *100*, 295–320

Weber, R., Camerer, C., & Knez, M. (2004). Timing and virtual observability in ultimatum bargaining and 'weak link' coordination games. *Experimental Economics*, *7*, 25–48

Zermelo, E. (1912). Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. In *Proceedings of the fifth international congress of mathematicians, vol. II* (pp. 501–504)

Part IV
Social Welfare and the Measurement of Unemployment and Diversity

Unemployment and Vulnerability: A Class of Distribution Sensitive Measures, its Axiomatic Properties, and Applications

Kaushik Basu and Patrick Nolen

1 Introduction

Traditional measures of unemployment were only concerned with the total number of people unemployed. In recent years such measures have come under criticism for ignoring those who may not currently be unemployed but are vulnerable, that is, they live under the risk of *becoming* unemployed (see Cunningham and Maloney (2000), Glewwe and Hall (1998), Thorbecke (2003)). Alongside this criticism a small but rapidly growing literature is emerging that looks at various aspects of vulnerability and tries to measure it (Amin, Rai, and Topa (2003), Ligon and Schechter (2003), Pritchett, Suryahadi, and Sumarto (2000)).[1]

There is a presumption in much of this literature and the policy statements of international organizations and governments that since vulnerability is bad, we should craft policy to rescue people from being vulnerable. We argue in this paper that such a prescription is wrong, or, at best, misleading. Under a variety of "normal" situations, having some people vulnerable to unemployment makes the *aggregate* problem of unemployment less severe (and more bearable).

The aim of this paper is to explain this normative stance of ours, to develop a class of unemployment measures that take account of this stance, and then to apply it to US and South African data.

The explanation of our normative position is not complicated and the general point can be made simply enough. Suppose there is a society in which, currently,

K. Basu
Cornell University, Ithaca, NY 14853
e-mail: kb40@cornell.edu

P. Nolen
University of Essex, Colchester CO4 3SQ, UK
e-mail: pjnolen@essex.ac.uk

[1] An important precursor of this literature is a body of writing that occurred around the theme of income mobility: see, for instance Fields (1996), Grootaert and Kanbur (1995), Shorrocks (1978).

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

some people are unemployed and some people are vulnerable to unemployment (that is, there is a probability that they will become unemployed in the next period). The presumption in much of the literature and in many World Bank policy discussions (see, for instance World Bank (2002)) is that the standard measure of unemployment, which ignores the vulnerable, effectively underestimates the aggregate pain of unemployment (which would, presumably, include the pain of its anticipation) in society. We, on the other hand, will argue that the standard measure of unemployment underestimates, not the *pain,* but the *inequity of the pain* of unemployment. Our argument is this – if unemployment holds constant over time and there are, currently, some people vulnerable to unemployment, then there must be some currently unemployed people who have a positive probability of becoming employed in the next period. If this is so, then an *aggregate* (that is, an economy-wide) measure of *effective* unemployment, while taking account of the pain of those who live under the risk of unemployment, must also take account of the hope of the currently unemployed who expect to find jobs soon. We argue that in an overall measure of unemployment there is reason to treat the latter as more than offsetting the former. We should clarify that, contrary to the impression that the above sentences might create, we do not take a welfarist approach in this paper but use the above argument concerning the pain of living under the risk of unemployment as *motivation* for creating a class of distribution-sensitive measures of unemployment.

Consider the point some would make that we are not right to assume that just because there are some people who are vulnerable to unemployment, there must be people currently unemployed but who have a positive probability of finding jobs in the next period. Our response to this is that if there were no such people, then having people who are vulnerable to unemployment is equivalent to saying that unemployment will rise tomorrow. If we then treat the situation as worse than what the standard measure captures, this does not show our valuation of vulnerability but the fact that the absolute amount of unemployment is about to rise. To isolate our attitude towards *vulnerability,* we must consider a case where the vulnerable population rises, but the total number unemployed remains unchanged. But this compels us to assume that a vulnerable population will be matched by a population expecting a converse shift – *out* of unemployment.

To close the argument consider two societies, x and y, in which unemployment is the same, say 10%, and this remains constant over time. However, in society x no one is vulnerable to unemployment, while in y, 10% are vulnerable, that is, they are currently employed but face a risk of unemployment. In other words, the total amount of the burden of unemployment to be shared in both societies is the same (10% of the people will have to be unemployed) but in y this burden is shared by 20% of the population, while in x this is borne entirely by only 10% of the population. The same way that, *ceteris paribus* (to use a term rapidly going into extinction), greater equality in the distribution of income and wealth ("good things," that is) is valued positively in most societies, we feel that there is reason to prefer a society where the "bads," such as unemployment, are more equally distributed. It follows that, starting with society $x,$ if vulnerability is increased and we reach society y, then we must consider this a change for the better. Therefore, the effective unemployment must be considered to be less in society y than in x.[2] The next section formalizes the above idea by suggesting a new measure of effective unemployment.

The example just given is the idea that will be tracked in this paper. A full-blown analysis of individual vulnerability (as in Calvo and Dercon (2005), Dercon (2005), Ligon and Schechter (2003)) entails "dynamics" – how a person who may be employed today fares tomorrow. But in this paper we stay away from a full-blown dynamic account of vulnerability. What we are, instead, interested in is the common ground that lies between questions of vulnerability and questions of distribution that arise when we are attempting to create an *aggregate* measure of effective unemployment. Hence, the critical question that we are concerned with is the normative issue of how the "distribution of some fixed amount of unemployment time may be captured," with an inclination to favor "any trend towards a more equal distribution of unemployment" (Shorrocks, 1994, p. 5).

These concerns are shared in a parallel and large literature on unemployment durations (see, for instance Akerlof and Main (1980), Clark and Summers (1979), Shorrocks (1992)). However, the effort to bring these concerns under one aggregate measure of effective unemployment is quite rare (Borooah (2002), Paul (1992), Shorrocks (1994), are the only ones that these authors can think of); and that is what is attempted in this paper.

2 Effective Unemployment

Consider a society with n persons. Let r_i be the fraction of a year during which person i is unemployed. Hence, by the measure of the "standard unemployment rate" this society's unemployment is

$$U \equiv \frac{r_1 + r_2 + \cdots + r_n}{n}. \tag{1}$$

The standard unemployment measure that one encounters in newspapers is usually the above measure (often multiplied by 100, since the measure is generally stated in percentage terms).

From the discussion in the previous section it should be evident that we are looking for a measure of unemployment that is distribution sensitive. That is, if the same aggregate unemployment is unevenly shared in one society, we shall consider the **effective** unemployment to be greater in the more unequal society. We codify this later, in Axiom E, as the "equity axiom."

Let us define an **unemployment profile** of a society to be a vector $(r_1, r_2, \ldots, r_n)$ such that, for all i, $r_i \in [0,1]$. Let Δ be the collection of all **unemployment profiles**.

[2] Another case for better sharing of "unemployment" can be made by arguing that, within each household, the unemployed are helped by the employed. In such a situation there arises the case for a better distribution of employment across households, as was argued in Basu and Foster (1998) in the context of literacy.

Hence, $\Delta = \{(r_1, r_2, \ldots, r_n) \,|\, n \in Z_{++}$ and $r_i \in [0,1], \forall i\}$, where Z_{++} is the set of strictly positive integers.

Formally, a **measure of unemployment** (hereafter referred to as MOU) is a mapping

$$M : \Delta \to R_+,$$

where R_+ is the set of nonnegative real numbers.

The MOU that we propose in this paper takes the following form:

$$M^{\beta}(r_1, \ldots, r_n) \equiv \frac{1}{\beta} - \prod_{i=1}^{n} \left(\frac{1}{\beta} - r_i\right)^{\frac{1}{n}}, \tag{2}$$

$$\text{where } \beta \in (0,1).$$

Since for every $\beta \in (0,1)$ we have a distinct measure M^{β}, what we have just proposed is a *class* of new measures of unemployment. We show that these measures have appealing properties, demonstrate, with some actual empirical examples, how using these new measures make a difference to the description of unemployment and then fully characterize these measures. Let us from now on call an MOU defined by (2), above, an **effective unemployment rate**.

One property of every member of the family of effective unemployment rates worth observing at the outset is that if $R = (r_1, r_2, \ldots, r_n)$ is such that $r_i = r \; \forall i$, then $M^{\beta}(R) = r$. In other words, if the burden of unemployment is perfectly equitably shared by everybody, then the effective unemployment rate is independent of $\beta \in (0,1)$ and equal to the standard unemployment rate defined in (1).

It is worth checking what the limits or boundaries of our class of measures look like. First consider the case where $\beta = 1$. This measure (which is not a part of the class we are recommending) is then represented by: $M^1(r_1, r_2, \ldots, r_n) = 1 - \prod_{i=1}^{n} (1 - r_i)^{\frac{1}{n}}$. Note that if, for some i, $r_i = 1$ (one person is fully unemployed) then $M^1 = 1$. Hence, this measure makes no distinction between the cases where 1 person is fully unemployed and where 10 persons are fully unemployed. It amounts to an extreme evaluation where a tragedy for one is a tragedy for all. This is the standard "multiplicative" form of an evaluation function.

Now, what about the other limit, that is as β goes to 0? It can be shown that as $\beta \to 0$, $M^{\beta} \to U$. That is, as β goes to 0, our measure converges to the standard unemployment rate as defined by (1). The first lemma establishes this result. Since the standard measure is one in which individuals' unemployment are aggregated by simply adding up, this could be thought of as a kind of utilitarian or additive representation of unemployment. Hence the class of measures that we are proposing is bounded at one end by a multiplicative representation and at the other end by an additive one.

Lemma 1. *For all $R = (r_1, r_2, \ldots, r_n) \in \Delta$, and for all $\beta \in (0,1)$,*
$\lim_{\beta \to 0} M^{\beta}(R) = \frac{\sum_{i=1}^{n} r_i}{n}$.

Proof.

$$\begin{aligned}
\lim_{\beta\to 0} M^{\beta}(R) &= \lim_{\beta\to 0}\left[\frac{1}{\beta}-\prod_{i=1}^{n}\left(\frac{1}{\beta}-r_i\right)^{\frac{1}{n}}\right],\\
&= \lim_{\beta\to 0}\frac{1}{\beta}\left[1-\prod_{i=1}^{n}\beta^{\frac{1}{n}}\left(\frac{1}{\beta}-r_i\right)^{\frac{1}{n}}\right],\\
&= \lim_{\beta\to 0}\frac{1}{\beta}\left[1-\prod_{i=1}^{n}(1-\beta r_i)^{\frac{1}{n}}\right],\\
&= \lim_{\beta\to 0}\frac{1-\prod_{i=1}^{n}(1-\beta r_i)^{\frac{1}{n}}}{\beta}=\frac{0}{0}.
\end{aligned}$$

So we may now use L'Hôpital's rule. Note that

$$\frac{\partial}{\partial\beta}\beta = 1$$

and

$$\frac{\partial}{\partial\beta}\left[1-\prod_{i=1}^{n}(1-\beta r_i)^{\frac{1}{n}}\right] = -\sum_{k=1}^{n}\frac{1}{n}(1-\beta r_k)^{\frac{1-n}{n}}(-r_k)\prod_{i\neq k}(1-\beta r_i)^{\frac{1}{n}}.$$

Taking the limit of this numerator we get

$$\begin{aligned}
\lim_{\beta\to 0}\left[-\sum_{k=1}^{n}\frac{1}{n}(1-\beta r_k)^{\frac{1-n}{n}}(-r_k)\prod_{i\neq k}(1-\beta r_i)^{\frac{1}{n}}\right] &= -\sum_{k=1}^{n}\frac{1}{n}(-r_k),\\
&= \frac{1}{n}\sum_{k=1}^{n}r_k.
\end{aligned}$$

Thus by L'Hôpital's rule,

$$\lim_{\beta\to 0}\frac{1-\prod_{i=1}^{n}(1-\beta r_i)^{\frac{1}{n}}}{\beta} = \frac{1}{n}\sum_{k=1}^{n}r_k,$$

which implies that

$$\lim_{\beta\to 0}M^{\beta}(R) = \frac{1}{n}\sum_{k=1}^{n}r_k.$$

□

We shall now demonstrate how the effective unemployment rate, as characterized by (2), satisfies some attractive axioms. First of all consider two routine axioms.

Axiom O (Monotonicity Axiom). An MOU, M, is said to satisfy the monotonicity axiom if for any $R = (r_1, r_2, \ldots, r_n) \in \Delta$ and $R' = (r'_1, r'_2, \ldots, r'_n) \in \Delta$ such that, $\forall i$, $r_i \geq r'_i$ and $\exists j$, where $r_j > r'_j$, then $M(R) > M(R')$.

Axiom P (Population Replication Axiom). An MOU, M, is said to satisfy the population replication axiom if for any $R = (r_1, r_2, \ldots, r_n) \in \Delta$ and $R^k = (r'_1, r'_2, \ldots, r'_{kn}) \in \Delta$, where R^k is a k-replica of R for some positive integer k (that is $r'_j = r_i$, $\forall j \in \{1 + (i-1)k, \ldots, ik\}$, $\forall i \in \{1, \ldots, n\}$), then $M(R) = M\left(R^k\right)$.

These two axioms are standard and we would expect a good measure to satisfy them. Fortunately – as is easy to see – the effective unemployment rate that we have proposed satisfies both these axioms. Observe that, given the monotonicity axiom, coupled with the fact that $M^\beta(1, 1, \ldots, 1) = 1$, we now know that our measure ranges from 0 to 1. That is, $M^\beta(\Delta) \subset [0, 1]$.

Our measure, and the need to break away from the standard unemployment concept, was motivated by using an equity argument, namely, that it is superior to have a society where the burden of a certain amount of aggregate unemployment is more widely shared. So it is important to check that the effective unemployment rate satisfies equity. The simplest idea of equity may be formalized as follows.

Axiom E (Equity Axiom). An MOU, M, is said to satisfy the equity axiom if for $R = (r_1, r_2, \ldots, r_n) \in \Delta$ and $R^* = (r^*, r^*, \ldots, r^*) \in \Delta$ such that $\sum_{i=1}^n r_i = nr^*$ and $R \neq R^*$, then $M(R) > M(R^*)$.

It can be shown that M^β satisfies the equity axiom for every $\beta \in (0, 1)$. But instead of showing this directly, we show that M^β satisfies another axiom and then show that the latter implies the equity axiom. This other axiom is the "transfer axiom" widely used in the literature on poverty and inequality measurement (see Sen (1976) for instance). This, in the context of unemployment, says the following. Suppose there are two people, one who is unemployed more than the other. Now if the more unemployed person becomes even more unemployed – say by ε amount of time – and the less unemployed person finds more work – again by ε amount of time – then the effective unemployment is higher. Formally,

Axiom T (Transfer Axiom). An MOU, M, is said to satisfy the transfer axiom if for any $R = (r_1, r_2, \ldots, r_n) \in \Delta$ and $R' = (r'_1, r'_2, \ldots, r'_n) \in \Delta$ such that $r_k = r'_k \;\; \forall k \neq i, j$, $r_i \geq r_j$ and $r'_i = r_i + \varepsilon \leq 1$ and $r'_j = r_j - \varepsilon \geq 0$ (for some $\varepsilon > 0$), then $M(R') > M(R)$.

Lemma 2. *For all $R = (r_1, r_2, \ldots, r_n) \in \Delta$, and for all $\beta \in (0, 1)$ every effective unemployment rate, M^β, satisfies the transfer axiom.*

Proof. $M^\beta(R') = \frac{1}{\beta} - \prod_{k=1}^{n} \left(\frac{1}{\beta} - r'_k\right)^{\frac{1}{n}}$

$$= \frac{1}{\beta} - \left(\frac{1}{\beta} - r'_i\right)^{\frac{1}{n}} \left(\frac{1}{\beta} - r'_j\right)^{\frac{1}{n}} \prod_{k \neq i,j}^{n} \left(\frac{1}{\beta} - r'_k\right)^{\frac{1}{n}},$$

$$= \frac{1}{\beta} - \left[\left(\frac{1}{\beta} - r_i - \varepsilon\right)\left(\frac{1}{\beta} - r_j + \varepsilon\right)\right]^{\frac{1}{n}} \prod_{k \neq i,j}^{n} \left(\frac{1}{\beta} - r_k\right)^{\frac{1}{n}},$$

$$= \frac{1}{\beta} - \left[\left(\frac{1}{\beta} - r_i\right)\left(\frac{1}{\beta} - r_j\right) - (r_i - r_j)\varepsilon - \varepsilon^2\right]^{\frac{1}{n}} \prod_{k \neq i,j}^{n} \left(\frac{1}{\beta} - r_k\right)^{\frac{1}{n}},$$

$$> \frac{1}{\beta} - \left[\left(\frac{1}{\beta} - r_i\right)\left(\frac{1}{\beta} - r_j\right)\right] \prod_{k \neq i,j}^{n} \left(\frac{1}{\beta} - r_k\right)^{\frac{1}{n}} \text{ (since } r_i \geq r_j,\ \varepsilon > 0, \text{ and } \beta \in (0,1)),$$

$$= \frac{1}{\beta} - \prod_{i=1}^{n} \left(\frac{1}{\beta} - r_i\right)^{\frac{1}{n}} = M^{\beta}(R). \qquad \square$$

The fact that M^{β} satisfies the equity axiom follows from Lemma 2 and the following lemma.

Lemma 3. *If an MOU satisfies the transfer axiom, it must satisfy the equity axiom.*

Proof. Suppose M is an MOU that satisfies the transfer axiom. Consider $\widetilde{R} = (r_1, r_2, \ldots, r_n)$ and $R^* = (r^*, r^*, \ldots, r^*)$, which satisfy the hypotheses of the equity axiom. That is, $\widetilde{R}, R^* \in \Delta$, $\widetilde{R} \neq R^*$, and $\sum_{i=1}^{n} r_i = nr^*$.

Define $S \subset \Delta$ such that

$$S \equiv \left\{ R = (r_1, r_2, \ldots, r_n) \in \Delta \mid \sum_{i=1}^{n} r_i = nr^* \right\}.$$

Note that for any $R \neq R^*$, $R = (r_1, r_2, \ldots, r_n) \in S \backslash \{R^*\}$. So we can define $\overline{r}(R) \equiv \max_i r_i$ and $\underline{r}(R) \equiv \min_i r_i$. Let $\varepsilon = \min\{\overline{r}(R) - r^*, r^* - \overline{r}(R)\}$. Now define a mapping $\Psi : S \rightarrow S$ as follows:

$\Psi(R^*) = R^*$ or, if $R = (r_1, r_2, \ldots, r_n) \neq R^*$, then $\Psi(R) = R'$,
where $R' = (r'_1, r'_2, \ldots, r'_n)$ such that $r'_k = r_k$, $\forall r_k \neq \overline{r}(R), \underline{r}(R)$,
and $r'_i = \underline{r}(R) + \varepsilon$ for $r_i = \underline{r}(R)$,
and $r'_j = \overline{r}(R) - \varepsilon$ for $r_j = \overline{r}(R)$.

By the transfer axiom we know that $M(R) > M(\Psi(R))$.

Now look at the infinite sequence $\{R^1, R^2, \ldots\}$ such that $R^1 = \widetilde{R}$ and $R^{t+1} = \Psi(R^t)$ $\forall t > 1$. There must exist some $\bar{t}$ such that $\forall t \geq \bar{t}$, $R^t = R^*$. Thus $M(R^1) > M(R^t)$ $\forall t > 1$, and therefore $M(\widetilde{R}) > M(R^*)$. $\square$

In the light of this result, the next lemma is obvious and stated only for completeness.

Lemma 4. *Every effective unemployment rate, M^{β}, satisfies the Equity Axiom.*

While the measure being suggested here has attractive axiomatic properties, which particular β should one use when applying this measure? One possibility is to study the sensitivity of ranking societies with respect to changes in β. The other is to pick some salient values of β from the interval $(0,1)$ and use those specific measures. This is the strategy that is often used vis-a-vis the Foster–Greer–Thorbecke family of poverty measures (see Foster, Greer, and Thorbecke (1984)).

For such salient β's an obvious one is the half-way mark, that is, $\beta = \frac{1}{2}$. There is another one, $\beta = \frac{8}{9}$, which appears unnatural at first sight, but has a natural explanation.

Consider a society of size n and suppose that x is the fraction of society that has to be unemployed. In other words, the total amount of jobs available is $(1-x)n$. For matters of illustration we are ignoring the fact that $(1-x)n$ may not be an integer. Let us fix x and consider different distributions of the total amount of unemployment nx, and their corresponding measures of effective unemployment. By using the equity axiom it is clear that effective unemployment is minimized if nx is distributed equitably, that is, if each person is unemployed a fraction x of her time.

Let $m(x)$ be the minimum effective unemployment rate for a society with a total burden of unemployment nx. It is easy to see this is independent of $\beta \in (0,1)$. Hence, writing this as $m(x)$, with no mention of β, is fine. It is obvious that $m(x)$ will be the 45° line as shown in Fig. 1. Thus if half the society has to be half unemployed (i.e., $x = \frac{1}{2}$), the lowest value M^β takes is when every person is half-time unemployed. In that case, for all $\beta \in (0,1)$, $M^\beta(x,\ldots,x) = \frac{1}{2}$.

Here is an interesting question. Let us pick any $x \in [0,1]$ and think of the worst distribution of this total burden of x unemployment (in the sense of the distribution that makes effective unemployment the maximum). By the transfer axiom, we know that this happens when some people are fully unemployed and the rest are fully employed. Hence, fix $\beta \in (0,1)$, consider this worst-distribution for every x and define $\overline{M}^\beta(x)$ as the value of M^β for a given unemployment profile, $(r_1, r_2, \ldots, r_n)$, which is the worst way to share the burden of nx. Clearly $\overline{M}^\beta(x) \in (x,1)$, $\forall x$. It is not hard to see that for a given β, $\overline{M}^\beta(x)$ will look something like the curve shown in Fig. 1. The higher the values of β, the higher the curve will be. And as β goes to 0, the line will converge to the $m(x)$ curve.

There are two ways of choosing β. One is to elicit this from individual choice. This involves asking individuals questions like: If you face a choice of two lotteries,

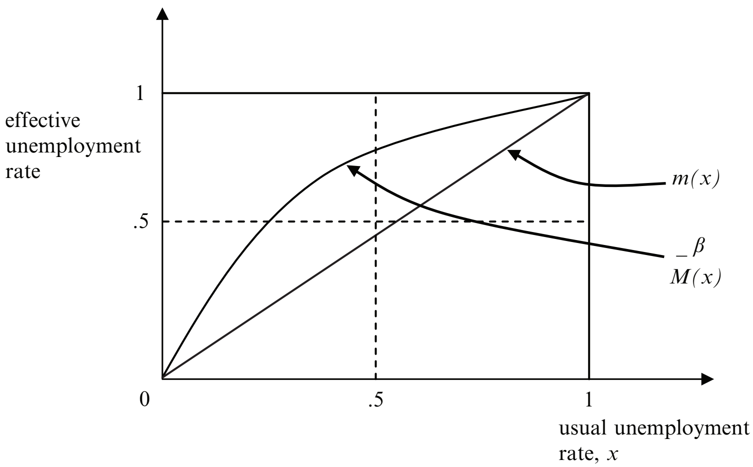


Fig. 1 This shows the relation between the usual and effective unemployment rate for a given value of β

one in which you will be unemployed all year with the probability $\frac{1}{4}$ or employed for the full year with probability $\frac{3}{4}$; and the other in which you will be employed for a fraction t of the year with certainty and unemployed for the remainder of the year, what value of t would you choose? This would be in the spirit of what is done by Ligon and Schechter (2003).

The other way to approach β is as a moral judgement of the policy maker. In the absence of data on individual risk-aversion, let us explore that moral approach here. Just to fix our thinking consider the case of $x = \frac{1}{2}$. We know that if every person is unemployed $\frac{1}{2}$ of the year then $M^{\beta}(\frac{1}{2}) = \frac{1}{2}, \forall \beta$. Now consider the worst distribution of this total burden. Clearly this is one where $\frac{n}{2}$ persons are fully employed and $\frac{n}{2}$ persons are fully unemployed. Let $R = (r_1, r_2, \ldots, r_n)$ signify such a distribution. We know that $M^{\beta}(R) \in (\frac{1}{2}, 1)$ as β varies from 0 to 1. We need to ask ourselves: what score we would like to give to $M^{\beta}(R)$? One simple strategy is to set this half-way in this interval. That is $M^{\beta}(R) = \frac{3}{4}$. In other words, we are making the judgement that a society where half the people are fully employed and half are fully unemployed is equivalent to one where everybody is employed with certainty for one quarter of the time. What would β have to be to yield this mid-way result?

The answer turns out to be, interestingly, $\frac{8}{9}$. To see this note:

$$\begin{aligned}
M^{\beta}(R) &= \frac{1}{\beta} - \left(\frac{1}{\beta} - 1\right)^{\frac{1}{n}\cdot\frac{n}{2}} \left(\frac{1}{\beta} - 0\right)^{\frac{1}{n}\cdot\frac{n}{2}}, \\
&= \frac{1}{\beta} - \left(\frac{1}{\beta} - 1\right)^{\frac{1}{2}} \left(\frac{1}{\beta}\right)^{\frac{1}{2}}, \\
&= \frac{1}{\beta} - \frac{(1-\beta)^{\frac{1}{2}}}{\beta}.
\end{aligned}$$

If $M^{\beta}(R) = \frac{3}{4}$, it follows that $\beta = \frac{8}{9}$. Hence, the $\frac{8}{9}$ rule. We shall use this in the empirical section as one of the salient values. Before moving on we should point out that in practice β could well vary from one country to another. Depending on the policies that nations follow to support the unemployed, the trauma of unemployment can vary from one nation to another. We shall, however, ignore this complication here.

3 Simple Data Exercise

To provide an illustrative example of how our measure works, we require certain information. First, a history of how much someone was unemployed over a certain period of time. For this exercise, we use the weeks or months one was unemployed over a year. Second, we require a value for the parameter β. As explained earlier, for the purpose of illustration, we use the values $\beta = \frac{1}{2}$ and $\beta = \frac{8}{9}$.

The March Current Population Surveys (hereafter referred to as the CPS) for the United States have the amount of weeks any member of the workforce was employed during the previous year. The Labour Force Surveys (hereafter referred to as LFS) for South Africa have the amount of months that one has been unemployed, if she was unemployed at the time of the survey, and when one started a job, if she is currently working at the time of the survey. These data allow us to illustrate the differences of the effective and the standard measures of unemployment. In the South African case one will have to make some assumptions to go from the available data to the measures we wish to calculate.

3.1 United States

We begin with the case of the United States. The CPS contains how many weeks a survey participant had been employed during the previous year. Therefore, since we have the data for the March CPS from 1976 through 2003, we are able to calculate the usual yearly unemployment rate and the effective yearly unemployment rate for the years of 1975 through 2002 (excluding 1993).

To get measures of unemployment as accurate as possible, we tried to exclude students and retired individuals by calculating the unemployment rates only for people between the ages of 25 and 54. Any persons who listed themselves as being unemployed for any of the following reasons were dropped from our data even if they were between the ages of 25 and 54: to take care of house or family; ill or disabled; to attend school; retired. Anyone who claimed to have not worked for the year, but had spent less than four weeks searching for a job was also not included in our data set. Thus, we were left only with people who were able to participate in the labor market during the full year and had been actively seeking work.

Table 1 shows the usual unemployment rate and the effective unemployment rates for $\beta = \frac{1}{2}$ and $\beta = \frac{8}{9}$ over the years for which we have data. Figure 2 puts this information into graphical form. To begin our discussion it is useful to focus attention on three sets of years: 1987–1989, 1991–1992, and 1999–2000. These three periods illustrate how even though the usual unemployment rate stayed roughly constant our effective unemployment measures tell a different story in these periods. The period from 1987 to 1989 has a continuous decrease in the effective unemployment rate at $\beta = \frac{8}{9}$. During this period the usual unemployment rate, though, rose slightly from 1987 to 1988 and then back to slightly below the 1987 level in 1989. Thus, while the usual unemployment measure would rank 1987, 1988, and 1989 as being roughly equivalent, if one takes vulnerability into account – as we have defined it – then 1989 would be ranked as having an unambiguously better employment situation than either of the other two years, and 1988 would be ranked higher than 1987. The effective measure at $\beta = \frac{1}{2}$ shows the same trend but there is almost no difference between 1987 and 1988. Therefore, during the years of 1987–1989, under President Reagan's administration, the burden of unemployment became more equitably shared according to the effective unemployment measure with $\beta = \frac{8}{9}$.

Table 1 Usual and efficient unemployment rates for individuals who were in the labor force, aged 25–54 and available for employment that entire year

Year	Usual unemployment rate	Effective unemployment rate at $\beta = \frac{1}{2}$	Effective unemployment rate at $\beta = \frac{8}{9}$
2002	0.0529	0.0624	0.0803
2001	0.0457	0.0533	0.0669
2000	0.0382	0.0442	0.0544
1999	0.0382	0.0443	0.0547
1998	0.0490	0.0572	0.0721
1997	0.0431	0.0502	0.0630
1996	0.0530	0.0619	0.0782
1995	0.0585	0.0686	0.0874
1994	0.0612	0.0719	0.0922
1993	–	–	–
1992	0.0707	0.0834	0.1079
1991	0.0705	0.0822	0.1039
1990	0.0602	0.0697	0.0864
1989	0.0529	0.0613	0.0761
1988	0.0543	0.0633	0.0797
1987	0.0537	0.0632	0.0819
1986	0.0621	0.0734	0.0958
1985	0.0639	0.0756	0.0987
1984	0.0677	0.0807	0.1071
1983	0.0829	0.0994	0.1341
1982	0.0916	0.1092	0.1455
1981	0.0680	0.0800	0.1037
1980	0.0635	0.0744	0.0953
1979	0.0500	0.0579	0.0726
1978	0.0501	0.0583	0.0735
1977	0.0565	0.0663	0.0850
1976	0.0646	0.0763	0.0996
1975	0.0708	0.0836	0.1089

The symbol " – " implies the unemployment rate could not be calculated for that period

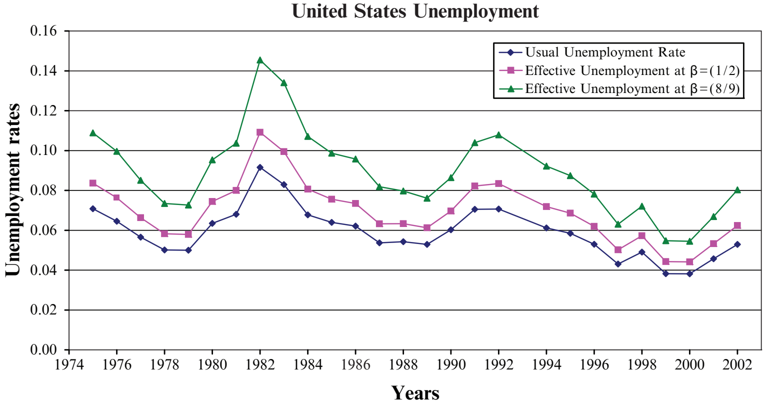


Fig. 2 The usual and effective yearly unemployment rates in the United States over 1975–2002

The period of 1991–1992, under President George Herbert Walker Bush's administration, our effective unemployment measures again deviate from the rankings provided by the usual measure of unemployment. Again during this period we have the usual unemployment rate staying roughly constant, but by both the effective unemployment measure with $\beta = \frac{8}{9}$ and with $\beta = \frac{1}{2}$ effective unemployment measures the burden of unemployment was being shared less equitably in 1992 than in 1991. In contrast to these two examples, the period of 1999–2000, during Bill Clinton's administration, shows that both the usual unemployment measure and the effective unemployment measures rank the years as being roughly equivalent. This shows that there was no significant change in the equity of how the unemployment burden was being shared.

These examples illustrate how our measure can be used to distinguish between years that might seem to be roughly equivalent in regards to unemployment, and shows how the role of vulnerability can cause a re-ranking of how one judges a country's unemployment situation over the years. This illustration uses only information from within one country. The effective unemployment measures could also be used to compare the unemployment situation between countries. Indeed, this will be done below.

3.2 South Africa

The South African LFS are twice yearly surveys that are representative of the Republic of South Africa. The survey collects data on people who are currently unemployed and those currently employed, but figuring out the yearly history of individuals is not as easy as with the CPS. Any person who is currently employed is asked when she began working at that job. Therefore we have a measure of her duration of current employment. Likewise, anyone unemployed is asked how long it has been since she last worked, if she had worked at all, giving us the duration of the current unemployment spell.

In South Africa, the labor unions are rather powerful and, because of the historical situation, firing an individual is rather difficult. Therefore, job turnover is a rarer phenomenon than in the United States. Thus, the duration periods are probably an accurate measures of a labor force participant's job status over the past year. With this in mind the usual and effective unemployment rates were calculated in this "moderate" case. The durations are by no means guaranteed to be accurate representatives of the employment history, though, so we have also constructed bounds to the measures of unemployment.

To understand how these bounds were calculated let us look at a worker who was employed for 6 months at the time of the survey. We know that she has worked at least 6 months over the past year, and the employment pattern for the other 6 months is lost. In the worst case scenario, she spent the previous 6 months searching for the one job she had at the time of the survey and was unemployed the rest of the time. This means that her duration of current employment is the worst history she could have. On the other hand, since we know she was only working at her current job for 6 months she could have found that job after only 1 month or less of unemployment – we count her as having had 1 month of unemployment even if it was less than that, though. Thus her best unemployment profile would be eleven months working and one month unemployed.

Likewise a currently unemployed person has a best and worse case scenario. If a person was unemployed for 6 months, then, in the best case, she was working for the 6 months before that and in the worst case she was working for only 1 month before her unemployment began. The usual and effective unemployment rates were therefore also calculated for both the "worst" and "best" case scenarios.

When comparing the usual unemployment rate under these three scenarios to that listed as the official unemployment rate in the LFS, the moderate rate was almost identical in all five periods.[3] Therefore, in our discussion below we use the results from the calculations done on the moderate data, though, using either of the other scenarios does not change our analysis.

The usual and effective unemployment rates for the moderate case are depicted graphically in Fig. 3. The effective unemployment seems to be simply a horizontal increase of the usual unemployment. This may be because of the coarseness of the LFS data with respect to the CPS data or it may be because of a lack of change in the equity of unemployment. Comparing this graph to that of the United States, though,

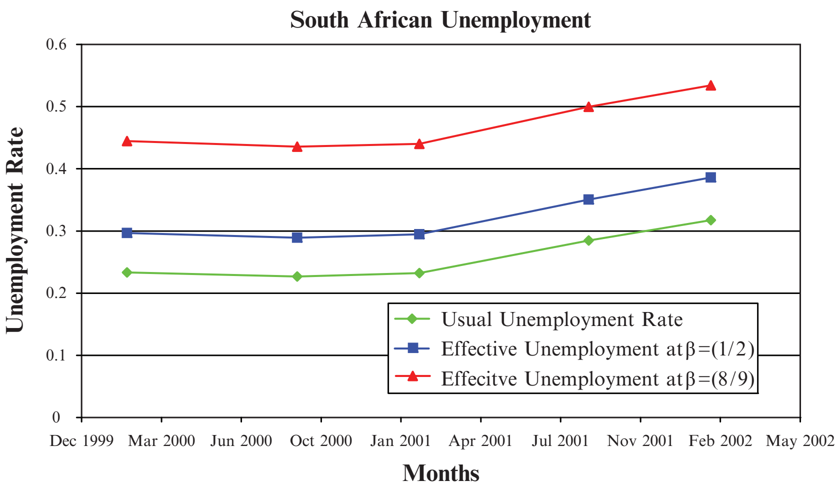


Fig. 3 The usual and effective unemployment rates in South Africa under the moderate scenario for the months of February 2000–2002 and September 2000 and 2001

[3] The official unemployment rates according to the LFS are 26.2%, 25.4%, 26%, 29.2%, and 29%, while the usual unemployment rate calculated using the "moderate" estimates of r_i are 23, 23, 23.2, 28.5, and 31.7 for February 2000, September 2000, February 2001, September 2001, and February 2002, respectively.

one can see that the jump from the usual measure of unemployment to the effective measures is much higher in both the $\beta = \frac{1}{2}$ and $\beta = \frac{8}{9}$ case for South Africa than for the United States. The effective unemployment rate at $\beta = \frac{1}{2}$ was, on average, 1.17 times the usual unemployment rate in the US and never above 1.199 times the usual rate. In South Africa, though, the effective unemployment rate at $\beta = \frac{1}{2}$ is on average 1.25 times the usual unemployment rate and never below 1.22 times the usual rate. Likewise at $\beta = \frac{8}{9}$, the effective unemployment rate is on average 1.50 times and 1.83 times the usual unemployment rate in the US and South Africa, respectively.

One may ask: What does this proportional difference show? Since our effective unemployment measure adjusts for the inequity of unemployment in a society, then the higher jump from the usual to effective rate in the case of South Africa suggests that the burden of unemployment is shared less equitably in South Africa than in the United States. Given South Africa's history, this is not a hard story to believe. To make this claim in a more rigorous manner, and not just to provide an illustration, one would have to test the confidence intervals of the proportions mentioned above. Given the size of the samples in both the LFS and the CPS, the discussion above will probably still be valid. Furthermore, one could ask what is causing the difference in these proportions of unemployment. For example, it may be that we are picking up only frictional unemployment in the United States, but structural unemployment in South Africa.

4 Full Characterization of the Effective Measure

After seeing the usefulness of this measure and its applicability it is worth returning to the theoretical discussion and addressing a natural question. We have seen that our measure of unemployment satisfies several attractive axioms, but is there a set of axioms that exactly characterize our measure or, more precisely, the class of measures we proposed and is stated in (2)? This is what we set out to answer in this section.

A property that we have already discussed but needs to be stated formally is codified in the next axiom.

Axiom C (Coincidence). A MOU, M, satisfies coincidence if $\forall R = (r_1, r_2, \ldots, r_n) \in \Delta$, such that $r_1 = r_2 = \cdots = r_n \equiv r,\ M(R) = r$.

This axiom says that if in some society everybody is unemployed to the same extent as everybody else, then the society's unemployment rate should be equal to the individual's unemployment rate. This is a normalization axiom, which says that if the distribution of unemployment is perfectly egalitarian, then our measure coincides with the standard unemployment rate.

Another axiom that we use and seems eminently reasonable is the following.

Axiom A (Anonymity). The MOU, M, satisfies anonymity if for all $R = (r_1, r_2, \ldots, r_n) \in \Delta$, and for all permutations $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, $M(R) = M\left(r_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(n)}\right)$.

And finally, a much stronger axiom.

Axiom R (Representation). For every individual i, there exists a utility function $u_i : [0,1] \to R_{++}$, and for every $n \in Z_{++}$, there exists an aggregation mapping $F : R^n_{++} \to R$, such that for all $R = (r_1, r_2, \ldots, r_n)$ where $r_i \in [0,1]$,
$M(R) = F(u_1(r_1), u_2(r_2), \ldots, u_n(r_n))$ and
(i) For every i, u_i is affine and decreasing.
(ii) F satisfies anonymity. That is, for all $u = (u_1, u_2, \ldots, u_n) \in R^n_{++}$ and for all permutations $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, $F(u_1, u_2, \ldots, u_n) = F(u_{\sigma(1)}, u_{\sigma(2)}, \ldots, u_{\sigma(n)})$.
(iii) F satisfies scale independence. That is, $F(u_1, u_2, \ldots, u_n) \geq F(u'_1, u'_2, \ldots, u'_n)$ and $(b_1, b_2, \ldots, b_n) \in R^n_{++}$ implies that $F(b_1u_1, b_2u_2, \ldots, b_nu_n) \geq F(b_1u'_1, b_2u'_2, \ldots, b_nu'_n)$.

If an MOU satisfies axiom R, we shall call each u_i function and the F function referred to in the axiom as person i's utility function and the society's aggregation function, respectively.

What Axiom R (iii) says is that a change in the unit for measuring one person's utility must be of no consequence in our social evaluation of the economy. This is not a normative axiom but an informational one, in the spirit of standard "invariance axioms" (Sen, 1974) used in social choice and bargaining theory and meant to capture the degree of measurability of an individual's utility.

Theorem 1. *An MOU, M, satisfies axioms A, C, O, and R if and only if it belongs to the class described in (2). That is,* $\forall R \in (r_1, r_2, \ldots, r_n) \in \Delta$,

$$M(r) = \frac{1}{\beta} - \left[\prod_{i=1}^{n}\left(\frac{1}{\beta} - r_i\right)\right]^{\frac{1}{n}}, \text{ where } \beta \in (0,1).$$

Proof. ($\Rightarrow$) We have already seen that the MOU described in (2) satisfies axioms C and O. Axiom A is obvious. To see that it satisfies axiom R, consider $u_i = \frac{1}{\beta} - r_i$, for all i, and define the welfare mapping F as follows:

$$F = \frac{1}{\beta} - \left[\prod_{i=1}^{n} u_i\right]^{\frac{1}{n}}. \tag{3}$$

It is easy to see that (3) satisfies axiom R.

($\Leftarrow$) Next assume that M is an MOU that satisfies axioms A, C, O, and R. By axiom R we know that $\exists$ is a welfare mapping F such that, $\forall R = (r_1, r_2, \ldots, r_n) \in \Delta$, $M(R) = F(u_1(r_1), u_2(r_2), \ldots, u_n(r_n))$. Now the proof will continue in a series of steps.

- *Step 1:* It will first be shown that F is a transformation of the product of the arguments. In other words, $\forall u = (u_1, u_2, \ldots, u_n) \in R^n_{++}$, $F(u) = \phi\left(\prod_{i=1}^{n} u_i\right)$. Consider a utility vector $u = (u_1, u_2, \ldots, u_n) \in R^n_{++}$. Then we know that $\left(u_1^{\frac{1}{n}}, u_2^{\frac{1}{n}}, \ldots, u_n^{\frac{1}{n}}\right) \in R^n_{++}$.

  Note $F\left(u_1^{\frac{1}{n}}, u_2^{\frac{1}{n}}, \ldots, u_n^{\frac{1}{n}}\right) = F\left(u_2^{\frac{1}{n}}, u_3^{\frac{1}{n}}, \ldots, u_n^{\frac{1}{n}}, u_1^{\frac{1}{n}}\right)$ by axiom R (ii).

  $F\left(u_1^{\frac{1}{n}} u_1^{\frac{1}{n}}, u_2^{\frac{1}{n}} u_2^{\frac{1}{n}}, \ldots, u_n^{\frac{1}{n}} u_n^{\frac{1}{n}}\right) = F\left(u_1^{\frac{1}{n}} u_2^{\frac{1}{n}}, u_2^{\frac{1}{n}} u_3^{\frac{1}{n}}, \ldots, u_{n-1}^{\frac{1}{n}} u_n^{\frac{1}{n}}, u_n^{\frac{1}{n}} u_1^{\frac{1}{n}}\right)$,
  by axiom R (iii).

  $F\left(u_1^{\frac{2}{n}}, u_2^{\frac{2}{n}}, \ldots, u_n^{\frac{2}{n}}\right) = F\left((u_1 u_2)^{\frac{1}{n}}, (u_2 u_3)^{\frac{1}{n}}, \ldots, (u_{n-1} u_n)^{\frac{1}{n}}, (u_n u_1)^{\frac{1}{n}}\right)$,

  $F\left(u_1^{\frac{2}{n}}, u_2^{\frac{2}{n}}, \ldots, u_n^{\frac{2}{n}}\right) = F\left((u_2 u_3)^{\frac{1}{n}}, (u_3 u_4)^{\frac{1}{n}}, \ldots, (u_n u_1)^{\frac{1}{n}}, (u_1 u_2)^{\frac{1}{n}}\right)$,
  by axiom R (ii).

  $F\left(u_1^{\frac{3}{n}}, u_2^{\frac{3}{n}}, \ldots, u_n^{\frac{3}{n}}\right) = F\left((u_1 u_2 u_3)^{\frac{1}{n}}, (u_2 u_3 u_4)^{\frac{1}{n}}, \ldots, (u_n u_1 u_2)^{\frac{1}{n}}\right)$,
  by axiom R (iii).

  Continuing in this manner we get
  $F\left(u_1^{\frac{n}{n}}, u_2^{\frac{n}{n}}, \ldots, u_n^{\frac{n}{n}}\right) = F\left(\left(\prod_{i=1}^{n} u_i\right)^{\frac{1}{n}}, \left(\prod_{i=1}^{n} u_i\right)^{\frac{1}{n}}, \ldots, \left(\prod_{i=1}^{n} u_i\right)^{\frac{1}{n}}\right)$.

  It follows that if $u, v \in R^n_{++}$ such that $\prod_{i=1}^{n} u_i = \prod_{i=1}^{n} v_i$, then $F(u) = F(v)$. Hence, $\exists$ a function ϕ, such that $\forall u \in R^n_{++}$, $F(u) = \phi\left(\prod_{i=1}^{n} u_i\right)$.
- *Step 2:* Now it will be shown that F is a negative monotone transformation of the product of the arguments. That is $\exists \phi : R \to [0,1]$ such that $x, y \in R$, $x > y$ implies $\phi(x) < \phi(y)$. Let R, $R' \in \Delta$ be such that $R = (r_1, r_2, \ldots, r_n)$ and $R' = (r'_1, r'_2, \ldots, r'_n)$, $r'_1 > r_1$ and $r'_i = r_i$, $\forall i \in \{2, 3, \ldots, n\}$. Then by axiom O, $M(R') > M(R)$. Now, by axiom R, in particular, that u_i is a decreasing function for all i, we know that $u_1(r_1) \prod_{i=2}^{n} u_i(r_i) > u_1(r'_1) \prod_{i=2}^{n} u_i(r_i)$. Furthermore, again by axiom R, we know

  $$M(R) = F(u_1(r_1), u_2(r_2), \ldots, u_n(r_n)) = \phi\left(u_1(r_1) \prod_{i=2}^{n} u_i(r_i)\right),$$
  $$M(R') = \phi\left(u_1(r'_1) \prod_{i=2}^{n} u_i(r_i)\right).$$

  Since $M(R') > M(R)$, it follows that ϕ is a decreasing function.
- Thus with these two steps we know that there exists a decreasing function ϕ such that $\forall R = (r_1, r_2, \ldots, r_n)$

  $$M(R) = \phi\left(\prod_{i=1}^{n} u_i(r_i)\right). \tag{4}$$

- *Step 3:* Now it will be shown that there is no loss of generality by requiring that every person's utility function is identical. Consider $R = (r^*, r_2, \ldots, r_n) \in \Delta$. By axiom A, $M(r^*, r_2, \ldots, r_n) = M(r_2, r^*, \ldots, r_n)$. Hence (4) implies that

$$M(r^*, r_2, \ldots, r_n) = M(r_2, r^*, \ldots, r_n),$$
$$\phi\left(u_1(r^*)\, u_2(r_2) \prod_{i=3}^{n} u_i(r_i)\right) = \phi\left(u_1(r_2)\, u_2(r^*) \prod_{i=3}^{n} u_i(r_i)\right)$$

and, since ϕ has been shown to be a decreasing function, we have

$$u_1(r^*)\, u_2(r_2) \prod_{i=3}^{n} u_i(r_i) = u_1(r_2)\, u_2(r^*) \prod_{i=3}^{n} u_i(r_i),$$
$$\text{which implies } u_1(r^*)\, u_2(r_2) = u_1(r_2)\, u_2(r^*),$$
$$u_2(r_2) = \frac{u_1(r_2)\, u_2(r^*)}{u_1(r^*)}.$$

Likewise, using the same argument as above we have that $u_j(r_j) = u_1(r_j)\frac{u_j(r^*)}{u_1(r^*)}$, $\forall j = 1, \ldots, n$. Therefore, $\forall (r_1, r_2, \ldots, r_n) \in \Delta$, $\prod_{i=1}^{n} u_i(r_i) = \theta \prod_{i=1}^{n} u_i(r^*)$, where $\theta \equiv \frac{\prod_{i=1}^{n} u_i(r^*)}{u_1(r^*)^n} > 0$. It follows that if there is a decreasing function ϕ satisfying (4), there must exist a decreasing function Ψ, such that $\forall R = (r_1, r_2, \ldots, r_n) \in \Delta$, $M(R) = \Psi\left(\prod_{i=1}^{n} u_1(r_i)\right)$.

For simplicity, we write $u(r_i)$ for $u_1(r_i)$ so we have

$$M(R) = \Psi\left(\prod_{i=1}^{n} u(r_i)\right). \tag{5}$$

- *Step 4:* We now complete the proof. By axiom C, we know that $\forall r \in [0,1]$, $r = \Psi(u(r_i)^n)$. If we write $x \equiv u(r_i)^n$, then $\Psi(x) = u^{-1}\left(x^{\frac{1}{n}}\right)$. By axiom R (i), we can write $u(r) = A - Br$, where $B > 0$. Hence, $\Psi(x) = \frac{A}{B} - \frac{x^{\frac{1}{n}}}{B}$. Therefore, by using (5) we have

$$M(r_1, r_2, \ldots, r_n) = \frac{A}{B} - \frac{1}{B}\left[\prod_{i=1}^{n} (A - Br_i)\right]^{\frac{1}{n}},$$
$$= \frac{A}{B} - \left[\prod_{i=1}^{n} \left(\frac{A}{B} - r_i\right)\right]^{\frac{1}{n}}.$$

By writing β for $\frac{B}{A}$, we have

$$M(r_1, r_2, \ldots, r_n) = \frac{1}{\beta} - \left[\prod_{i=1}^{n} \left(\frac{1}{\beta} - r_i\right)\right]^{\frac{1}{n}}.$$

Since $u : [0,1] \to R_{++}$, then $u(1) = [A - B \cdot 1] > 0$, which implies $A > B$. Therefore $1 > \beta$. Since, in addition $B > 0$ (by axiom R(i)), then $\beta > 0$. What we have left therefore is precisely the class of MOUs described in (2). □

One advantage of a full axiomatization of the kind just undertaken is that it helps us evaluate the measure by factorizing it to its constituents. In this case the strong assumption is clearly axiom R. This requires individual utility to be cardinal but does not impose interpersonal comparability. This kind of an axiom is used to derive the Nash bargaining solution and is also widely used by Sen in social choice theory (see Sen (1974, 1977) for instance). What may appear more contentious is the requirement that u_i be affine.

Some may treat this as reason to look for a different measure of unemployment, but there are two points worth keeping in mind. First, there are alternate ways of axiomatizing the same measure. So there may be other ways of visualizing our measure that do not require one to use an affine utility function as an input.

Second, we must not think of the utility function of each person, u_i, as the person's own evaluation of her utility. Instead, it should be viewed as society's evaluation of a person's employment status, which may well be different from the person's own utility evaluation (this is elaborated upon further in the next section). Once we take this approach and note that there are two steps to get to a final measure – (i) the assessment of each person's utility, u_i, and (ii) aggregation of these using a function, F – it becomes evident that the concavity of F acts as a substitute for diminishing marginal utility of the individual and that is the route we are taking here.

Moreover, our approach has some natural interpretational advantages. Consider person i's utility function: $u_i = A - Br_i$. Let Δu_i be the change in this person's utility if her status changed from fully unemployed ($r_i = 1$) to fully employed ($r_i = 0$). Clearly $\Delta u_i = B$. Now let $\widetilde{u}_i$ be this person's reservation utility, meaning the utility this person gets if she is without any work ($r_i = 1$). Clearly $\widetilde{u}_i = A - B$, that is, a person without work has a utility of $A - B$. Hence, the ratio of the utility from other things (i.e., other than work) to utility from being able to work is given by $\frac{\widetilde{u}_i}{\Delta u_i} = \frac{A}{B} - 1$.

Hence, an increase in A denotes how the other things in life are more important than work. An increase in A is thus associated with moving to a society where there is reasonable social welfare and other sources of income (for instance, through equity ownership) or where work is not as much a source of a person's social recognition.[4] Now note that since $\beta = \frac{B}{A}$, an increase in A is equivalent to β going towards zero. This, as we have already seen, pushes us towards the additive, or utilitarian, case where egalitarianism in unemployment matters less in our MOU given by (2). Likewise, as A becomes smaller, β becomes larger. In the limit, employment achieves enormous importance and our MOU converges towards a multiplicative evaluation.

[4] The "social" cost of unemployment does not always get its due. But it is arguable that once our basic economic needs are satisfied, loss of face becomes a dominant cost of unemployment (see Sen (1997) for instance).

5 Discussion

Given the full characterization of the effective measure of unemployment, it is now useful to draw out some of the distinctions between the existing literature and our paper. First note that the bulk of the existing writing – the theoretical Banerjee (2000), Ligon and Schechter (2003) and the empirical Kamanou and Morduch (2002) – mainly focuses on isolating individual vulnerability. It asks questions like "Who is vulnerable to becoming unemployed?" and "How do we estimate the number of vulnerable individuals?"

Our interest, on the other hand, is to assume that we know how many vulnerable or potentially unemployed persons there are in society and then to develop an **aggregate** measure of **effective** unemployment, that is, to find a single number that captures the total unemployment – actual and potential. Among the few papers that share our concern with the aggregate are Borooah (2002); Paul (1992); Shorrocks (1992, 1994).[5] Borooah (2002) develops a measure, drawing on the work of Atkinson (1970, 2005) in which aggregate, effective unemployment is derived from an aggregation of separable individual utilities.

Our measure charts out a different course based on a rejection of this separability. Take a look at our proposed MOU again. Recall, $M^{\beta}(R) = \frac{1}{\beta} - \prod_{i=1}^{n} \left(\frac{1}{\beta} - r_i\right)^{\frac{1}{n}}$ and let us examine society's view of one person's unemployment load, or pain – as referred to by Borooah. Using r_1 as an example, we can see that $\frac{\partial M^{\beta}(R)}{\partial r_1} = \frac{1}{n\left(\frac{1}{\beta} - r_1\right)} \left[\frac{1}{\beta} - M^{\beta}(R)\right]$ depends on the total effective unemployment as measured by $M^{\beta}(R)$. Hence, if total unemployment is higher, then $\frac{\partial M^{\beta}(R)}{\partial r_1}$ is lower. Therefore "the level of pain" that society associates with person i's unemployment depends on the level of effective unemployment in society. This essential relativity is not there in Borooah's measure.

Further, this paper takes the view that concepts like unemployment and even inequality cannot be reduced to pure welfarism. These are concepts that cannot be located entirely in the welfares of individuals and their aggregation. The same distinction that Sen (as in Sen (1976)) drew between ethical and descriptive features of inequality arise here in the context of unemployment.

We take the view that a greater amount of aggregate unemployment must not be **equated** with diminished aggregate social welfare.[6] This leads to an important difference between our approach and that of much of the literature on vulnerability

[5] Interestingly, Shorrocks (1994) is among the few papers that, like ours, uses an axiomatic approach, though the measure that he develops is different from ours.

[6] Some economists would go even further and argue that a small amount of unemployment may reflect flexibility in the labor market and so be good for the economy overall. This is not to deny that there may be a mathematical isomorphism between the welfarist approach and our approach. This is evident from Theorem 1 if we interpret u_i's as each person's own evaluation of her utility and think of F as a welfare function. But such interpretations are not necessary and indeed we would resist them here.

that uses the concept of "certainty-equivalence" to evaluate vulnerability (see, for example, Ligon and Schechter (2003)). To see the difference consider two states of employment: employed in a good job and employed in a bad job. We describe these as, respectively, "well employed" and "barely employed." Now consider a person facing the following job prospect: she will be unemployed on average 6 months in a year (because in each month the probability of being unemployed is $\frac{1}{2}$), and in the other 6 months she will be well employed. Viewed over the full year, should she be counted as employed or not? According to the certainty-equivalence approach, we simply have to ask this person if she would prefer to change her position with that of another person who will be barely employed with certainty for the entire year. If she says no, then this vulnerable person is presumably effectively employed.

This sounds like a reasonable exercise if our interest is in welfare. But it is clear that the enormous literature on unemployment measurement rejects such welfarism.[7] This is evident in, for instance, the work of Calvo and Dercon (please refer to Calvo and Dercon (2005)), since, as their "risk-sensitivity axioms" makes evident, in their approach it is only (and rightly so in our view) the downside risk that matters.

To understand this consider a society, x, with 12 persons, of whom six are unemployed and six are each well employed. Now transport all these 12 persons to a society, y, where they are all barely employed. Give each of them the choice of being born into society x without saying which position she will have. Let us say the probability that she will be unemployed is $\frac{1}{2}$ and the probability that she will be well employed is also $\frac{1}{2}$.

It is entirely possible that all 12 persons prefer society x to society y. Hence, in an *ex ante* sense x Pareto dominates y. Since there is no unemployment in y and everybody prefers x to y, if we were equating unemployment totally with welfare, we would be forced to say there is no unemployment in x. But that would be absurd and indeed with six unemployed people in this society at all times, no one would say that x has no unemployment.

Hence, in developing an aggregate measure of unemployment (treating this as description of society), we may be justified in rejecting the welfarism inherent in the certainty-equivalence approach.

6 Conclusion

We have offered an alternative way to look at vulnerability to what is currently being discussed in the literature and by policy makers. That is, vulnerability, when viewed as a part of an aggregate measure, need not always be treated as a "bad." Given this

[7] For one, how one comes to be unemployed may matter, which would immediately take us in the direction of the procedural approach discussed lucidly by Suzumura in Suzumura (1999). But we would go even further and argue that the measure of aggregate unemployment is distinct from welfarism and procedures. Moreover, the relation between aggregate unemployment and aggregate welfare need not invariably be positive monotonic.

perspective, we have provided a way of measuring "effective" unemployment. This measure is bounded on one side by the additive or utilitarian measure and on the other side by a multiplicative measure. Furthermore, our measure satisfies axioms that most people would agree are what one would want from a measure motivated by equity concerns. We have fully characterized our measure and shown how the measure can be applied to data in both the US and South Africa and what insights can be gained by comparing the usual and effective measures of unemployment. This paper then serves two purposes. First, it suggests that the current debate on vulnerability needs to examine not only the effect of vulnerability on people currently employed, but also the *hope* that vulnerability provides to people who are currently unemployed. Second, this paper provides a way of taking account of these concerns in a single measure of unemployment and shows how the measure can actually be put to use.

Acknowledgements The authors thank Kuntal Banerjee, Rajat Deb, Peter Hammond, Magnus Hatlebakk, Nanak Kakwani, Ethan Ligon, Francesca Molinari, Felix Naschold, Prasanta Pattanaik, Dorrit Posel, Furio Rosati, Karl Shell, Tony Shorrocks, S. Subramanian, Kotaro Suzumura, Erik Thorbecke, and the participants of the Cornell/PennState Macroeconomics Workshop, "The Conference on 75 Years of Development Research" at Cornell University, the 9th Australasian Macroeconomics Worshop, the "International Conference on Rational Choice, Individual Rights, and Non-Welfaristic Normative Economics" at Hitotsubashi University, Tokyo, March 2006, and the TWIPS seminar at Cornell University. We also benefited greatly from the comments and suggestions of an anonymous referee.

References

Akerlof, G. A., and Main, B. G. M. (1980). Unemployment spells and unemployment experience. *American Economic Review*, *70*, 885–893

Amin, S., Rai, A. S., and Topa, G. (2003). Does microcredit reach the poor and vulnerable? Evidence from Northern Bangladesh. *Journal of Development Economics*, *70*, 59–82

Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*, 244–263

Atkinson, A. (1983). *Social justice and public policy*. Cambridge, MA: WiMIT Press

Banerjee, A. (2000). The two poverties. *Nordic Journal of Political Economy*, *26*, 129–141

Basu, K., and Foster, J. (1998). On measuring literacy. *Economic Journal*, *108*, 1733–1749

Borooah, V. K. (2002). A duration-sensitive measure of the unemployment rate: Theory and application. *Labour*, *16*, 453–468

Calvo, C., and Dercon, S. (2005). *Measuring individual vulnerability* (Discussion Paper No. 229). Oxford: Department of Economics, Oxford University

Clark, K. B., and Summers, L. H. (1979). Labor market dynamics and unemployment: A reconsideration. *Brookings Papers on Economic Activity*, *1979*, 13–72

Cunningham, W., and Maloney, W. F. (2000). *Measuring vulnerability: Who suffered in the 1995 Mexican crisis?* (mimeo). IBRD

Dercon, S. (2005). *Vulnerability: A micro perspective*. In: Paper presented at the ABCDE Europe conference. Amsterdam, May 2005

Fields, G. S. (1996). The meaning and measurement of income mobility of income mobility. *Journal of Economic Theory*, *71*, 349–377

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, *52*, 761–766

Glewwe, P., and Hall, G. (1998). Are some groups more vulnerable to macroeconomic shocks than others? Hypothesis tests based on panel data from Peru. *Journal of Development Economics*, *56*, 181–206

Grootaert, C., and Kanbur, R. (1995). The lucky few amidst economic decline: Distributional change in Cote d'Ivoire as seen through panel datasets, 1985–1988. *Journal of Development Economics*, *31*, 603–619

Kamanou, G., and Morduch, J. (2002). *Vulnerability: A micro perspective* (Discussion Paper No. 2002/58). WIDER

Ligon, E., and Schechter, L. (2003). Measuring vulnerability. *Economic Journal*, *113*, C95–C102

Paul, S. (1992). An illfare approach to the measurement of unemployment. *Applied Economics*, *24*, 739–743

Pritchett, L., Suryahadi, A., and Sumarto, S. (2000). *Quantifying vulnerability to poverty: A proposed measure, with applications to Indonesia* (Working Paper). SMERU

Sen, A. (1974). Informational bases of alternative welfare approaches. *Journal of Public Economics*, *3*, 387–403

Sen, A. (1976). Poverty: An ordinal approach to measurement. *Econometrica*, *44*, 219–231

Sen, A. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica*, *45*, 1539–1572

Sen, A. (1997). *The penalties of unemployment* (Rome Paper No. 307). San Francisco, CA: Bank of Italy

Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, *46*, 1013–1024

Shorrocks, A. F. (1992). *Spell incidence, spell duration and the measurement of unemployment* (Discussion Paper 409). Essex, UK: Department of Economics, University of Essex

Shorrocks, A. F. (1994). *On the measurement of unemployment* (Discussion Paper 418). Department of Economics, University of Essex

Suzumura, K. (1999). Consequences, opportunities, and procedures. *Social Choice and Welfare*, *16*, 17–40

Thorbecke, E. (2003). *Conceptual and measurement issues in poverty analysis* (Paper prepared for the Second Nordic Conference in Development Economics (NCDE-2) Copenhagen). Denmark, June 23–24

World Bank (2002). *Pakistan poverty assessment. Poverty in Pakistan: Vulnerabilities, social gaps and rural dynamics* (Report No. 24296-PAK). Washington, DC: World Bank

Ordinal Distance, Dominance, and the Measurement of Diversity

Prasanta K. Pattanaik and Yongsheng Xu

1 Introduction

The purpose of this chapter is to consider a class of rules for comparing sets of objects[1] in terms of the degrees of diversity that they offer. Such comparisons of sets are important for many purposes. For example, in discussing biodiversity of different ecosystems, one is interested in knowing whether or not one ecosystem is more diverse than another. Similarly, when discussing issues relating to cultural diversities of various communities, one may be interested in knowing how these communities compare with each other in terms of cultural diversity. In the economics literature, there have been several contributions to the measurement of diversity. Weitzman (1992, 1993, 1998) develops a measure of diversity based on cardinal distances between objects. Among other things, Nehring and Puppe (2002) provide a conceptual foundation for cardinal distances in Weitzman's framework. Weikard (2002) discusses an alternative measure of diversity; Weikard's measure is based on the sum of cardinal distances between all objects contained in a set.

Underlying much of our everyday discussion of diversity, we have some intuition regarding the extent to which objects are dissimilar,[2] though, in its coarsest form, this intuition may distinguish between only two degrees of similarity by declaring that two objects are either similar or dissimilar. It is difficult to see how one can compare the diversity of one group of objects with that of another without some

P.K. Pattanaik
Department of Economics, University of California, Riverside, CA 92521, USA
e-mail: ppat@ucrac1.ucr.edu

Y. Xu
Department of Economics, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA 30303, USA
e-mail: yxu3@gsu.edu

[1] We use the term "objects" rather broadly so that people and animals can also be objects.

[2] The basis for assessing dissimilarity or similarity of two objects will, of course, vary depending on the specific notion of diversity in which one is interested.

P.K. Pattanaik et al. (eds.) *Rational Choice and Social Welfare: Theory and Applications*, Studies in Choice and Welfare.

notion, however coarse, of the distances or the degrees of dissimilarity between different objects in these sets. The requirement that distances be cardinally measurable does, however, seem rather restrictive in many contexts. Thus, in considering linguistic diversity, we may not be able to compare the extent to which the (linguistic) dissimilarity between an English person and a Chinese person exceeds the difference between a Chinese person and a Hindi-speaking person, on the one hand, and the extent to which the difference between an Italian person and a Hindi-speaking person exceeds the dissimilarity between a Spanish-speaking person and a Hindi-speaking person, on the other. This is not to claim that cardinal distance functions never have sound intuitive foundations. In general, however, the requirement of a cardinal distance function for the measurement of diversity seems quite strong. It is, therefore, of considerable interest to see how far one can proceed in measuring diversity on the basis of an ordinal distance function.

The existing literature has some contributions on the measurement of diversity based on ordinal distance. Pattanaik and Xu (2000) introduce a coarse ordinal distance function to measure the diversity of different sets: under their distance function, any two objects are either similar or dissimilar. Bossert, Pattanaik, and Xu (2003) elaborate further on ordinal distances between objects. Bervoets and Gravel (2007) also use ordinal distances to provide some rules for ranking sets of objects in terms of diversity; these rules focus on the objects that are most dissimilar in a set. In this chapter, we use an ordinal distance function to develop a notion of domination between sets, to characterize the class of all rules for ranking sets in terms of diversity that satisfy the property of dominance, and to characterize a specific ranking rule belonging to this class.

The chapter is organized as follows. In Section 2, we present the basic notation and definitions of our analysis. Section 3 introduces our central concepts of weak domination and domination between sets. In Section 4, we provide a characterization of the class of all rankings of sets of objects, which satisfy the property of dominance. In this section, we also present a characterization of a distinguished member of this class. Section 5 considers some related analytical frameworks. Section 6 contains some brief concluding remarks.

2 Notation and Definitions

Let X be the universal set of objects; X is assumed to contain a finite number of elements. It is clear that the interpretation of the objects in X will depend on the specific context in which one is interested in the notion of diversity. Thus, if one is interested in changes in biodiversity over time in a given region of China, then X can be the set of all animals that have been known to exist in that region of China at different points of time over that period (note that, for the purpose of this interpretation, we would consider, say, two different Chinese tigers as two distinct animals in the universal set). If one is interested in comparing the degrees of linguistic diversity in different countries, then X may be the set of all people living in all these countries (again, two different persons speaking exactly the same language will be considered to be two distinct elements of the universal set). The formal structure of the problem of measuring diversity as we consider it in this chapter does not, however, depend on the specific type of diversity in which we may be interested or on any specific interpretation of the universal set.

Let $\mathcal{K}$ be the class of all nonempty subsets of X and let $\mathcal{K}_2$ be the class of all sets A in $\mathcal{K}$ such that $\#A \leq 2$. Our problem is one of ranking the different subsets, $A, B, \ldots$ of $\mathcal{K}$ in terms of the degrees of diversity that they offer. For example, does a forest with 24 tigers, 22 bears, 2003 deer, and 5000 rabbits offer more biodiversity than a forest with 14 tigers, 40 bears, 4500 deer, and no rabbits? To analyze this type of questions, let $\succeq$ be a binary relation over $\mathcal{K}$; for all $A, B \in \mathcal{K}$, $A \succeq B$ means that the set A is at least as diverse as the set B. The symmetric and asymmetric parts of $\succeq$ are denoted, respectively, by $\sim$ and $\succ$. $\succeq$ is assumed to be a quasi-ordering, that is, it is assumed to be reflexive and transitive, but not necessarily connected. Much of this chapter will be concerned with the structure of the binary relation $\succeq$.

3 Ordinal Distance Function, Indistinguishable Objects, and Dominance

In analyzing the structure of $\succeq$, our starting point will be that of an ordinal distance function. Suppose we are interested in biodiversity so that our universal set is a set of animals, and suppose, on the basis of some criteria, we believe that an elephant and a panther are more dissimilar than a panther and a leopard. The notion of an ordinal distance function is intended to capture the comparison of the dissimilarity or distance between any two such elements in the set under consideration with the dis-similarity or distance between two other elements in the set. More formally, an *ordinal distance function* is a function $d : X \times X \to [0, \infty)$ such that

$$(a, a) = 0 \text{ for all } a \in X \tag{1}$$

and

$$d(a, b) = d(b, a) \text{ for all } a, b \in X. \tag{2}$$

The function d has the following interpretation: for all $x, y, z, w \in X$, $d(x, y) > d(z, w)$ denotes that the degree of dissimilarity between x and y is greater than the degree of dissimilarity between z and w, and $d(x, y) = d(z, w)$ denotes that the degree of dissimilarity between x and y is the same as the degree of dissimilarity between z and w. As the name "ordinal distance function" suggests, we do not attach any meaning to the comparison of $d(x, y) - d(z, w)$ and $d(x', y') - d(z', w')$ for any $x, y, z, w, x', y', z', w' \in X$. Equation (1) says that the degree of dissimilarity between an object and itself is 0; this is simply a convention. Equation (2) requires that the degree of dissimilarity between an object x and an object y is the same as the degree of dissimilarity between y and x (so that the distance function is symmetric).

Several points may be noted about our ordinal distance function d. First, while, for the sake of convenience of exposition, we start with the distance function d, the intuitively primitive notion is really that of an ordering of dissimilarties between different alternatives, which underlies the distance function d. d can be thought of as a convenient real-valued representation of this ordering of dissimilarities (this, of course, involves the assumption that the ordering of dissimilarities between different alternatives is representable by a real-valued function). Second, we do not enquire into the criteria that constitute the basis of d, but it is clear that the criteria underlying d will be very different depending on the type of diversity (e.g., biodiversity, cultural diversity, linguistic diversity, and so on) under consideration. Third, the definition of the ordinal distance function d implies an implicit assumption, namely, that the dissimilarity or distance between any two objects can be compared with the dissimilarity or distance between any two other objects. This may be a strong assumption. It is possible that, sometimes in practice, we may not have such universal comparability of the dissimilarities involved. In Section 5 below, we indicate how one can relax the assumption of universal comparability of distances. Finally, our definition of d permits $d(x,y)$ to be 0 for distinct objects x and y in X.

For illustrating several subsequent concepts, we use the following example of a distance function over a set $\{x,y,z,w\}$.

Example 1. We represent a specific distance function d over the set $\{x,y,z,w\}$ in the form of a table, where $d(y,z)$ is the number (0) that figures in the row corresponding to y and the column corresponding to z, $d(w,z)$ is the number (0.8) that figures in the row corresponding to w and the column corresponding to z, and so on.

	x	y	z	w
x	0	0	0	0
y	0	0	0	0.8
z	0	0	0	0.8
w	0	0.8	0.8	0

(Note that the specification of the numbers for different ordered pairs of alternatives in the above table satisfies (1) and (2).)

Let I be a binary relation (*"indistinguishable, in relation to other alternatives, from"*) defined over X as follows: for all x, $y \in X$, xIy iff for all $z \in X$, $d(x,z) = d(y,z)$. I is clearly an equivalence relation. We say that x and y are *distinguishable in relation to other alternatives* iff not (xIy). In Example 1, y is indistinguishable, in relation to other alternatives, from z, while z and w are distinguishable, in relation to other alternatives. Note that, if xIy, then $d(y,x) = d(x,x) = 0$. In the absence of further restrictions on d, $d(x,y) = 0$ does not necessarily imply xIy. Consider, however, the following rather mild restriction on d:

$$\text{For all } x,y \in X, \text{ if } d(x,y) = 0, \text{ then } d(x,z) = d(y,z) \text{ for all } z \in X. \tag{3}$$

It is easy to check that, if d satisfies (3), then, for all $x,y \in X$, $[d(x,y) = 0]$ is equivalent to xIy. Though we believe that (3) is a relatively mild and "natural" property of an ordinal distance function, we do not need it for our results and we do not impose it on d.

An object $x_0 \in X$ will be said to be a *null object* iff $d(x_0, x) = 0$ for all $x \in X$. In Example 1, x is a null object. For all $A \in \mathcal{K}$, the set of all null objects in A will be denoted by A_0. We say that a set A is *heterogeneous* iff [A does not contain any null object, and, for all distinct $x, y \in A, d(x,y) \neq 0$]. In Example 1, $\{y, w\}$ is hetererogeneous, but $\{x\}$ is not heterogeneous. For all $A \in \mathcal{K}$ with $\#A \geq 2$, if $A = A_0$, then let $e(A) \equiv \{a\}$ for some $a \in A$; otherwise, partition $A - A_0$ into I-equivalence classes, $A_1, A_2, \ldots, A_{m(A)}$, and let $e(A)$ denote $\{a_1, a_2, \ldots, a_{m(A)}\}$, where for all $i \in \{1, 2, \ldots, m(A)\}$, a_i is some object (arbitrarily) chosen from A_i. In Example 1, $X_0 = \{x\}$ and $X - X_0 = \{y, z, w\}$ can be partitioned into the following I-equivalence classes: $X_1 = \{y, z\}$ and $X_2 = \{w\}$. One can then take $e(X)$ to be either $\{y, w\}$ or $\{z, w\}$.

Our task is to use the information contained in the ordinal distance function d to rank various subsets of X in terms of diversity. This is a rather complex exercise, to say the least. To make the maximum possible use of our initial intuition, we shall start with the notion of "weak domination."

Definition 1. For all $A, B \in \mathcal{K}$, we say that A *weakly dominates* B iff $\#e(A) \geq \#e(B)$ and there exist a subset $e'(A)$ of $e(A)$ and a one-to-one correspondence f between $e'(A)$ and $e(B)$ such that [for all $x, y \in e'(A), d(x,y) \geq d(f(x), f(y))$].

Definition 2. For all $A, B \in \mathcal{K}$, we say that A *dominates* B iff A weakly dominates B, but B does not weakly dominate A.

It can be easily checked that, in our Example 1, $\{z, w\}$ dominates $\{x, y, z\}$.

Several features of our notions of weak domination and domination may be noted. First, when $d(x,y) = 0$, but x and y are distinguishable, $\{x, y\}$ dominates $\{z\}$ for all $z \in X$: a set of two objects that have 0 distance from each other but are distinguishable dominates every singleton set. Second, if [$d(x,y) = d(y,z) = d(x,z) = 0$ but x, y, z are pairwise distinguishable] and [$d(a,b) > 0$], there does not exist any weak domination relation either way between $\{x, y, z\}$ and $\{a, b\}$. Third, for all $x, y, z, w \in X$, $\{x, y\}$ weakly dominates $\{z, w\}$ or $\{z, w\}$ weakly dominates $\{x, y\}$ (note that this is true irrespective of whether $x = y$ or $z = w$).

4 Ranking Rules Satisfying the Property of Dominance

Definition 3. We say that $\succeq$ satisfies *dominance* (D) iff, for all $A, B \in \mathcal{K}$, [if A weakly dominates B, then $A \succeq B$] and [if A dominates B, then $A \succ B$].

Note that D, by itself, does not identify a unique quasi-ordering over $\mathcal{K}$; instead, we have a class of quasi-orderings over $\mathcal{K}$, each of which satisfies D. It is easy to check that every quasi-ordering $\succeq$ over $\mathcal{K}$, which satisfies D, satisfies the following intuitively appealing properties:

(P.1) For all $x, y \in X, \{x\} \sim \{y\}$;

(P.2) For all $x, y, x', y' \in X$, if x and y are distinguishable and x' and y' are distinguishable, then $d(x,y) > d(x',y') \Rightarrow \{x,y\} \succ \{x',y'\}$;
(P.3) For all $A \in \mathcal{K}$ and all $x \in X$, if [x is indistinguishable from some $a \in A$ or x is a null object, then $A \cup \{x\} \sim A$,] and [if for every $a \in A, x$ is distinguishable from a and x is not a null object, then $A \cup \{x\} \succ A$];
(P.4) for all $A, B \in \mathcal{K}$ with $\#A = \#B$, for all $x \in X \setminus A$ and all $y \in X \setminus B$ such that both $A \cup \{x\}$ and $B \cup \{y\}$ are heterogeneous, if there exists a one-to-one mapping f from A to B such that $d(x,a) \geq d(y, f(a))$ for all $a \in A$, then $[A \succeq B \Rightarrow A \cup \{x\} \succeq B \cup \{y\}]$ and $[A \succ B \Rightarrow A \cup \{x\} \succ B \cup \{y\}]$.

(P.1) states that every singleton set $\{x\}$ is as diverse as every other singleton set $\{y\}$. This property has been used by many writers implicitly or explicitly in measuring diversity. (P.2) requires that the diversity of every doubleton set containing two distinguishable objects depends exclusively on the ordinal distance between the two objects. (P.3) stipulates that the addition of an object x to a set A will not change the degree of diversity already offered by A when x is indistinguishable from some existing object in A or x is a null object, while the addition of x to A will increase the degree of diversity offered by A when x is distinguishable from every object in A and x is not a null object. This property needs careful interpretation. In real life, people sometimes make the remark that an increase in the population of, say, giant pandas will increase biodiversity. Such remarks seem to go counter to the intuition of (P.3). The remark, however, seems to be based on the belief that there is a critical level of present population below which giant pandas have no reasonable chance of surviving in the future, that the current population level of giant pandas is below this critical level, and that, as a consequence, an increase in the population of giant pandas now will increase biodiversity by ensuring the survival of giant pandas. (P.3) seems to be a reasonable property when such intertemporal issues are considered in a framework analogous to the standard analytical framework in economics where we consider the same physically identifiable commodity available at two different points of time as two different commodities. Suppose, we have only one species, giant pandas, and two periods, "the present" (0) and "the future" (1). The number of giant pandas in the present is denoted by g_0, and the number of giant pandas in the future is denoted by g_1. Let the minimum number of giant pandas required in the present for its survival in the future be 100. Suppose g_0 is 60. This cannot ensure the survival of giant pandas in the future. Hence, we have a set of animals consisting of 60 giant pandas in the present and 0 giant pandas in the future. On the one hand, if g_0 increases to 80, then we shall have a set of animals consisting of 80 giant pandas in the present and none in the future. On the other hand, if g_0 increases to 100, then that will ensure the survival of giant pandas in the future and we would have a set of animals consisting of 100 giant pandas in the present and, say, 40 giant pandas in the future. If a giant panda in the future is considered to be "different" from a giant panda in the present, while two giant pandas in the present are considered "exactly similar," then it is not unintuitive to say that biodiversity will not increase in the first case but will increase in the second case. This is consistent with (P.3). Finally, (P.4) is a type of independence property.

Theorem 1. $\succeq$ *satisfies D if and only if it satisfies (P.1), (P.2), (P.3), and (P.4).*

Proof. The proof is given in the appendix. □

As we noted earlier, property D does not determine a unique rule for ranking sets in terms of diversity. Instead, it identifies a (non-singleton) class rules. The following definition introduces a rule, denoted by $\succeq^*$, which belongs to this class and is of some interest.

Definition 4. For all $A, B \in \mathcal{K}$, $A \succeq^* B$ if and only if A weakly dominates B.

It is clear that $\succeq^*$ satisfies D. The feature of $\succeq^*$ that distinguishes it from other rules satisfying D is that, for every pairs of sets in $\mathcal{K}$, if neither of the two sets weakly dominates the other, then they are declared noncomparable by $\succeq^*$.

Now, consider the following properties of a ranking $\succeq$ over $\mathcal{K}$.

(P.5) For all $A, B \in \mathcal{K}$ with $\#A > 1$ and $\#B > 1$, if $A \succeq B$, then $A \setminus \{a^*\} \succeq B \setminus \{b^*\}$ for some $a^* \in A$ and some $b^* \in B$.
(P.6) For all $A, B \in \mathcal{K}$, if both A and B are heterogeneous, $\#A = \#B \geq 2$, $A \succeq B$, and $A \setminus \{a'\} \succeq B \setminus \{b'\}$ for some $a' \in A$ and some $b' \in B$, then there exist $a^* \in A$ and $b^* \in B$ such that $A \setminus \{a^*\} \succeq B \setminus \{b^*\}$ and for some one-to-one correspondence f from $A \setminus \{a^*\}$ to $B \setminus \{b^*\}$, $d(a^*, a) \geq d(b^*, f(a))$ for all $a \in A \setminus \{a^*\}$.
(P.7) For all $A, B \in \mathcal{K}$, if both A and B are heterogeneous, $A \succeq B$, and $\#A > \#B$, then there exists a proper subset A' of A such that $A' \succeq B$.

Theorem 2. $\succeq = \succeq^*$ *if and only if* $\succeq$ *satisfies (P.1) through (P.7).*

Proof. The proof is given in the appendix. □

5 Revealed Ordinal Distance Functions and the Comparability of Distances Between Objects

In this section, we indicate some alternatives to the approach that we have adopted in the earlier sections.

So far, we have treated the ordinal diastance function d as a primitive concept in our framework and based our ranking of the sets in $\mathcal{K}$ on this exogenously given d. One can, however, follow an approach where the quasi-ordering $\succeq$ over $\mathcal{K}$ is the primitive concept and the ordinal distance function is defined in terms of $\succeq$. Let $\succeq$ ("at least as diverse as") be a given quasi-ordering over $\mathcal{K}$. Let $\succeq_2$ be a binary relation over $\mathcal{K}_2$ such that for all $A, B \in \mathcal{K}$, $A \succeq_2 B$ iff $A \succeq B$. Assume that

$$\forall x, y \in X, \{x, y\} \succeq \{x\} \sim \{y\} \tag{4}$$

and

$$\forall x, y, z, w \in X, \{x, y\} \succeq \{z, w\} \text{ or } \{z, w\} \succeq \{x, y\}]. \quad (5)$$

Since $\succeq$ is assumed to be a quasi-ordering over $\mathcal{K}$ (i.e., $\succeq$ is reflexive and trasitive over $\mathcal{K}$), (5) implies that $\succeq_2$ is an ordering over $\mathcal{K}_2$ (i.e., $\succeq_2$ satisfies reflexivity, connectedness, and transitivity over $\mathcal{K}_2$). Define a function $d' : X^2 \Rightarrow R_+$ such that

$$\forall (x, y), (z, w) \in X^2, d'(x, y) \geq d'(z, w) \Leftrightarrow \{x, y\} \succeq \{z, w\} \quad (6)$$

and

$$\forall x, y \in X, d'(x, y) = d'(y, x) \text{ and } d'(x, x) = 0. \quad (7)$$

It can be easily checked that, given the finiteness of X and given (4), such a real-valued function d' can be found. One can then treat this function d' as an ordinal distance function which is "revealed" by $\succeq$. Our concepts of weak domination and domination of sets and the property of dominance can now be developed in terms of this revealed ordinal distance function, d', and the property of dominance for $\succeq$, in its turn, can be defined in terms of these newly defined relations of weak domination and domination between sets. One can then prove the counterparts of Theorems 1 and 2 in this framework, the proofs being exactly analogous to the proofs of Theorems 1 and 2, respectively.

It may be worth noting some possible extensions of Theorems 1 and 2. The ordinal distance function introduced in Section 3 implicitly assumes that, for all $x, y, z, w \in X$, the extent of dissimilarity between x and y can be compared with the extent of dissimilarity between z and w. This intuitive assumption regarding the comparability of dissimilarities between objects is also inherent in all approaches based on cardinal distance functions. The assumption of universal comparability of dissimilarities may, however, be considered rather strong in some contexts. It is, therefore, of interest to note that results analogous to our Theorems 1 and 2 can be proved without this intuitive assumption. To do this, we can start with a given reflexive and transitive, but not necessarily connected, binary relation $\trianglerighteq$ defined over X^2 such that, for all $x, y \in X$, $(x, y) \trianglerighteq (y, x) \trianglerighteq (x, x)$. For all $x, y, z, w \in X$, $(x, y) \trianglerighteq (z, w)$ means that the distance between x and y is at least as great as the distance between z and w. It is possible to develop the counterparts of Theorems 1 and 2 using the binary relation $\trianglerighteq$ instead of the real-valued ordinal distance function d, but we do not undertake the exercise here since the reasoning involved is very similar to the reasoning in Section 4.

6 Concluding Remarks

In this chapter, we have used an ordinal concept of distance between objects to provide a characterization of all diversity-based rankings of sets of objects that satisfy a plausible property of dominance; we have also characterized a specific member of this class, which declares two sets of objects to be noncomparable if they are not comparable in terms of the relation of weak domination. We have discussed a parallel approach where we rely on the "revealed ordinal distances" rather on an exogenously given ordinal distance function. Our results constitute only the first step in the exploration of diversity-based rankings of sets of objects, using ordinal distance functions. The class of diversity-based rankings that satisfy dominance is a very wide class. Can we narrow down this class by imposing other reasonable properties in addition to dominance, but requiring only information about ordinal distances? This "natural" extension of our analysis requires a separate study.

Acknowledgments For helpful comments, we are grateful to Nick Baigent, Rajat Deb, Mark Fleurbaey, Peter Hammond, Clemens Puppe, Kotaro Suzumura, and other participants in the Conference on Rational Choice, Individual Rights, and Non-Welfaristic Normative Economics held in Tokyo in March 2006.

Appendix

Proof of Theorem 1

Suppose that $\succsim$ satisfies (P.1), (P.2), (P.3), and (P.4). We show that $\succsim$ is a dominance-based rule.

Let $\succsim$ satisfy (P.1), (P.2), (P.3), and (P.4). By the repeated use of (P.3) and transitivity of $\succsim$, it is straightforward to show that $A \sim e(A)$ for every $A \in \mathcal{K}$.

We first show that, for all $A, B \in \mathcal{K}$, if $\#e(A) = \#e(B)$ and there is a one-to-one correspondence f between $e(A)$ and $e(B)$ such that [for all $x, y \in e(A), d(x,y) = d(f(x), f(y))$], then $A \sim B$. Let $A, B \in \mathcal{K}$ be such that $\#e(A) = \#e(B)$ and, for some one-to-one correspondence f between $e(A)$ and $e(B)$, we have [for all $x, y \in e(A), d(x,y) = d(f(x), f(y))$]. Let $e(A) = \{a_1, \ldots, a_m\}$ and $e(B) = \{b_1, \ldots, b_m\}$ be such that $f(a_i) = b_i$ for $i = 1, \ldots, m$. Then, $d(a_i, a_j) = d(b_i, b_j)$ for all $i, j = 1, \ldots, m$. By (P.1), $\{a_1\} \sim \{b_1\}$. By (P.4) and noting that $d(a_1, a_2) = d(b_1, b_2)$, it follows that $\{a_1, a_2\} \sim \{b_1, b_2\}$. By (P.4) and noting that $d(a_3, a_2) = d(b_3, b_2)$ and $d(a_3, a_1) = d(b_3, b_1)$, we obtain $\{a_1, a_2, a_3\} \sim \{b_1, b_2, b_3\}$. By the repeated use of (P.4) if necessary, and noting that $d(a_i, a_j) = d(b_i, b_j)$ for all $i, j = 1, \ldots, m$, we have $\{a_1, \ldots, a_m\} \sim \{b_1, \ldots, b_m\}$. That is, $e(A) \sim e(B)$. Noting [$A \sim e(A)$ and $B \sim e(B)$], the transitivity of $\succsim$ then gives us $A \sim B$.

Next, we show that, for all $A, B \in \mathcal{K}$, if $\#e(A) = \#e(B)$ and there is a one-to-one correspondence f between $e(A)$ and $e(B)$ such that [for all $x, y \in e(A), d(x,y) \geq d(f(x), f(y))$] and [for some $x, y \in e(A), d(x,y) > d(f(x), f(y))$], then $A \succ B$. Let $A, B \in \mathcal{K}$ be such that $\#e(A) = \#e(B)$ and, for some one-to-one correspondence f between $e(A)$ and $e(B)$, we have [for all $x, y \in e(A), d(x,y) \geq d(f(x), f(y))$] and [for some $x, y \in e(A), d(x,y) > d(f(x), f(y))$]. Again, let $e(A) = \{a_1, \ldots, a_m\}$ and $e(B) = \{b_1, \ldots, b_m\}$ be such that $f(a_i) = b_i$ for $i = 1, \ldots, m$. Then, $d(a_i, a_j) \geq d(b_i, b_j)$ for all $i, j = 1, \ldots, m$, and for some $h, k = 1, \ldots, m$, $d(a_h, a_k) > d(b_h, b_k)$. Without loss of generality, let $h = 1$ and $k = 2$. Then, $d(a_1, a_2) > d(b_1, b_2)$. By

(P.2), noting that $d(a_1,a_2) > d(b_1,b_2)$, $\{a_1,a_2\} \succ \{b_1,b_2\}$ follows immediately. By (P.4) and noting that $d(a_3,a_1) \geq d(b_3,b_1)$, $d(a_3,a_2) \geq d(b_3,b_2)$, it follows that $\{a_1,a_2,a_3\} \succ \{b_1,b_2,b_3\}$. By the repeated use of (P.4) if necessary, and noting that $d(a_i,a_j) \geq d(b_i,b_j)$ for all $i,j = 1,\ldots,m$, we obtain $\{a_1,\ldots,a_m\} \succ \{b_1,\ldots,b_m\}$. That is, $e(A) \succ e(B)$. By the transitivity of $\succeq$, $A \succ B$ follows from $e(A) \sim A$ and $e(B) \sim B$.

Finally, let $A,B \in \mathcal{K}$, let $\#e(A) > \#e(B)$, and let there be a subset $e'(A)$ of $e(A)$ and a one-to-one correspondence f between $e'(A)$ and $e(B)$ such that [for all $x,y \in e'(A), d(x,y) \geq d(f(x),f(y))$]. We show that $A \succ B$. Let $e'(A) = \{a_1,\ldots,a_m\}$, $e(A) = \{a_1,\ldots,a_m,a_{m+1},\ldots,a_{m+n}\}$, and $e(B) = \{b_1,\ldots,b_m\}$ be such that $f(a_i) = b_i$ for $i = 1,\ldots,m$ and $n \geq 1$. From the above analysis, we have $e'(A) \succeq e(B)$. By (P.3), it follows that $e'(A) \cup \{a_{m+1}\} \succ e'(A)$, $e'(A) \cup \{a_{m+1}\} \cup \{a_{m+2}\} \succ e'(A) \cup \{a_{m+1}\}$, $\ldots$, $e(A) = e'(A) \cup \{a_{m+1}\} \cup \cdots \cup \{a_{m+n-1}\} \cup \{a_{m+n}\} \succ e'(A) \cup \{a_{m+1}\} \cup \cdots \cup \{a_{m+n-1}\}$. By the transitivity of $\succeq$, it follows that $e(A) \succ e'(A)$. Another application of transitivity of $\succeq$ yields $e(A) \succ e(B)$. Now, $A \succ B$ follows from the transitivity of $\succeq$ by noting that $A \sim e(A)$ and $B \sim e(B)$.

To complete the proof of Theorem 1, we note that it is straightforward to check that every $\succeq$ that satisfies D must satisfy (P.1), (P.2), (P.3), and (P.4). □

Proof of Theorem 2

It can be verified that $\succeq^*$ satisfies (P.1) through (P.7). In what follows, we show that, if $\succeq$ satisfies (P.1) through (P.7), then $\succeq = \succeq^*$.

Let $\succeq$ satisfy (P.1) through (P.7). We first note that, by Theorem 1, $A \sim e(A)$ for all $A \in \mathcal{K}$.

Consider any $A,B \in \mathcal{K}$. If either of the two sets, A and B, weakly dominates the other, then, by Theorem 1 and the definition of $\succeq^*$, it is clear that $A \succeq B$ iff $A \succeq^* B$ and $B \succeq A$ iff $B \succeq^* A$. To complete the proof, therefore, we need only to show that, if neither of the two sets, A and B, weakly dominates the other, then A and B are noncomparable. Assume that neither A weakly dominates B nor B weakly dominates A. Given this, it can be checked that $\#e(A) \geq 2$ and $\#e(B) \geq 2$.

There are several cases that need to be considered. First, we note that, if $\#e(A) < \#e(B)$, then we must have not$[e(A) \succeq e(B)]$. This is because, if $e(A) \succeq e(B)$, by (P.5), we obtain $e(A) \setminus \{a_1\} \succeq B \setminus \{b_1\}$ for some $a_1 \in e(A)$ and some $b_1 \in e(B)$. If $e(A) \setminus \{a_1\}$ contains one object, then there is an immediate contradiction with $e(B) \setminus \{b_1\} \succ e(A) \setminus \{a_1\}$, given Theorem 1. If $e(A) \setminus \{a_1\}$ contains more than one element, by repeated application of (P.5), we obtain a similar contradiction. Therefore, when $\#e(A) < \#e(B)$, it must be true that not$[e(A) \succeq e(B)]$. By transitivity of $\succeq$, it follows that, if $\#e(A) < \#e(B)$, then not$[A \succeq B]$ holds.

Next, we consider A and B such that (i) $\#e(A) = \#e(B)$; and (ii) for every one-to-one correspondence f from $e(A)$ to $e(B)$, there exist $x,y,z,w \in e(A)$ such that $d(x,y) > d(f(x),f(y))$ and $d(z,w) < d(f(z),f(w))$. We need to show that $e(A)$ and $e(B)$ are not comparable. Suppose to the contrary that they are comparable. If $e(A) \succeq e(B)$, by (P.5), there exist $a_1 \in e(A)$ and $b_1 \in e(B)$ such that

$e(A) \setminus \{a_1\} \succeq e(B) \setminus \{b_1\}$. Then, by (P.6), there exist $a^* \in e(A)$ and $b^* \in e(B)$ such that $e(A) \setminus \{a^*\} \succeq e(B) \setminus \{b^*\}$ and, for some one-to-one correspondence g from $e(A) \setminus \{a^*\}$ to $e(B) \setminus \{b^*\}$, $d(a^*, a) \geq d(b^*, g(a))$ for all $a \in e(A) \setminus \{a^*\}$. If $e(A) \setminus \{a^*\}$ contains two objects, say a and a', then, by Theorem 1, it must be true that $d(a, a') \geq d(g(a), g(a'))$. Consider the one-to-one correspondence g' from $e(A) = \{a^*, a, a'\}$ to $e(B) = \{b^*, g(a), g(a')\}$ defined as $g'(a^*) = b^*$, $g'(a) = g(a)$, and $g'(a') = g(a')$. Then, for the correspondence g' from $e(A)$ to $e(B)$, we have $d(u, v) \geq d(g'(u), g'(v))$ for all $u, v \in e(A)$, which contradicts the fact that there exist $x, y, z, w \in e(A)$ such that $d(x, y) > d(f(x), f(y))$ and $d(z, w) < d(f(z), f(w))$. Therefore, in this case, it cannot be true that $e(A) \succeq e(B)$. If $e(A) \setminus \{a^*\}$ contains more than two objects, then by the repeated use of (P.5) and (P.6), a similar contradiction can be derived. Therefore, $e(A) \succeq e(B)$ does not hold. Similarly, it can be shown that $e(B) \succeq e(A)$ cannot hold. Consequently, we must have that $e(A)$ and $e(B)$ are noncomparable. By transitivity of $\succeq$, it follows that A and B are not comparable.

Finally, we consider A and B such that (i) $\#e(A) > \#e(B)$; and (ii) for every subset $e'(A)$ of $e(A)$ with $\#e'(A) = \#e(B)$, and every one-to-one correspondence f from $e'(A)$ to $e(B)$, there exist $x, y, z, w \in e'(A)$ such that $d(x, y) > d(f(x), f(y))$ and $d(z, w) < d(f(z), f(w))$. Suppose $e(A) \succeq e(B)$. Then, by (P.7), there exists a proper subset C of $e'(A)$ such that $C \succeq e(B)$. We can assume that $\#C = \#e(B)$ since (i) $\#C \geq \#e(B)$ and (ii) if $\#C > \#e(B)$, then, by possibly several applications of (P.7), we can reduce the cardinality of C to $\#e(B)$. Note that, C is a subset of $e(A)$ and that $\#C = \#e(B)$. Then, for every one-to-one correspondence f from C to $e(B)$, there exist $x, y, z, w \in e'(A)$ such that $d(x, y) > d(f(x), f(y))$ and $d(z, w) < d(f(z), f(w))$. It then follows that C and $e(B)$ are noncomparable, a contradiction. Therefore, $e(A) \succeq e(B)$ cannot be true. Since $\#e(B) < \#e(A)$, it must be true that not$[e(B) \succeq e(A)]$. Therefore, $e(A)$ and $e(B)$ are not comparable. The transitivity of $\succeq$ now implies that A and B are not comparable. This completes the proof of Theorem 2. □

References

Bervoets, S., and Gravel, N. (2007). Appraising diversity with an ordinal notion of similarity: An axiomatic approach. *Mathematical Social Sciences*, *53*, 259–273

Bossert, W., Pattanaik, P. K., and Xu, Y. (2003). Similarity of objects and the measurement of diversity. *Journal of Theoretical Politics*, *15*, 405–421

Nehring, K., and Puppe, C. (2002). A theory of diversity. *Econometrica*, *70*, 1155–1190

Pattanaik, P. K., and Xu, Y. (2000). On diversity and freedom of choice. *Mathematical Social Sciences*, *40*, 123–130

Weikard, H. (2002). Diversity functions and the value of biodiversity. *Land Economics*, *78*, 20–27

Weitzman, M. (1992). On diversity. *Quarterly Journal of Economics*, *107*, 363–406

Weitzman, M. (1993). What to preserve? An application of diversity theory to crane conservation. *Quarterly Journal of Economics*, *108*, 157–183

Weitzman, M. (1998). The Noah's ark problem. *Econometrica*, *66*, 1279–1298