Macroeconomic Theory and Economic Policy

Essays in honour of Jean-Paul Fitoussi

Edited by K. Vela Velupillai

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Macroeconomic Theory and Economic Policy

An impressive array of leading macroeconomists have come together in a single book to pay tribute to Jean-Paul Fitoussi. That they have done so here is a tribute to the consummate macroeconomic theorist, Jean-Paul Fitoussi – in whose honour this book is written.

The development of macroeconomic theory and the implications for economic policy is the main focus of this book. Topics that are given new theoretical underpinnings, as well as alternative perspectives, in this book include:

- growth theory
- labour market theory
- rational expectations theory
- the development of macroeconomic theory and thought.

Coming together to honour Fitoussi, the contributors have pulled off an impressive coup in creating a book which will interest advanced students of macroeconomics, economists generally as well as those involved in the formulation of economic policy.

K. Vela Velupillai is John E. Cairnes Professor of Economics at the National University of Ireland, Galway and also holds a Chair at the University of Trento, Italy, as a 'Professore di Chiara Fama'. He is also a Senior Visiting Professor at the Madras School of Economics, Chennai, India.

Routledge frontiers of political economy

1 Equilibrium Versus Understanding
Towards the rehumanization of economics within social theory
Mark Addleson

2 Evolution, Order and Complexity
Edited by Elias L. Khalil and Kenneth E. Boulding

3 Interactions in Political Economy
Malvern after ten years
Edited by Steven Pressman

4 The End of Economics
Michael Perelman

5 Probability in Economics
Omar F. Hamouda and Robin Rowley

6 Capital Controversy, Post Keynesian Economics and the History of Economics
Essays in honour of Geoff Harcourt, volume one
Edited by Philip Arestis, Gabriel Palma and Malcolm Sawyer

7 Markets, Unemployment and Economic Policy
Essays in honour of Geoff Harcourt, volume two
Edited by Philip Arestis, Gabriel Palma and Malcolm Sawyer

8 Social Economy
The logic of capitalist development
Clark Everling

9 New Keynesian Economics/Post Keynesian Alternatives
Edited by Roy J. Rotheim

10 The Representative Agent in Macroeconomics
James E. Hartley

11 Borderlands of Economics
Essays in honour of Daniel R. Fusfeld
Edited by Nahid Aslanbeigui and Young Back Choi

12 Value, Distribution and Capital
Essays in honour of Pierangelo Garegnani
Edited by Gary Mongiovi and Fabio Petri

13 The Economics of Science
Methodology and epistemology as if economics really mattered
James R. Wible

14 Competitiveness, Localised Learning and Regional Development
Specialisation and prosperity in small open economies
Peter Maskell, Heikki Eskelinen, Ingjaldur Hannibalsson, Anders Malmberg and Eirik Vatne

15 Labour Market Theory
A constructive reassessment
Ben J. Fine

16 Women and European Employment
Jill Rubery, Mark Smith, Colette Fagan and Damian Grimshaw

17 Explorations in Economic Methodology
From Lakatos to empirical philosophy of science
Roger Backhouse

18 Subjectivity in Political Economy
Essays on wanting and choosing
David P. Levine

19 The Political Economy of Middle East Peace
The impact of competing trade agendas
Edited by J.W. Wright Jnr

20 The Active Consumer
Novelty and surprise in consumer choice
Edited by Marina Bianchi

21 Subjectivism and Economic Analysis
Essays in memory of Ludwig Lachmann
Edited by Roger Koppl and Gary Mongiovi

22 Themes in Post Keynesian Economics
Essays in honour of Geoff Harcourt, volume three
Edited by Claudio Sardoni and Peter Kriesler

23 The Dynamics of Technological Knowledge
Cristiano Antonelli

24 The Political Economy of Diet, Health and Food Policy
Ben J. Fine

25 The End of Finance
Capital market inflation, financial derivatives and pension fund capitalism
Jan Toporowski

26 Political Economy and the New Capitalism
Edited by Jan Toporowski

27 Growth Theory
A philosophical perspective
Patricia Northover

28 The Political Economy of the Small Firm
Edited by Charlie Dannreuther

29 Hahn and Economic Methodology
Edited by Thomas Boylan and Paschal F. O'Gorman

30 Gender, Growth and Trade
The miracle economies of the postwar years
David Kucera

31 Normative Political Economy
Subjective freedom, the market and the state
David Levine

32 Economist with a Public Purpose
Essays in honour of John Kenneth Galbraith
Edited by Michael Keaney

33 The Demise of Involuntary Unemployment
Michel De Vroey

34 The Fundamental Institutions of Capitalism
Ernesto Screpanti

35 Transcending Transaction
The search for self-generating markets
Alan Shipman

36 Power in Business and the State
An historical analysis of its concentration
Frank Bealey

37 Editing Economics
Essays in honour of Mark Perlman
Hank Lim, Ungsuh K. Park and Geoff Harcourt

38 Money, Macroeconomics and Keynes
Essays in honour of Victoria Chick, volume one
Philip Arestis, Meghnad Desai and Sheila Dow

39 Methodology, Microeconomics and Keynes
Essays in honour of Victoria Chick, volume two
Philip Arestis, Meghnad Desai and Sheila Dow

40 Market Drive and Governance
Reexamining the rules for economic and commercial contest
Ralf Boscheck

41 The Value of Marx
Political economy for contemporary capitalism
Alfredo Saad-Filho

42 Issues in Positive Political Economy
S. Mansoob Murshed

43 The Enigma of Globalisation
A journey to a new stage of capitalism
Robert Went

44 The Market
Equilibrium, stability, mythology
S.N. Afriat

45 The Political Economy of Rule Evasion and Policy Reform
Jim Leitzel

46 Unpaid Work and the Economy
Edited by Antonella Picchio

47 Distributional Justice
Theory and measurement
Hilde Bojer

48 Cognitive Developments in Economics
Edited by Salvatore Rizzello

49 Social Foundations of Markets, Money and Credit
Costas Lapavitsas

50 Rethinking Capitalist Development
Essays on the economics of Josef Steindl
Edited by Tracy Mott and Nina Shapiro

51 An Evolutionary Approach to Social Welfare
Christian Sartorius

52 Kalecki's Economics Today
Edited by Zdzislaw L. Sadowski and Adam Szeworski

53 Fiscal Policy from Reagan to Blair
The Left veers Right
Ravi K. Roy and Arthur T. Denzau

54 The Cognitive Mechanics of Economic Development and Institutional Change
Bertin Martens

55 Individualism and the Social Order
The social element in liberal thought
Charles R. McCann Jnr

56 Affirmative Action in the United States and India
A comparative perspective
Thomas E. Weisskopf

57 Global Political Economy and the Wealth of Nations
Performance, institutions, problems and policies
Edited by Phillip Anthony O'Hara

58 Structural Economics
Thijs ten Raa

59 Macroeconomic Theory and Economic Policy
Essays in honour of Jean-Paul Fitoussi
Edited by K. Vela Velupillai

Macroeconomic Theory and Economic Policy

Essays in honour of Jean-Paul Fitoussi

Edited by K. Vela Velupillai

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2004
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2004.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested

ISBN 0-203-60046-0 Master e-book ISBN

ISBN 0-203-34513-4 (Adobe eReader Format)
ISBN 0-415-32336-3 (Print Edition)


Cinquantenario del titolo di
Professore Ordinario
"Cattedra Ernesto Kroeber"
a Jean-Paul Edmond

This publication was grant-aided by the Publications Fund of the National University of Ireland, Galway

Contents

Contributors

Jean-Pascal Bénassy is Directeur de Recherche at CNRS and CEPREMAP, Paris.

Olivier Blanchard is the CLASS OF 1941 Professor in the Department of Economics at the Massachusetts Institute of Technology.

Robert W. Clower was, until recently, the Hugh C. Lane Professor of Economic Theory at the University of South Carolina. He is an Honorary Fellow of Brasenose College, Oxford and Emeritus Professor of Economics at UCLA.

Jacques Le Cacheux is Professor of Economics at the Université de Pau et des Pays de l'Adour and Director of the Département des études, Observatoire français des conjonctures économiques (OFCE), Paris.

Axel Leijonhufvud is Professor of Monetary Theory and Policy at the University of Trento. He was the founding Director of the Center for Computable Economics at UCLA.

Edmond Malinvaud is currently associated with CREST, in Paris, and was formerly Director of INSEE and Professor at the College de France.

Edmund S. Phelps is McVickar Professor of Political Economy at Columbia University in the City of New York.

Robert M. Solow, a Nobel Laureate, is Institute Professor at the Massachusetts Institute of Technology.

Jean Tirole is Professor and Scientific Director of the Institut d'Economie Industrielle a the University of Sciences in Toulouse.

K. Vela Velupillai is John E. Cairnes Professor of Economics at the National University of Ireland, Galway and Professor of Economics at the University of Trento.

1 Preface and summary

K. Vela Velupillai

One of the distinguished contributors[1] to this volume of essays in honour of Jean-Paul Fitoussi, when initially approached, responded warmly but with a genuine puzzle:

> I'm shocked that Jean-Paul is anywhere near 60! Have you ever checked your dates? It was only a few years ago that he moved into his 50s.

Tempus fugit! It was, indeed a few years ago that almost all of us 'moved into our 50s', and not long before that into our 30s.

When I first met Jean-Paul Fitoussi, at the *Badia Fiesolana* in Fiesole, almost a quarter of a century ago, he was a young man in his 30s; when I last met him in Paris, a few months ago, he was still a young man, but had entered his 60s. This volume was conceived before his sixtieth birthday, as a tribute by a few close friends and admirers, to honour his sixtieth birthday, on 19 August 2002. I am sure I speak for all the contributors, and many others, when I wish him many happy years through his 60s and many more decades of professional and personal adventures.

Whenever we meet, particularly in Paris, he asks me to recite, appropriately, a stanza from Auden's *Many Happy Returns*:

> So I wish you first a
> Sense of theatre: only
> Those who love illusion
> And know it will go far:
> Otherwise we spend our
> Lives in a confusion
> Of what we say and do with
> Who we really are.

The generosity, kindness and promptness with which all the contributors to this volume agreed to participate in this interesting and rewarding tribute to a friend, sometimes colleague and an always interesting

companion and coauthor to many of us, made my potentially arduous task a pleasure and a privilege. We have all shared memorable moments of intellectual and other adventures with Jean-Paul Fitoussi. This is only a very small token of our friendship, esteem and warmth for the person he is, was and always will be.

The rest of this chapter is an attempt at summarizing and collating the different contributions in a concise way.

The chapters have been arranged in alphabetical order by author's surname, although they could have been arranged in thematic or other ways. Had I arranged them in terms of purely thematic content I may have grouped the contributions by Bénassy, Clower and myself together as belonging to the more purely theoretical end of macroeconomic theory; the Leijonhufvud and Malinvaud chapters together as belonging to a broad-based narrative of the development of the interaction between the development of visions of macroeconomic theory and their impact on macroeconomic policy, the chapters by Le Cacheux, Phelps and Solow on the macrotheoretical underpinnings of macroeconomic policy and the Blanchard-Tirole chapter, closely related in structure to the latter three, as one concentrating on the theoretical underpinnings of a particular institutional aspect of macroeconomic policy. However informing these delineations may well be and may have been useful in providing structure to a collection of essays, there is a danger that it may place some or even all of the contributions in unwarranted straitjackets. It may also carry with it the danger that I would be distorting some of the subtle messages in the individual intents and aims. Hence, I have chosen to order and organize them alphabetically and provide, in the ensuing few pages, brief summaries to try to link them together.

The central concern of all contributions is the interaction between macroeconomic theory and economic policy – two themes that have characterized Fitoussi's lifelong research, and academic and administrative efforts.

Bénassy's contributions takes up a theoretical theme that has concerned Fitoussi's own macrotheoretic concerns, in a slightly different context, for a long time (cf. [1], p. 25 and note 40, p. 40). It is the undoubted empirical fact of the existence, in advanced industrial economies of all shades, of staggered wage and price contracts. Given this empirical fact, how can we theorize about, and answer questions on, their impact on output and employment persistence? Theoretical frameworks were suggested in a series of pioneering papers by Stanley Fischer, Jo-Anna Gray, Edmund Phelps, John Taylor, Guillermo Calvo and others during the late 1970s and early 1980s. That early research seemed to have lost some of the initial steam till it was revived in more recent years. Bénassy carries on this revival and, in the context of a solvable dynamic stochastic general equilibrium model, he incorporates a Calvo-type staggered wage and price contract scheme. The explicit optimal solution he obtains suggests that price

staggering produces less output persistence, for the same average duration of contracts, than wage staggering.

Blanchard and Tirole, in a chapter with rich and copious documentation of the historical facts and experiences of labour market institutions, tackle, against that backdrop, the crucial but thorny problem of designing 'optimal employment protection' institutions, norms and regulations. Their chapter addresses the problem, appropriately in a volume in honour of Jean-Paul Fitoussi, in the specific context of institutional reform in France. The general principle that underpins their approach is squarely in the tradition of underpinning a macroproblem in a rigorous micro setting: 'that of making firms internalize, to the extent possible, the social cost of unemployment'. With this general principle as a guiding light, they construct, first a benchmark model and, then, 'explore three main deviations from the benchmark, and their implications for employment protection'. Having done this, and informed and guided by it, they look at the 'employment protection system in place in France today'. This enables them, finally, to outline the contours of a possible scheme for reforming the institutional basis for employment protection in France. The Blanchard-Tirole chapter places the Bénassy framework in its institutional context and makes it possible to read it in a precise policy oriented way. Moreover, it is highly complementary to the issues raised in the chapter by Phelps. Above all, the 'usual culprits', invoked in the popular political press, are given short shrift in a meticulously documented, theoretically impeccable framework, and an institutionally informed and underpinned discussion. In the clarity with which the argument proceeds, in the way the evidence – empirical and institutional – is marshalled and corralled into a theoretical framework, it is a masterpiece of pedagogy, as well.

Thomas Sowell opened his magisterial book, *Say's Law*, with the remark ([6], p. 3) that the 'idea' of Say's Law, 'that supply creates its own demand, appears on the surface to be one of the simplest propositions in economics, and one which should be readily proved or disproved'. And goes on:

> Yet this doctrine has produced two of the most sweeping, bitter, and long-lasting controversies in the history of economics – first in the early nineteenth century and then erupting again a hundred years later in the Keynesian revolution of the 1930s. Each of these outbursts of controversy lasted more than twenty years, involved almost every noted economist of the time, and had repercussions on basic economic theory, methodology and sociopolitical policy.

Clower, in his chapter, returns to this topic, in its second incarnation, one that has remained central to macroeconomic theory *and* economic policy: the correct understanding, formalization and interpretation of Say's Law. We have debated – and, in this, Fitoussi, too has played his part –

endlessly about 'what Keynes actually meant', and variations on that theme. Much of that debate has taken for granted a particular codification, if not an ossification, of what Keynes meant with his strictures against Say's Law. But no one seems to have taken the trouble to go back and check with Say, himself, whether he did mean what all and sundry seem to attribute to him! Clower, with characteristic pungency, and scholarly accuracy, has done just that and stripped a host of emperors, particularly Lange and Patinkin, of their clothes in the process.

Jacques Le Cacheux is unique among the contributors to this volume in not only being an ongoing collaborator and consistent coauthor with Fitoussi on matters at the frontiers of macroeconomic policy issues in the European Community in general and in France in particular, but also in having been his doctoral pupil at the European University Institute. In a felicitous and almost balanced blend of theory and policy, Le Cacheux revisits a question that had occupied Fitoussi's efforts – together with Phelps – for quite some time in the 1980s: *the Slump in Europe*. To put it in more contemporary terms, the question is not 'why are we so rich and they so poor', but, 'why are we so productive and they so unproductive' and 'why do we tax so little and they so much'. Hence, in effect, 'why are we enjoying prosperity and they depression' – the 'we' being the US and the 'they' being 'old' Europe and Japan.[2] Le Cacheux returns to the framework suggested by Fitoussi and Phelps in the 1980s, but with due attention paid to the changed contexts and circumstances brought about by financial globalization. There is an underlying reliance on the theoretical vision provided by an overlapping generations framework to highlight the role played by demographic factors in the dynamics of assets and debts and their induced macroeconomic effects via their impacts on behaviour at the individual level of firms and households.

In a sense, Leijonhufvud, Malinvaud and Solow tackle an issue that is at the heart of macroeconomics: the kind of vision that dominates the macroeconomics of an era or of a generation.[3] Time was when macroeconomic policy was synonymous with stabilization policy and an active government role in the policy-making and implementing process. Now, as Leijonhufvud ruefully ruminates, the pendulum has swung and the drive is to constrain government actions. A generation and an era ago, the vision of industrial societies was permeated by pathological metaphors: instabilities, fluctuations, market failures and the like; to the present generation, reared on theoretical technologies emphasizing equilibrium stochastic dynamics, pathologies have been banished. Economies reside on intertemporal equilibrium paths, buffeted, now and then by shocks of one sort or another, but containing enough self-adjustment capabilities to return to tranquillity. How and why did – and do – visions change? Leijonhufvud tries to pose this challenging question and answers it with a wide ranging, panoramic sweep over the recent development in macroeconomics to extract lessons, precepts and prescriptions.

Malinvaud's questions are closely related to the concerns expressed in Leijonhufvud's chapter, but with a more specific focus: the twenty-year period, 1975–1995, spanning the two decades in which there was the emergence, consolidation and dominance of macroeconomic theory by, first, the rational expectations 'revolution' and, then, the refocusing of the problem of economic fluctuations as one of *real business cycles*. His focus is not just the changing visions that imply, for any generation, a particular way of doing macroeconomic theory; but it is also how this particular way of doing theory may impact on the kind of issues that could and would be dealt with in the framework of macroeconomic policy. He does not deny or belittle the substantial contributions of these two related movements in macroeconomic theory – not least in conceptual innovations and methodological, especially mathematical, sophistications. However he rues the fact that the mathematical sophistications seem to have become an end in themselves, without any anchoring in empirical phenomena or policy concerns, indeed, even contributing to a nihilistic attitude towards policy. A wedge is being driven between what was once a seamless continuum, between theory and policy, largely due also to the internal criteria the macroeconomic profession seems to be according the kind of work to be prized, rewarded and published. The poignant point he tries to make is most vividly illustrated in the way he discusses the Lucasian obsession with computing welfare gains – or losses – of a representative consumer, as a result of stabilization policies attempting to control or eliminate business cycles.[4] Malinvaud, almost plaintively, suggests that economists should also heed the testimonies of scientists in the ancillary social and human sciences when trying to discuss and theorize, using intertemporal stochastic equilibrium models, the welfare consequences of economic booms and depressions.

Phelps, starting from the foundations that underpinned his work with Fitoussi in the 1980s, where they tried to provide a novel theory about the deep slump in which Europe was mired at that time, goes on to explain, first, the genesis of what he has come to call *structuralism*. This sets the stage for him, second, to be able to delineate and differentiate its non-monetary bases in comparison with supply-side and real business cycle models and thereby exposing the weaknesses and flaws in the latter and showing how structuralism avoids these pitfalls. *Structuralism* was set in its felicitous paces in the 1994 book by Phelps, *Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest and Assets*. The title makes explicit the difference with *The General Theory* in its substitution of 'assets for money'.[5] Phelps has been indefatigable in emphasizing the importance of taking an 'asset view' of business life – a theme taken up in Le Cacheux's chapter, largely inspired by the work of Phelps and Fitoussi–Phelps. Indeed, it may not even be an exaggeration to summarize the structuralist programme – a previous generation would have been familiar with a different kind of structuralism, just as there was, once,

another kind of New Keynesianism – in one phrase: the importance of modelling investment in business assets and the theory of expectation formation that underpins it. I myself have come to view this new Phelpsian vision as a genuine development of *The General Theory*, for a generation imbued with New Classical economics and its concepts, myths, theoretical technologies and the concomitant policy prejudices.[6] Phelps, with deftness, diplomacy and nuanced panache takes the new classicals on squarely, on almost every front, but within their framework. This is offence at its civilized best, reminiscent of Keynes at his best, and is elegantly exemplified by the way he first, apparently with approval, cites Lucas's warning to 'beware of theorist's bearing free parameters', and, then, in almost the same breath, to invoke Occam to deflect Lucas, by making a case against the warning. Finally, and most felicitously, Phelps ends his essay by noting: 'It is a reasonable hypothesis that Europe's employment levels would pick up to permanently higher levels if several economic institutions were retired or reformed and several new institutions installed. The modelling that will make this view operational and testable largely remains to be done.' A part, at least, of 'modelling that ... remains to be done' is what is attempted by Blanchard and Tirole, in this volume.

Solow, elegant and concise as always, dissects a simple theoretical point at the heart of growth theory, to illustrate its profound significance in macroeconomic policy contexts. Growth theory, Solow observes, has been 'locked' into the strategy of starting with those macroeconomic assumptions that are necessary to ensure the existence of 'one or more exponential steady states'. Once locked into this mode of thinking, then whenever growth models are used as guides to policy, one is further constrained to think as if the only alternative open to policy makers is to fashion instruments to 'increase the long-term growth rate'. Solow suggests, with characteristically persuasive prose of supreme simplicity, that this is not only unnecessary but also 'dysfunctional both for theory and policy'. In speculating as to why we may have got locked into what may be called the 'exponential vision' of growth theory, he states that there may have been empirical, technical and casual reasons. Now, among the possible 'casual' reasons there is, as Solow notes, Harrod's pioneering role – both for good and bad.[7] But 'Harrod had little command of analytical technique. An ordinary differential equation would not have been a natural mode of thought for him', points out Solow. Perhaps I can add a possible factual element to this casual, but perceptive speculation, by Solow. A few months after *The General Theory* was published Harrod came out with his own, little, book, *The Trade Cycle*. Tinbergen reviewed this latter book in the May 1937 issue of the *Weltwirtschaftliches Archiv*. In the review Tinbergen made the rather devastating observation that the economic assumptions underlying Harrod's attempt to build a model of the trade cycle implied a *first order differential equation* which, in turn, is incapable of generating fluctuations but can only generate exponential paths as its solutions.[8] The

review was written in German and Harrod, quite apart from having only 'little command of analytical technique', had equally little command of the German language.[9] However, Tinbergen's review and its above crucial point was brought to Harrod's attention by his then pupil, Richard Goodwin. Harrod is reputed to have told Goodwin that he would 'think about it', and did, indeed, go off and 'think about it' for two years and came up with the *Essay in Dynamic Theory* to lock growth theory into its exponential vision. The rest, as the saying goes, is history!

My own chapter traces the origins of the way rational expectations formalisms got 'locked' into its topological mode of theorizing and suggests how a recursion theoretic formalization could make the theory more open-ended from the policy point of view. The main message of the chapter is to make the case for alternative mathematical formalisms for macroeconomic concepts and issues and to point out how that might lead to diametrically opposite policy implications.

All of the chapters above pay homage to Fitoussi in a very particular sense, a sense that underlines his own attitude to theory in its relation to policy. All theories are permeated by visions of the workings of the macroeconomic system. These visions characterize an era – others may use the word *Zeitgeist* – and all those of us working in macroeconomic theory are, often, prisoners of the visions. To that extent the envisioned macroeconomic policies are equally imprisoned within that vision. Leijonhufvud's chapter told the story of how these visions may have their own dynamics. To understand, therefore, that there are alternative visions, over time, should encourage the more audacious theorists to speculate, at a point in time, of the possibility of alternative visions – as Fitoussi constantly struggles to point out, in his tireless efforts to propagate more inclusive policy alternatives. Malinvaud's chapter emphasizes the damage that can be caused to the relationship between theory and policy when visions are determined by narrow theoretical, conceptual and mathematical technologies – a theme further developed, in the context of growth theory by Solow. All of them suggest that an acknowledgement of the existence of such dominant visions is also a first step in being able to transcend them. The essays by Bénassy, Blanchard–Tirole, Le Cacheux, Phelps and myself are attempts to show how, concretely, one can transcend the dominant vision, within its own strictures and underpinnings, thus broadening the policy horizons. This is, essentially, the philosophy advocated by Fitoussi – in his academic writings, in his policy pamphlets and in his exhortations to policy makers in every conceivable medium.

Notes

1 Edmund Phelps.

2 Le Cacheux opens his essay by framing the questions he poses in this way and he seems to have been influenced in this by Prescott's recent *Richard T. Ely Lecture*, titled, 'Prosperity and Depression' (*pace* Haberler!) ([4]).

3 I have, quite consciously, eschewed reliance on terminology from the philosophy of science in this Summary. I did not want to use clichés like 'normal science', paradigm and so on, lest the unwary reader interpret my intentions in ways that might not sit comfortably with the aims of the various distinguished contributors to this volume.

4 Malinvaud refers to the 1987 Lucas discussion of this issue in *Models of Business Cycles*. He could equally felicitously have made the same point – perhaps even more forcefully – by referring to the recent Presidential address by Lucas to the American Economic Association ([2]).

5 Of course Keynes also chose to use the word 'employment' in his famous title, rather than 'unemployment'.

6 Phelps thinks that although the idea for the name *structuralism* may have been, partly, 'a joke', with origins in the genesis of the word in the Sorbonne circles around Claude Levi-Strauss, it could also, partly, have emanated from reflections on a Hayekian admonishment that Mr Keynes did not have 'enough structure'!

7 In another context, Paul Samuelson made a similar observation regarding Frisch

> In leaving Frisch's work of the 1930s on stochastic difference, differential and other functional equations, let me point out that a great man's work can, in its impact on lesser men, have bad as well as good effects. Thus, by 1940, Lloyd Metzler and I as graduate students at Harvard fell into the dogma … that all economic business-cycle models should have *damped* roots.… What was so bad about the dogma? Well, it slowed down our recognition of the importance of non-linear atorelaxation models of van der Pol-Rayleigh type, with their characteristic amplitude features lacked by linear systems ([5]).

8 The relevant sentence is ([7]):

> Die Kombination also der »relation« mit dem Satz über den »multiplier« in der oben beschriebenen Weise gibt essentiell keine Theorie des Zyklus.

9 I cannot help recalling, in this context, Myrdal's acid remark on Keynes ([3], pp. 8–9),

> the attractive Anglo-Saxon kind of unnecessary originality, which has its roots in certain systematic gaps in the knowledge of the German language on the part of the majority of English economists.

Fitoussi's many policy writings have suffered from a similar problem due to the difficulties faced, not only by Anglo-Saxons, with their 'systematic gaps in the knowledge of the French language'!

References

[1] Jean-Paul Fitoussi (1983) 'Modern Macroeconomic Theory: An Overview', in: *Modern Macroeconomic Theory* edited by Jean-Paul Fitoussi. Basil Blackwell, Oxford.

[2] Robert E. Lucas Jr (2003) 'Macroeconomic Priorities', *The American Economic Review* 93 (1) March: 1–14.

[3] Gunnar Myrdal (1939) *Monetary Equilibrium*, William Hodge & Co., London.

[4] Edward C. Prescott (2002) 'Prosperity and Depression', *The American Economic Review* 92 (2) May: 1–15.
[5] Paul A. Samuelson (1974) 'Remembrances of Frisch', *European Economic Review* 5 (1): 7–23.
[6] Thomas Sowell (1972) *Say's Law: An Historical Analysis*, Princeton University Press, Princeton, New Jersey.
[7] Jan Tinbergen (1937) 'Review of *The Trade Cycle: An Essay* by R.F. Harrod', *Weltwirtschaftliches Archiv* XLV, May: 89–91.

2 The consummate macroeconomist

Jean-Paul Fitoussi*

K. Vela Velupillai

1 Preamble

> Suddenly a sparrow darts in
> Through a door, flits across the hall
> And flies out through another one.
> Inside, cocooned in light and warmth
> It can enjoy a moment's calm
> Before it vanishes, rejoining
> The freezing night from which it came
>
> James Harpur: *The Flight of the Sparrow*

One fine, crisp, winter's day, in February, 1981, I received a telephone call from Jean-Paul Fitoussi's secretary. I was in my office at Peterhouse, having just succeeded my own *teacher*, *mentor* and *friend*, Richard Goodwin, as Director of Studies in Economics at that most ancient of colleges – indeed the oldest in Oxbridge. Fitoussi was in Fiesole, at the European University Institute, having himself succeeded the late Andrew Schonfield as the head of their small economics department. The telephone call contained the serendipitous[1] invitation to come for an interview, preceded by a seminar which I was to give on any macroeconomic topic of my choice, for the post of an Assistant Professor. I had also just been offered a Fellowship at Fitzwilliam College and we, my wife and three young children (the last one not quite eight months old), had only moved back to Cambridge less than one year before that fateful telephone call. In spite of this, I agreed, almost with alacrity, to come for the interview.[2]

In those late winter months of 1981, Fitoussi was fully immersed in editing the proceedings of a very successful conference he had organised at the EUI, in May, 1980. The book that emerged from that effort ([7]),

contained the lead chapter by Fitoussi himself, 'Modern Macroeconomic Theory: An Overview'[3] and it was being crafted and forged with extreme diligence and care, when I first met him, in late February, 1981. Naturally, I decided to speak about *Disequilibrium Macrodynamics*, but being a product of 'old' Cambridge, it was in the tradition of Kaldor–Goodwin models and a sprinkling of ideas from the, then influential, Solow–Stiglitz paper of 1967 ([32]). However, I did have two trump cards up my thin and short sleeves: I was fairly well equipped in the Swedish economics of the late 1920s and 1930s; and I had been trained by Goodwin on nonlinear dynamics. These were worked into the model I presented at the seminar, particularly in showing a very general cyclical result using Kolmogorov's extension of the Lotka–Volterra model ([22][4]).

Fitoussi liked the presentation and the model[5] and I was offered the post – which I accepted with pleasure. Thus began a friendship, occasionally punctuated by collaborative writings, often sealed by prolonged discussions on matters economic, and, at least during those common years in Fiesole, constant interaction on a variety of intellectual matters and, not least, a shared concern for the economic predicament of the underdog, the underprivileged and the less fortunate. While I have retreated to intellectual reclusiveness and obscurity, Fitoussi has gone on to conquer a variety of new worlds, putting into practice those early and shared concerns, underpinning enlightened practice with imaginative theories and broaching new frontiers of political, social and economic debate, across academic and national frontiers. He has, almost single-handedly, resurrected the tradition of the 'Consultant Administrators' and the 'Political Arithmetician', a tradition extolled by Schumpeter in his monumental *History of Economic Analysis* ([30]), esp. chapters 3 and 4, to new levels of rigour and passion, with a series of tightly reasoned pamphlets and booklets on issues of current concern.[6]

This chapter is not a survey of Fitoussi's varied and multi-faceted contributions to macroeconomic theory and economic policy. The time has not come for that; he is in full flow and, even as I write, his interests are veering towards the Social Philosopher whose reasoning is underpinned and informed by serious economic theory, combined with political philosophy, at the frontiers. Instead, this chapter attempts to chart the rough contours of parts of the formative years[7] and, perhaps, to provide enough of that background to enable one to understand the passions that motivate this consummate macroeconomic theorist and policy activist, who is also becoming a social philosopher of considerable influence, at least in European academic intellectual circles.

The next section charts the contours of the early and formative years, largely personal, but partly also outlining the early academic influences that determined the eventual direction Fitoussi's research and work took. In section 3, I try to summarise the impetus that led to the work and career of the Fitoussi who became one of France's – and Europe's – leading

macroeconomists. The final section contains ruminations on his present broader concerns and speculations on where, and to what, they may lead.

2 'What, be a singer born and lack a theme?'[8]

> In older, traditional societies, one's reach into the future was usually determined by one's reach into the past: instant mobility was the exception. . . . Few Jews had ancestors they knew anything about: one knew about one's grandparents and one knew about the common descent from Adam and Abraham; the interval was marked by obscurity of the Diaspora.
>
> Fritz Stern ([33], p.4)

The origins of the name Fitoussi remains murky and mysterious, lost in the annals of time, like so many Jewish names. But there are small hints and fragile threads that link its origins to the Jabal Nefussa mountains in modern Libya. In more recent times there are links with the Livornese Jewish community in Tunisia of Sephardic origins. It is possible that an Issaskhar Fitoussi of Setif in Algeria published a book in Hebrew with the *Casa Editrice* of Livorno.[9]

Jean-Paul Fitoussi was born on 19 August 1942 in La Goulette, about 10 kilometres from Tunis, in Tunisia, the youngest of four children. His father, Joseph Fitoussi, was born in 1902 in Tunis and died there in 1963. His mother, Mathilde, née Cohen-Balalum, was born in Tunis in 1911 and died in Paris. Joseph Fitoussi was, early in his career, a primary school teacher, but later on worked as an actuary for an insurance company. Jean-Paul Fitoussi, in various personal conversations, when we have reminisced about our growing years and their formative influences, has always spoken fondly of his father's gentle character and subtle influences. His father had, for example, even in family conversations, emphasised the importance of statistics in understanding life's vicissitudes and, hence, had inculcated in the young Jean-Paul the desire to understand the foundations of statistics. On his mother's side, a cousin, Roger Cohen, was a distinguished journalist and a director of the biggest French newspaper in Tunisia, *La Presse de Tunisia*. Perhaps the quantitative side was inherited from his father's passion for statistics and the flowing, prolific, 'consultant administrator' cum 'pamphleteering' talents descended from the Cohen-Balalum lineage.

He married the proverbial childhood love, also of Tunisian origin, Annie Krief, in 1964. Their two children, Lisa and David Lawrence, were born in 1974 and 1982, respectively.

His brother chose to study medicine and a sister chose law. Jean-Paul recalled writing from Paris to his parents in Tunis that he had decided to study economics and that this choice was not completely arbitrary. It was because he thought that by studying economics he would be combining the

logic of law with the quantitative foundations of statistics that may be given by the mathematics of economics.

Jean-Paul Fitoussi gives handsome credit to two of his teachers, from his undergraduate and doctoral years in Strasbourg and Paris, for the lifelong passion for rigorous economic thinking and macroeconomics: Paul Chamley and Paul Coulbois.[10] The former, a professor in the faculty of law and economics, was instrumental in teaching Jean-Paul the importance of money and trade. He lectured on trade theory and the history of economic thought. His approach seems to have been philosophical and, thus, on money would start from Hegel[11] and reach Simmel in his monetary courses. His lectures and reasonings were, as Jean-Paul recalls, models of clarity, rigour and displayed great and deep learning and familiarity with the classics. Thus, his informed discourses on trade were solidly based on history – both of thought and facts of trade and tempered by philosophy. He was, as far as a political philosophy could be imputed to him, an enlightened liberal.

Paul Coulbois taught macroeconomic theory and economic policy. It was from him, during the dawning years of Lucasian macroeconomics, that Jean-Paul Fitoussi learned the classical theory of economic policy. Those lectures, he recalled during a conversation many years ago, were given with passion and conviction and it was difficult not to be persuaded of the power of that framework. That learning seems to have left an indelible mark on Fitoussi's way of thinking about policy, tempered by a respect for the informing hand of history that was the message of Chamley. Even though both of these mentors were supremely logical in their reasonings, clearly evident in the nature of their lectures, neither of them possessed or harnessed mathematical tools of any complexity. They belonged to a generation that had learned its economics in faculties that were amalgams of law and economics, with the philosophy and the logic of law giving the rigorous underpinnings that a later generation, to which Fitoussi himself had to teach, absorbed from mathematics.

The economic-philosophic bent, inherited from Paul Chamley, the quantitative talents, descending from paternal influences, the pamphleteering penchant, a possible talent from the maternal lineage, a passion for logical and rigorous reasoning, an autonomous development and a concern and a conviction of the efficacy and desirability of active macroeconomic policy, deeply influenced by the teachings of Coulbois, have all, each in their own small way, gone into the synthesis that is Jean-Paul Fitoussi. The Fitoussi who is the author of the more economic–political–philosophical treatises, pamphlets and booklets from the mid-1990s – from, say, *Le Débat Interdit* ([8]) through *Le Nouvel Âge des Inégalites* ([9]) right up to the most recent *La Régle et Le Choix* ([10]) – reflect those early influences of Chamley, bolstered by the teachings of Coulbois, but given a personal and passionate bent by the logic of economic theory, particularly macroeconomic theory, and a profound knowledge of the workings of the institutions and traditions of the European economies.

But the more explicit macroeconomic concerns, above all with feasibility of full employment in civilised societies, after Keynes, was to become a perennial issue in the intellectual and professional life of Fitoussi – as a teaching professor, as a researching intellectual, as a consultant to ministers and bureaucrats at national, transnational and international levels and as a concerned citizen and a daunting pamphleteer. It is not surprising, then, that the Hegelian quote to which I referred to earlier was bracketed, a few pages later, with the concerned observation by the Scitovskys, to which Fitoussi has, time and time again, referred – at least in personal conversations:

> Full employment, growth and price stability are three aims of economic organisations about whose desirability most people seem to be agreed. . . . *Of the three, full employment is the only one whose desirability we regard as self-evident*. We consider it the prime function of an economy to enable everybody willing and able to work to earn a living, and only a fully employed economy performs this function.
>
> [28], pp. 429–30; italics added

Fitoussi, having lived through the golden quarter-century of Keynesian policy optimism of 1947–1972, and now forced to witness the obduracy of institutions and policies that have reversed the priorities from full employment to price stability at any cost, finds himself incensed by the lack of compassion in ruling circles. It is not for nothing that the French title of his last booklet ([10]) in its Italian translation comes out as *The Benevolent Dictator* ([11])!

The themes that have driven his intellectual pursuits, the concerns that have informed his public policy stances, the passions that rule his visions for a future have all their roots in those formative years of liberal, enlightened, upbringing, education and research.

3 'The point of noon is past. Outside: light is asleep. . . .'[12]

> I remember one said there were no sallets in the lines to make the matter savory; nor no matter in the phrase that might indict the author of affectation, but called it an honest method, as wholesome as sweet, and by very much more handsome than fine.
>
> Hamlet, Act. II, Sc. ii

Fitoussi entered the macroeconomic arena at an opportune time. The mid-1960s initiated the period of turmoil and dissatisfaction with the tranquil era of the *Neoclassical Synthesis*. Clower had just unleashed his series of challenges to Neowalrasian complacencies ([3]) and Leijonhufvud's own *magnum opus* was about to be completed ([23]). The second edition of Patinkin's *MIP* ([27]) had also just been published. And, of course, the

series of papers by Phelps had immediate impact on Fitoussi,[13] as well as Hahn's piece in the same IEA volume in which Clower (op. cit.) had appeared ([18]). The Phillips classic and, some time later, Lipsey's analytic piece ([25]), in the *Phillips Memorial Volume*, outlining the theoretical genesis of the Phillips' curve, also made a deep impression on Fitoussi. This was particularly influential, as I recall, in the way he taught macroeconomic theory, especially because Lipsey, skilfully and with refined subtleties, showed how it emerged naturally from Phillips' earlier work on *stabilisation policies*. Naturally, the policy context from which it was derived by Lipsey resonated well with Fitoussi's perennial policy concerns in any macroeconomic issue. But, in French circles, Bent Hansen's much earlier work ([19]) developed along Lindahlian lines,[14] was in circulation and Fitoussi was well versed in it, even before he came to the more standard challenges to the orthodox macroeconomic theories of the times. I must, however, mention that, over the years, Fitoussi has consistently maintained that Arrow's extraordinarily perceptive paper in the Bernard Haley *Festshcrift* ([2]), was one of the great eye-openers for him.

Thus the central problems of macrodynamics of the age was the stage on which Fitoussi crafted and honed his analytical, empirical and policy skills. The issues he tackled, with a variety of techniques and concepts, but always faithful to a fundamental Keynesian vision, were the traditional grand themes of Macroeconomics: *unemployment*, *inflation*, *growth*, *fluctuations* and *functional income distribution*, in those early, heady, years. These issues were set against the backdrop of the 'raging' debates on the appropriate 'microeconomic foundations for macroeconomics' and whether or not the natural dynamics of macro issues should be framed as equilibrium or disequilibrium phenomena and whether or not voluntary or involuntary decision underpinnings were relevant, especially in dealing with unemployment.[15] From the years I shared as a colleague at EUI in Fiesole, I think I am right in remembering that he came down squarely in favour of disequilibrium dynamics and involuntary unemployment as backdrops for macrodynamic analysis, again especially in the context of unemployment.

This was also a time of great revitalisation of various French traditions in economic theory. For years the macroeconomic mantles at the frontiers had been worn by the Anglo-Saxons, the Swedes and the Austrians. Although the microeconomic and general equilibrium theoretic traditions seem always to have had a French presence and component, the macroeconomic and macrodynamic lineage had by-passed French circles during the consolidation of the Neoclassical Synthesis, even though there was a clear NeoWalrasian element in that synthesis – as Clower, more than anyone else, took great pains to point out.

With Malinvaud's influential and deceptively simple *Yrjö Jahnsson Lectures* ([26]), synthesising diverse contemporary French traditions in macroeconomic theory and microeconomic foundation, there was a veritable blossoming of French Keynesian macroeconomics that literally took

the macroeconomic profession by the proverbial storm. Fitoussi, of course, was ideally placed and well-equipped, intellectually, to be at the frontiers of this revival of French macroeconomics. Almost all his work prior to that particular event could, with hindsight, even be said to have been a presaging for it. He embraced the framework suggested by Malinvaud,[16] extended it in various imaginative ways, and took it upon himself to codify it – even as a pedagogical synopsis.[17] This led to that amazing *tour de force* of the early 1980s, when he undertook a comprehensive survey of the state of macroeconomic theory and produced what many regarded as a masterpiece: 'Modern Macroeconomic Theory: An Overview' ([7]). The clear and unambiguous message in that majestic paper was that the business of microeconomic foundations of macroeconomics was too complex to be left to the general equilibrium theorists; that there were many ways to look for foundations and one would need a metatheory – which we neither had nor had the remotest idea how to go about devising one – to choose between them. Fitoussi reiterated Keynesian and Hicksian wisdoms: there was an independent and autonomous macroeconomic logic that had to be pursued in its own way, on the basis of its own accounting categories.

With hindsight, and on reflection, having been close to Fitoussi and worked intimately with him during those exciting years of French Keynesian macroeconomic revival, I think that majestic 'Overview', and the enormous work that went into putting it together, marked a watershed in his own macroeconomic visions for the future. In a sense he had worked on completing that paper for several years and, although when he conceived it, and even in the early years while he was writing it, fixed-price macroeconomics seemed on the verge of becoming the dominant macroeconomic school, by the time he had finished it, the Lucasian dominance was almost sealed. The resurgence in *endogenous business cycle theory*, even that with a large and distinguished French component at the hands of Bénassy, Malgrange, Grandmont and others, came to be swamped by the tornado that hit the profession in the form of the *real business cycle theory* paradigm. Fitoussi, whose disquiet had been extreme in the face of the nihilistic policy stand of New Classical economics, and had been devising theoretical structures and conceptual frameworks to nullify it was, perhaps, taken as much by surprise as by bewilderment at the ease with which the hard-won alternative paradigms, yet in formative stages, were being surrendered. He had taken a courageous stand against the New Classical visions in his long and closely reasoned review article on 'Lucas Collected'.[18] But the tide had turned and the article had no repercussions whatsoever and fell off, proverbially, like water off a duck's back.

Throughout the 1980s and early 1990s, in spite of the increasing dominance of New Classical macroeconomics, Fitoussi continued his macrotheoretical work with a series of innovative approaches to unemployment and fluctuations – particularly the European slump of the 1980s. The paper with Georgescu-Roegen (op. cit.) broached several seminal directions and

re-emphasised earlier concepts that Fitoussi himself had developed in related contexts. Thus long before the emergence of a codified 'New Keynesian School' of macroeconomics, Fitoussi had worked into a model of unemployment (cf. [6]), elements of asymmetries due to irreversibilities and sought the essential 'local-global' divide without which interesting dynamics is impossible, in these elements. Lucas had, of course, located that divide via the use of the Phelpsian 'Island Paradigm'. This particular paper was also prescient in its embryonic discussion of concepts of *emergence*, a term that has become fashionable in so-called 'complexity circles', and attempts to incorporate it in the context of Schumpeterian innovations.

Then there was the attempt, this time jointly with Phelps, to try to provide a theoretically satisfactory and empirically implementable explanation of the slump in Europe in the 1980s – especially the *persistence* of unemployment ([17]). It was a book discussing *disemployment*, i.e., *disinflation + unemployment*, as against the, by then, orthodox problem of *stagflation*, i.e., *inflation + unemployment*. The authors went about devising an interesting impulse-propagation framework with an ingenious mechanism for the lower turning point in a model of fluctuations:

> A decline in capital stock will not produce a decline of employment which will produce in turn another decline of the capital stock and so on, in an endless spiral. Rather, the decline of capital will *outstrip* that of unemployment until capital touches bottom. At that time there will no longer be a force countering recovery of employment. Barring fresh stocks, employment and capital will rise strongly together on a route to complete recovery.
>
> ibid., p. 94; italics added

I recall, when first reading that fine, albeit dense book of 131 pages that crammed into it 19 diagrams, 13 tables and 44 equations, that this mechanism for the lower turning point was nothing other than a Hicksian 'floor' and, therefore, looked for a similarly clear mechanism for the 'ceiling'. I do not recall finding it easily; but had I found it I would have suggested to the authors that the impulse-propogation dichotomy, although sufficient would not be necessary for their model of the European slump. But they were writing in an age when *ad hoc* shock theories of the cycle were the frameworks of discourse and it was appropriate that they chose a similar language because they were trying to convince policy makers who had donned RBC blinkers, almost without exception. The authors did identify, in the manner that has become fashionable to talk about *stylised facts*, eight such things and, for the *impulse* mechanisms, two and three for propagations. I suspect the three crucial stylised facts, in the context of the discussion of the times, were fluctuations in the mark-up rate in the US and Europe,[19] abnormally high real interest rates and fluctuations in functional shares.

I may mention, in passing, that in our two joint papers with disequilibrium flavouring to them, we did tackle a couple of unusual issues. In the first of them ([14]) we tried to find a way to model and estimate the *size* of a *Leijonhufvudian corridor*, in a model which dealt explicitly with fluctuations in functional distributive shares. The technical originality of our paper lay in the way we used the Hopf bifurcation theorem, using Tobin's q as the critical parameter, in *three dimensions*.[20] In our effort for the *Patinkin Festschrift*, on the other hand ([16]) we tried to correct what we called a slip in *MIP*.[21] Patinkin's discussion and formalisation of a *distribution cycle* (op. cit., p. 501ff and, especially, figure A-8, p. 502) required the full paraphernalia of *homoclinic dynamical formalisation* to encapsulate his eminently reasonable economic intuitions. But, of course, such formalisations, although available and buried in the French works of Poincaré, had not become part of the folklore of modern dynamical systems theory till the late 1950s and early 1960s – long after Patinkin had conceived his economic problem. What was surprising was that such a famous and oftenquoted book, as *MIP* was, had not unearthed a reader who had spotted this infelicity in it! But our 'correction', in the *Patinkin Festschrift*, was a simple 'patching up' affair.[22]

There were many other forays into theoretical questions but, in a sense, the die had been cast with the move from purely Professorial duties at the European University Institute to the joint duties of Professor and a fulltime administrator at two prestigious institutes in Paris: Professor at the Institut d'Etudes Politiques de Paris and, first Director of Research and, then, Président de Observatoir Français des Conjonctures Economiques.[23] Fitoussi was obviously thrust into the thick of policy debates and discussions in the highly centralised French administrative machinery and, also, at the European level, as advisor to Jacques Attali in the formative years of the EBRD. I believe he had to develop a wholly different vision and attitude to economics and economic theory as a result of close encounters with people who were the architects of national and international policy proposals and with people who had to implement them. This is, of course, a phase of his life that has, in a sense, only just begun and it would be foolish to try to detect a pattern or infer a contour. I shall, therefore, in the next section, only scratch the surface that is very much in the making and try to provide a glimpse of the forces that may be driving towards becoming almost a veteran social and economic philosopher.

4 'What next? After this triumph, what portends?'[24]

I'm not such an idiot
As to claim the power
To peer into the vistas
Of your future
W.H. Auden: *Many Happy Returns*

Fitoussi has always maintained faith in the classical framework of the theory of economic policy, not uncritically, but as a basis for framing a reasonable discourse about feasibilities and desirabilities when thorny problems of normative values interact with descriptive complexities. In my own discussions of such themes, mostly at a purely theoretical level, I have often narrated to him, the story of the genesis of the original framework for the theory of economic policy. It was this that became codified as the Tinbergen–Frish model of targets and instruments. This, in turn, had its origins in the memoranda that were prepared by Gunnar Myrdal[25] and Erik Lindahl for the newly elected Social Democratic government of Sweden, in the depths of the depression of the 1930s. The story[26] goes something like this.

The Minister of Finance of the newly elected Social Democratic Government of Sweden, around 1933, Ernst Wigforss, facing the task of tackling the massive problems of the effects of the depression, had come to Gunnar Myrdal and Erik Lindahl with a conundrum which he wanted them to help him resolve. The problem was, as he put it to the two economists, that he – Wigforss – thought he knew what he had to do to solve the problem of unemployment but he did not have a *theory* to underpin it. What Wigforss thought he ought to do, at least to begin with, was to underbalance the budget.[27] But, he reputedly told Myrdal and Lindahl, he had *no theory* for such an action and he would face, in Parliament, as the Leader of the Opposition, the Professor of Economics at Stockholm, Gösta Bagge. Wigforss was worried that Bagge would counter his budgetary proposals with the full force of *orthodox economic theory* which, as Wigforss understood it, had no scope for underbalanced budgets!

I do not know whether the story is literally true but, of course, stories need not be true to be believed; they need only become parables. I think this story, as a parable, illustrates the precise nature of Fitoussi's perplexities and stances against the practices and policies being advocated and implemented by the Commission of the European Communities. In particular, he often singles out the triumvirate of the Commissioner for Competition, the Commissioner for Budgets and the President of the European Central Bank and their offices,[28] for scathing, but well reasoned and meticulously documented criticism, from one particularly principled standpoint: these officers and offices devise and implement policies on competition, budget balances and price stability *as if* reality *must* conform to a textbook norm which has the status of an eternal truth. To put it more bluntly, policies on these fronts are devised, by the respective Commissioners and the President of the European Central Bank, *as if* the market and all other relevant economic institutions conform to some ideal norm that is represented in a mythical arcadia and codified in a textbook. If a market or any other institution in any member country does not do so, then it must be 'whipped' into conformity before it strays too far into territories uncharted by the textbook 'first approximations'.

Almost all of Fitoussi's polemical writings and pamphleteering since about the mid-1990s, right up to his most recent publications (cf. [10], [11]), are best understood in this light. Fitoussi's basic point is that the kind of reality that, for example, Wigforss as a Minister of Finance faced, could not be tamed by the textbook paradigms that were available at that time. Given such a conundrum, the task of good government was not to enforce policies that would force institutions to conform to the straitjackets of the textbook paradigms but to enlarge, broaden and modify the textbook representations to reflect new realities.[29]

But Fitoussi provides further structure to this basic stance. As argued in one of his most recent papers, reflecting an underlying philosophy in the other writings and pamphlets of the last decade or so, and drawing on parallels for public or aggregative policies from one of Akerlof's imaginative pieces ([1]):

> Norms constraining public behaviour have rather their origin in some economic doctrine considered as true at a moment in time, regardless of its short-term economic consequences. In other words, 'Private social customs' are grounded on rules of fair behaviour, whereas 'Public-social customs' are supposed to be consistent with dominant economic doctrine.
>
> [13], p. 2

Thus, in addition to the simple observation that a 'dominant economic doctrine considered as true at a moment in time' may no longer be relevant at another 'moment in time', the truth of which seems beyond the 'triumvirate' to comprehend, Fitoussi broadens the underpinnings of his wide ranging criticisms of the pervasive economic and social policies of the EC and other national and international organs to include other elements that go beyond the capabilities of a narrow economic theory, of any variety, to be able to justify. His point here is that individual and social behaviour, whether by rational agents or by evolved or designed institutions, work within a framework of tacit norms, some of which are private and others are social. Either way:

> Whether of public or private origin, a social norm does not generally follow sound economic reasoning, being grounded on other motivations.
>
> ibid., p. 2

Who at the European Central Bank wonders where and how the economic doctrine that underpins the stability part of the straitjacket that has become the 'stability and growth pact' originated? How many within the hallowed halls of the Bank's offices in Frankfurt would take the trouble to understand the reasons why Wicksell devised his monetary policy norms,

by developing the framework of the *unstable cumulative process* to understand and tame the 20-year deflation of the late nineteenth century with a monetary policy rule that was underpinned by *his* enlightened amalgam of a reconstructed Austrian capital theory and an innovative monetary theory? Who within those Frankfurt walls and behind the glass façades in Brussels, can recall that it was for Wicksell, as it was for Irving Fisher, the implications on *income* and *wealth distributions* of the great twenty-year deflation of the late nineteenth century that were the great motivating forces for the theories of monetary stability they devised?

Jean-Paul Fitoussi, a consultant administrator and a pamphleteer, in addition to being an intellectual of great power and considerable reputation, and a university professor at a French institution of impeccable and international standing and fame, in the grand tradition that Schumpeter extolled, continues to take courageous stands and speak out against obdurate policies, the designing of untenable and unfair institutions and the ahistorical practices and narrow visions of policy makers and their handmaidens – whether it be outdated theories or ignorant advisors, whatever their reputation. Inequalities of every sort, the erosion of evolved practices of social cohesion, the preservation of traditions that foster social norms, the inequities that are foisted on unborn generations due to unenlightened interest rate policies, etc., are the broad canvasses on which he has begun to draw his visions for a just, democratic and prosperous society. I cannot but recall, in conclusion, the contextually apt and poignant words with which Keynes, from whom Fitoussi continues to draw much of his inspiration, himself ended his own great manifesto against the obduracy of an earlier generation ([21]), p. 383):

> [T]he ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed the world is ruled by little else. Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back. I am sure that the power of vested interests is vastly exaggerated compared with the gradual encroachment of ideas. Not, indeed, immediately, but after a certain interval; for in the field of economic and political philosophy there are not many who are influenced by new theories after they are twenty-five or thirty years of age, so that the ideas which civil servants and politicians and even agitators apply to current events are not likely to be the newest. But soon or late, it is ideas, not vested interests, which are dangerous for good or evil.

Precisely the message Fitoussi, as a social and economic philosopher, with a solid understanding of the history of economic theory, and its

contextual developments, has been trying to tell the 'civil servants, politicians and even agitators' who 'apply to current events' theories distilled 'from some academic scribbler of a few years back'.

A note on the origins of the name 'Fitoussi'

> *O, be some other name!*
> *What's in a name? That which we call a rose*
> *By any other name would smell as sweet.*
>
> *Romeo and Juliet*, Act 2, Sc. 1

Fifty years ago, Jean-Paul Fitoussi was living near Tunis in Tunisia; I was living in Colombo in what was then called Ceylon. About forty years ago, he was in Paris; I was in Tokyo. Thirty years ago, he was in Strasbourg; I was in Cambridge. Twenty years ago we were both together in Florence! Ten years ago, he was back in Paris, on his way to becoming an uncrowned king; I was on the fringes of Los Angeles. Now, here we are, on either side of an ocean! Such are the *serendipitous* effects of a kind of globalisation that is common to people of different kinds of diaspora.

'Serendip' is an old name for Ceylon, from which Walpole coined the word that has its modern significance as 'the faculty of discovering pleasing things by sagacity and chance'.[30]

I want to tell you a story about a name.

Many years ago, as a young boy in old Colombo, I saw a wonderful film, without understanding anything about its significance as a social commentary on a dying epoch and a disappearing order. It was Luchino Visconti's rendering of the book *Il Gattopardo*, renamed *The Leopard*. Many will recall Burt Lancaster as Fabrizio, Alain Delon as Tancredi and Claudia Cardinale as Angelica. I, of course, promptly forgot all about it and went about life and living with other interests.

Some years ago, my friend and colleague Elisabetta De Antoni rekindled my interest in the subject matter of *Il Gattopardo* by insisting, almost relentlessly, that I should read it. As is usual when I read books of such subtle themes, I decided to make myself acquainted with the background — social, intellectual, etc. – for the author of the book, Tomasi di Lampedusa. In reading about his life, in the book by David Gilmour, my thoughts also went back to other readings of Sicilian themes of Anglo-Saxon origin, particularly Raleigh Trevelyan's wonderful book *Princes Under the Volcano*.

More recently, when I began to wonder about the origin of the word 'fitoussi' my fading memory cells were rekindled in a most unusual way. I recalled Raleigh Trevelyan referring to Tina Whitaker,[31] in her diary, calling Tomasi di Lampedusa's grandfather, also called Giuseppe, *a piedi-fitusi*. This was requoted in the biography of Tomasi di Lampedusa by Gilmour which I had read very recently. I tracked down the reference in

Raleigh Trevelyan and began to try to find out how the suffix 'fitusi' originated.

Now, before I continue with my adventures in tracking down the origins of the word 'fitoussi', I want to sidetrack the reader for a moment.

Solow, in his 'Reminiscence and Ruminations' on having Richard Goodwin as a teacher, recalled a couple of interesting precepts that he had learned from that *maestro* ([31], pp. 32–3):

The first was on what it was that characterised a 'theorist's frame of mind':

> I seem to recall that he – Goodwin – sometimes suggested that, well, one could not actually believe this or that, but it was an ingenious line of thought, perhaps worth following just to see where it came out. One could always reject it later, and then one would have a better idea of what one was rejecting.

The second was about a particular form of 'intellectual style':

> The unspoken message was that if a thing was worth doing it is worth doing playfully. Do not misunderstand me: 'playful' does not mean 'frivolous' or 'unserious'. It means, rather, that one should follow a trail the way a puppy does, sniffing the ground, wagging one's tail, and barking a lot, because it *smells interesting* and it would be fun to see where it goes.
>
> (italics added)

I want to assure you that I followed, almost religiously, these two precepts in my pursuit of the origins of the word 'fitoussi' – and, in particular, 'rejected' one particular find, because it did not 'smell' interesting!

To return to my main story, the result of my speculative search for the origins of the word 'fitoussi' led me, naturally, to Sicilian dialects! There are two possibilities: either from the singular form of the noun, *fituso*, and its plural, *fitusi*, in usage in Sicilian dialects, and these, in turn, related to the Italian word '*fetore*'; or, my own preference, coming from the subjunctive form of the verb '*fiutare*' – *fiutassi*.

Take your pick! I know what I prefer!

Why does it seem more interesting to opt for this last alternative? Again, let me substantiate my own preference with an allegory and an analogy. Many of you will recall the famous episode when Colin Clarke, one of the pioneers of National Income Accounting, spent years collecting data and estimating the marginal propensity to consume for the English economy. After years of toil he went with his results to Keynes – who simply glanced at the values and told him that they could not be right and advised him to go back to the drawing boards. Keynes had a feel or a sense for what could be plausible numbers for the kind of economy England was, at that time.

Fitoussi is a modern Keynes or Wicksell in that same sense – and more. Not only, like Keynes and Wicksell, is Fitoussi a supreme pamphleteer and a consultant administrator in the noble Schumpeterian sense; he also has that 'good sense' (*avere buon fiuto per la macroeconomia*, as one might say in Italian) to intuit the relevant numbers that characterise a complex aggregate economy. Without this 'good sense' or a 'feel', one can only theorise in a vacuum. Hence, *FIUTARE* → *FIUTASSI* → *FITOUSSI*.

One can either believe Patinkin's sophisticated conjecture about the origins of the word in the distant Jabal Nefussa mountain ranges in Libya; or one can look at a school geography map or atlas and figure out for yourself how close Sicily and Lampedusa are to Tunisia and wonder about social intercourse between communities in the two societies. The Goodwin precepts, recalled above by Solow, point clearly to accepting the 'interesting conjecture' – whether it is true or not is quite another matter!

Notes

* This is very much my own reconstruction of the contours of Jean-Paul Fitoussi's general intellectual path, with some minor forays into details of a more personal nature. Naturally, neither he nor anyone else is responsible for the way I have reconstructed, in a general and simplified way, what must have been – and, surely, will continue to be – a complex, rich and varied professional and personal life. In particular, I do not even touch tangentially upon Fitoussi's deep concern and continuing important work on 'transition economies' and their almost intractable problems.

1 Few seem to be aware that the origins of the word 'serendipity' lie in that old name for Ceylon – Sri Lanka, now – *Serendip*. As concisely summarised in the OED, the word was formed by Horace Walpole, from the tale, *The Three Princes of Serendip*, via the allusions to their propensities and proclivities to make 'discoveries, by accidents and sagacity, of things they were not in quest of'. Naturally, I feel 'strongly' about the etymology of this enchanting word because I happen, serendipitously, to be from Serendip!

2 The telephone call was not a complete surprise. A week or so before the call, Richard Goodwin rang me from Siena and asked whether I would agree to move to Italy, near him, if the opportunity presented itself. Of course, my answer was in the affirmative. He knew that there was a vacancy at the EUI and had, obviously, made a case for me with some members of the economics department of the time. He – Goodwin – himself had moved to Siena to take up a professorship after retirement from Cambridge. When, eventually, Fitoussi offered me the post of Assistant Professor at the EUI, his qualifying caveat was: 'I want you here, in Florence, not in Siena! I know that Goodwin is in Siena.'

3 'A veritable *tour de force*', as a leading French macroeconomist described it, and an essay that entered into many a reading list in graduate courses at leading universities, on both sides of the Atlantic – and in other continents, as well.

4 I had first come to terms with this brilliant paper during my first term at Cambridge, in 1973. At that time I was not proficient in Italian but had the fortune to have a good friend, a contemporary at Cambridge, Guglielmo Chiodi, who took the enormous trouble of translating that paper into English. My acquain-

tance with the Kolmogorov paper was through a footnote in Minorsky's classic book, *Nonlinear Oscillations*, which Goodwin lent me after our first or second meeting. I mention these apparently irrelevant details simply because Fitoussi was mightily impressed by the general, nonlinear, model I presented at that seminar in Fiesole; but also because I was surprised that few in economics seem to have been familiar with that paper at that time – *circa* early 1970s.

5 A variation of which went into the structure of our joint paper of a few years later [14].

6 In particular, *Le Débat Interdit: Monnaie, Europe, Pauvreté* ([8]).

7 It is also not my intention to record, chronologically or otherwise, details and facts, such as years of graduations, lists of appointments and so on, that can be found in formal CVs and easily available 'Who's Whos'. That, the interested reader can investigate for herself.

8 W.B. Yeats: *Vacillation*.

9 For the Livornese–Tunisian–Jewish connection see [24] and [29].

10 Paul Coulbois wrote a generous Preface to Fitoussi's '*Thèse complémentaire*' of 1973 (cf. [5]).

11 It is, therefore, not surprising that his prize, doctoral dissertation, *Inflation, Équilibre et Chômage* ([4]), opens with a famous quote from Hegel's '*Principles of the Philosophy of Law*'. Fitoussi has often told me that Chamley would be tireless in his emphasis, following Hegel, that 'money is an abstraction' and talk of the rate of interest, in a Smithian vein, in the homely aphorism that its 'level measures our depreciation of the future'; that it required an enlightened institutional setting to make sense of it and, thus, the path from Adam Smith towards Georg Simmel was natural.

12 Christopher Brennan: *The Forest of Night*.

13 I recall with undiminished vividness the first time Fitoussi and I met Phelps and the occasion that led to that meeting, at the European University Institute, in Fiesole. During a regular staff meeting – we were a small department of only six members at that time, in the early 1980s – Manfred Streit received a telephone call from Ned Phelps, who was then visiting Germany and expressed interest in spending a few weeks with us, in Fiesole. Fitoussi, at that time, did not know that 'Ned' was the diminutive for 'Edmund' and immediately asked Streit whether this was *the* Phelps! That was the beginning of yet another friendship that led to much collaborative work.

14 Indeed, Bénassy had, partially, done his own doctoral work under Bent Hansen at UC, Berkeley.

15 I remember well that Haavelmo's little known paper in the *Econometrica* of 1950 was important, in these contexts, for Fitoussi ([20]).

16 There was, of course, that unfortunately neglected but analytically brilliant paper Fitoussi wrote, jointly with Georgescu-Roegen ([12]), which, on the surface seemed hostile to fix price macroeconomics of the French variety. Indeed, even a sympathetic discussant, of that paper, like Modigliani was led to conclude that the authors (ibid., p. 267):

> 'began by rejecting disequilibrium theories like those of Professor Malinvaud.'

But a close reading of that dense and innovative paper suggests that Fitoussi and Georgescu-Roegen took the Malinvaud framework as a benchmark from which to diverge and show how a disequilibrium theory with Schumpeterian (and Marshallian) elements could better account for the kind of unemployment that was prevalent at that time in the Western economies. Fitoussi and

Georgescu-Roegen were aiming to go beyond the tripartite division of the famous partition of the p-w space in [26] and suggest that periodic Western (mass) unemployment was a 'structural phenomenon', almost entirely due to novelties and innovations of a Schumpeterian variety. Of course, these 'structural' elements have remained an ingredient of Fitoussi's thinking ever since and is evident, as backdrops, in his scathing criticisms of labour market policies in contemporary European ruling circles.

17 I recall an interesting conversation between Fitoussi and Phelps, around May, 1983, in Paris. The three of us were walking back to Fitoussi's office on Quay d'Orsay, after lunch, and Fitoussi casually mentioned that he would be visiting UCLA for the Fall term of that year. Then the following dialogue took place:

> Phelps: 'What will you be teaching?'
> Fitoussi: 'Fixed-price macroeconomics.'
> Long pause and silence all around. Then:
> Phelps: 'And the second and subsequent lectures?'

Anyone even remotely familiar with the theoretical underpinnings of Fitoussi's works will know that he has never quite abandoned the basic macroeconomic framework that emerged in those years, from various French sources.

18 In [15], in which we were joint authors, and which may not have seen the light of day without the late Don Patinkin's unstinting support against scepticism by the editors of the *Journal of Money, Credit and Banking*. I must take this opportunity to mention that he – Don Patinkin – would, I am sure, have been an enthusiastic contributor to this volume had fortune been kinder with time. Patinkin was a constant visitor to the EUI during those enthusiastic years of the early 1980s, as, indeed, were many other of the contributors to this volume: Bénassy, Leijonhufvud, Malinvaud, Phelps and, of course, Le Cacheux, who was one of the early doctoral products from that institute, under Fitoussi.

19 Roughly, inverse to each other.

20 Using a result of Swinnerton–Dyer ([34]), we were able to suggest ways of estimating the size of a corridor of bifurcation, in the early drafts of that paper. The final, published version, had to eliminate that section on the advice of an eminent and sympathetic referee – Richard Goodwin – who thought it was adding disproportionate complications to an already lengthy and complex paper!

21 MIP: Money, Interest and Prices.

22 Again, partly, because an eminent and sympathetic referee advised us to 'downplay the slip' – but also because we felt that the machinery of homoclinic and heteroclinic dynamics would have been a disproportionate addition to a paper in a volume in honour of a great economist who was wary of unnecessary mathematics!

23 He was also, almost simultaneously with that move from Fiesole to Paris, appointed General Secretary of the prestigious International Economic Association.

24 W.H. Auden: *Secondary Epic*.

25 Myrdal's *Finanspolitikens Ekonomiska Verkningar* of 1933, the classic of this tradition, remains untranslated. It was prepared at the explicit request of the then Minister of Finance, Ernst Wigforss.

26 A story that was narrated to me by Mrs Gertrud Lindahl, when I was working on the 'Lindahl archives' which were then neither catalogued nor collected in any formal way at any library but were simply and neatly kept by her in a well-ordered way at her modest home in Lund. Whether this story was confirmed in one of Myrdal's memoirs, I no longer recall. I did meet and discuss such issues with Myrdal in summer 1981, but that was before I met Mrs Gertrud Lindahl.

27 The Swedish innovations on budgetary policy – as a supplement to their much better known novelties in stabilisation and monetary policies – which entailed, of course, much more than simple underbalancing of a budget, came about, partly as a result of the previously mentioned memorandum prepared by Myrdal.
28 At present these offices are occupied by Michael Schreyer, Mario Monti and Wim Duisenberg, respectively.
29 In such contexts I have often cited a characteristically perceptive observation on police behaviour, confronting potential criminals and a crime scene in *Great Expectations*, by Charles Dickens (italics added):

> They took up several wrong people, and they ran their heads very hard against wrong ideas and *persisted in trying to fit the circumstances to the ideas, instead of trying to extract ideas from circumstances*.

30 See footnote 1.
31 Tina Whitaker, née Scalia, was the wife of Giuseppe (Joseph) Whitaker. They built what is today known as the *Villa Malfitano*. There is also a 'Violet' (flower) named after Tina Whitaker, who found it growing in her garden in the Villa Malfitano, and sent a sap to friends in Kent, in England, around the turn of the twentieth century!

References

[1] Akerlof, George A. (1980) 'A Theory of Social Custom of Which Unemployment May be One Consequence', *The Quarterly Journal of Economics*, 94: 749–75.
[2] Arrow, Kenneth J. (1959) 'Toward a Theory of Price Adjustment', in: *The Allocation of Economic Resources: Essays in Honor of Bernard Francis Haley* edited by Moses Abramovitz, Stanford University Press, Stanford, CA.
[3] Clower, Robert W. (1965) 'The Keynesian Counter-Revolution: A Theoretical Appraisal', in: *The Theory of Interest Rates* edited by F. Brechling and F.H. Hahn. Macmillan, London.
[4] Fitoussi, Jean-Paul (1972) *Inflation, Équilibre et Chômage*, Editions Cujas, Paris.
[5] Fitoussi, Jean-Paul (1973) *Le Fondement Micro-Économique De La Théorie Keynésienne*, Editions Cujas, Paris.
[6] Fitoussi, Jean-Paul (1979) 'Emploi, Structure et Régulation', *Revue Economique*, No. 1.
[7] Fitoussi, Jean-Paul (1983) 'Modern Macroeconomic Theory: An Overview', in: *Modern Macroeconomic Theory* edited by Jean-Paul Fitoussi, Chapter 1, pp. 1–46. Basil Blackwell, Oxford.
[8] Fitoussi, Jean-Paul (1995) *Le Débat Interdit: Monnaie, Europe, Pauvreté*, Éditions Arléa, Paris.
[9] Fitoussi, Jean-Paul (1996) *Le Nouvel Âge des Inégalités* (jointly with Pierre Rosanvallon), Seuil, Paris.
[10] Fitoussi, Jean-Paul (2002) *La Règle et Le Choix: De La Souveraineté Économique en Europe*, Seuil, Paris.
[11] Fitoussi, Jean-Paul (2003) *Il Dittatore Benevolo: saggio sul Governo dell'-Europa*, Il Mulino, Bologna.
[12] Fitoussi, Jean-Paul and Nicholas Georgescu-Roegen (1980) 'Structure and Involuntary Unemployment', in: *Unemployment in Western Countries* edited by Edmond Malinvaud and Jean-Paul Fitoussi. Macmillan, London.

[13] Fitoussi, Jean-Paul and Francesco Saraceno (2002) 'A Theory of Social Custom of Which Soft Growth May be One Consequence: Tales of the European Stability Pact', *O.F.C.E. Discussion Paper* 2002–07, October.
[14] Fitoussi, Jean-Paul and Kumaraswamy Velupillai (1987) 'Growth and Cycles in Distributive Shares and Employment', in *Macro-Dynamique et Déséquilibres* edited by Jean-Paul Fitoussi and Pierre-Alain Muet, Chapter 3, pp. 31–54. Economica, Paris.
[15] Fitoussi, Jean-Paul and Kumaraswamy Velupillai (1987) 'Studies in Business-Cycle Theory: A Review Article', *Journal of Money, Credit and Banking*, 19 (1), February: 121–30.
[16] Fitoussi, Jean-Paul and Kumaraswamy Velupillai (1993) 'Macroeconomic Perspectives', in *Monetary Theory and Thought: Essays in Honour of Don Patinkin* edited by H. Barkai, S. Fischer and N. Liviatan. Macmillan, London.
[17] Fitoussi, Jean-Paul and Edmund S. Phelps (1988) *The Slump in Europe: Reconstructing Open Economy Theory*, Basil Blackwell, Oxford.
[18] Hahn, Frank H. (1965) 'On Some Problems Proving the Existence of Equilibrium in a Monetary Economy', in: *The Theory of Interest Rates* edited by F. Brechling and F.H. Hahn. Macmillan, London.
[19] Hansen, Bent (1951) *A Study in the Theory of Inflation*, George Allen & Unwin, Ltd, London.
[20] Haavelmo, Trygve (1950) 'The Notion of Involuntary Economic Decisions', *Econometrica*, 18, January.
[21] Keynes, John Maynard (1936) *The General Theory of Employment, Interest and Money*, Macmillan, London.
[22] Kolmogorov, A.N. (1936) 'Sulla Teoria Di Volterra Della Lotta Per L'Esistenza', *Giorn. dell' Istituto Italiano degli Attuari*, 14: 73–80.
[23] Leijonhufvud, Axel (1968) O*n Keynesian Economics and the Economics of Keynes*, Oxford University Press, Oxford.
[24] Levy, Lionel (1999) *Les Nation Juive Portugaise: Livourne, Amsterdam, Tunis, 1591–1951*, l'Harmattan, Paris.
[25] Lipsey, Richard G. (1968) 'The Place of the Phillips Curve in Macroeconomic Models', in: *Stability and Inflation: A Volume of Essays to Honour the Memory of A.W.H. Phillips* edited by A.R. Bergstrom, *et. al.*, pp. 49–75. John Wiley, Chichester.
[26] Malinvaud, Edmond (1977) *The Theory of Unemployment Reconsidered*, Basil Blackwell, Oxford.
[27] Patinkin, Don (1965) *Money, Interest and Prices: An Integration of Monetary and Value Theory* (2nd edition), Harper & Row, New York and London.
[28] Scitovsky, Tibor and Anne Scitovsky (1964) 'Inflation Versus Unemployment: An Examination of their Effects', in: *Inflation, Growth and Employment*, a series of research studies prepared for the Commission on Money and Credit, Prentice-Hall, Inc., New York.
[29] Sebag, Paul (2002) *Les Noms des Juifs de Tunisie*, l'Harmattan, Paris.
[30] Schumpeter, Joseph A. (1954) *History of Economic Analysis*, George Allen & Unwin Ltd, London.
[31] Solow, Robert (1990) 'Goodwin's Growth Cycle: Reminiscence and Rumination', in: *Nonlinear and Multisectoral Macrodynamics – Essays in Honour of Richard Goodwin*. Macmillan, London.
[32] Solow, Robert M. and Joseph E. Stiglitz (1967) 'Output, Employment and

Wages in the Short Run', *Quarterly Journal of Economics*, 82 (4), November: 537–60.

[33] Stern, Fritz (1977) *Gold and Iron: Bismarck, Bleichröder, and the Building of the German Empire*, Alfred A. Knopf Inc., New York.

[34] Swinnerton-Dyer, Peter (1977) 'The Hopf Bifurcation Theorem in Three Dimensions', *Mathematical Proceedings of the Cambridge Philosophical Society*, 82: 469–83.

3 Price versus wage stickiness and the issue of persistence

*Jean-Pascal Bénassy**

We compare in this chapter the relative abilities of price and wage staggering to produce output persistence. For that we construct a dynamic stochastic general equilibrium model, and integrate into it wage or price contracts à la Calvo ([3]). We derive explicit solutions for both the optimal wage and price contracts, and the resulting output and employment dynamics. It is found that usually, for the same average duration of contracts, price staggering will produce less persistence than wage staggering. The difference between the two depends crucially on the households' labor supply.

1 Introduction

Staggered wages and prices are two mechanisms that have been extensively used to build dynamic models displaying employment and output persistence[1]. Periodically the question arises of which one is the most suited to that purpose. Two issues are actually at stake. The first is an empirical one: between wages and price contracts, which ones are the most long lived? We shall not investigate this issue, but casual empiricism suggests that usually wages remain predetermined for longer periods than prices. The second issue, which is the one we shall address here, is theoretical: for a given contract life, which of wage and price contracts gives most persistence?

This issue was first tackled in an insightful paper by Andersen ([1]), who showed in a simple Taylor ([10], [11]) type model that staggered wages usually lead to more persistence than staggered prices. We shall extend this research in two directions: (a) We shall use Calvo ([3]) contracts, so that the average duration of contracts is not two periods, as in the simple Taylor model, but can take any positive value. (b) We shall build a full-fledged dynamic stochastic general equilibrium model embedding these contracts. In spite of these extensions we shall be able to obtain explicit solutions throughout. Another important aspect of this investigation is that we use exactly the same underlying structure for both price and wage staggering models. Our results will validate Andersen's ([1]) initial

insight: unless labor supply is highly elastic, price staggering produces less persistence than wage staggering[2].

2 The model

The economy studied is a monetary economy with markets for goods at the (average) price P_t and for labor at the (average) wage W_t. The goods and labor markets function under a system of imperfectly competitive contracts, which we now describe.

2.1 Wage and price contracts

Let us begin our description with the wage contracts. As in Calvo ([3]), in each period there is a random draw for all wage contracts, after which any particular contract will continue unchanged (with probability γ), or be terminated (with probability $1-\gamma$). In this last case the corresponding contract wage is renegotiated on the basis of all information currently available.

We shall denote by X_t the wage contract newly signed in period t (as we shall find out below, all workers who sign a new contract in t choose the same wage level, so that we do not need to index X_t by i or k).

Symmetrically price contracts are terminated stochastically with probability $1-\phi$, and continue unchanged with probability ϕ. We shall denote by Q_t the new price contract signed in period t.

2.2 The agents

Let us describe first the production side. The output index Y_t is an aggregate of a continuum of output types, indexed by $i \in [0, 1]$:

$$\operatorname{Log} Y_t = \int_0^1 \operatorname{Log} Y_{it} di \tag{1}$$

Each index Y_{it} is itself the aggregate of another infinity of output types indexed by k[3]:

$$Y_{it} = \left(\int_0^1 Y_{ikt}^{\theta} dk \right)^{1/\theta} \tag{2}$$

One can think of the index i as representing sectors, whereas the index k represents firms within these sectors. All firms in the same sector i face exactly the same situation in terms of price contracts, which means in particular that their contracts are renewed at the same time. On the contrary firms in different sectors may have signed their contracts at different times.

The representative firm has a Cobb–Douglas type technology:

$$Y_{ikt} = Z_t N_{ikt}^{\alpha} \tag{3}$$

where Z_t is a common technological shock.

The representative household (we omit the indexes i or k at this stage) works N_t, consumes C_t and ends the period with a quantity of money M_t. He maximizes the expected value of his discounted utilities, with the following intertemporal utility:

$$U = \Sigma \beta^t \left[\text{Log}\, C_t + \omega \text{Log} \frac{M_t}{P_t} - \xi \frac{N_t^{\nu}}{\nu} \right] \qquad \nu \geq 1 \tag{4}$$

At the beginning of period t there is a stochastic multiplicative monetary shock as in Lucas ([7]): money holdings carried from the previous period M_{t-1} are multiplied by the same factor μ_t for all agents, so that the representative household starts period t with money holdings $\mu_t M_{t-1}$. His budget constraints in t is thus:

$$C_t + \frac{M_t}{P_t} = \frac{W_t}{P_t} N_t + \frac{\mu_t M_{t-1}}{P_t} + \Pi_t \tag{5}$$

where Π_t are profits distributed.

3 The Walrasian regime

We shall now, as a benchmark for what follows, compute the Walrasian equilibrium of this economy. The real wage is equal to the marginal productivity of labor:

$$\frac{W_t}{P_t} = \frac{\partial Y_t}{\partial N_t} = \alpha \frac{Y_t}{N_t} \tag{6}$$

The households maximize the expected value of the utility function (4) subject to the budget constraints (5). The Lagrangean of this maximization program is:

$$\Sigma \beta^t \left[\text{Log}\, C_t + \omega \text{Log} \frac{M_t}{P_t} - \xi \frac{N_t^{\nu}}{\nu} \right]$$

$$+ \Sigma \beta^t \lambda_t \left[\frac{W_t N_t}{P_t} + \frac{\mu_t M_{t-1}}{P_t} + \Pi_t - C_t - \frac{M_t}{P_t} \right] \tag{7}$$

and the first order conditions:

$$\frac{1}{P_t C_t} = \frac{\omega}{M_t} + \beta E_t\left(\frac{\mu_{t+1}}{P_{t+1} C_{t+1}}\right) \tag{8}$$

$$\xi N_t^{\nu-1} = \frac{W_t}{P_t C_t} \tag{9}$$

Using the fact that $\mu_{t+1} = M_{t+1}/M_t$ equation (8) above can be solved as:

$$\frac{M_t}{P_t C_t} = \frac{\omega}{1-\beta} \tag{10}$$

We see further from (9) that Walrasian employment is constant and equal to N, which is given by:

$$N = \left(\frac{\alpha}{\xi}\right)^{1/\nu} \tag{11}$$

and the Walrasian wage W_t^* and price P_t^* are equal to:

$$W_t^* = \frac{\alpha(1-\beta)}{\omega}\left(\frac{\xi}{\alpha}\right)^{1/\nu} M_t \tag{12}$$

$$P_t^* = \frac{1-\beta}{\omega}\left(\frac{\xi}{\alpha}\right)^{\alpha/\nu} \frac{M_t}{Z_t} \tag{13}$$

4 The demand for goods and labor

We shall study below our model under wage or price contracts. It is assumed that households, possibly through trade unions, decide on the level of wages, and supply the amount of labor demanded by firms at these wages. Similarly firms set prices and supply the amount of goods demanded.

In this section we shall derive the demand for goods addressed to firms and the demand for labor addressed to households. Optimal wage contracts will be derived in section 5, optimal price contracts in section 6.

4.1 The demand for goods

At any time there may be a multiplicity of prices. This variety of prices can be due to two causes: first, there may be staggered prices, and thus there are different prices because price contracts have been signed at different points in time. Second, even if prices are fully flexible in each period, the workers in different firms may have different wage contracts, so that prices will differ even if all other economic conditions are the same.

Consider first the firms producing final output. They competitively maximize profits, i.e. they solve the following program:

$$\text{Max}\, P_t Y_t - \int_0^1 P_{it} Y_{it} di \qquad \text{s.t.} \qquad \int_0^1 \text{Log}\, Y_{it} di = \text{Log}\, Y_t$$

whose solution is:

$$Y_{it} = \frac{P_t Y_t}{P_{it}} \tag{14}$$

$$\text{Log}\, P_t = \int_0^1 \text{Log}\, P_{it} di \tag{15}$$

Now firms indexed by i will similarly maximize profits, i.e. they solve:

$$\text{Max}\, P_{it} Y_{it} - \int_0^1 P_{ikt} Y_{ikt} dk \qquad \text{s.t.} \qquad \left(\int_0^1 Y_{ikt}^{\theta} dk \right)^{1/\theta} = Y_{it}$$

whose solution is:

$$Y_{ikt} = Y_{it} \left(\frac{P_{ikt}}{P_{it}} \right)^{-1/(1-\theta)} \tag{16}$$

$$P_{it} = \left(\int_0^1 P_{ikt}^{-\theta/(1-\theta)} dk \right)^{-(1-\theta)/\theta} \tag{17}$$

Putting together equations (14) and (16) we obtain the expression of the demand for goods:

$$Y_{ikt} = \frac{P_t Y_t}{P_{it}} \left(\frac{P_{ikt}}{P_{it}} \right)^{-1/(1-\theta)} \tag{18}$$

An important thing to remember for what follows is that, in view of equation (14), all sectors have exactly the same value of sales:

$$P_{it} Y_{it} = P_t Y_t \qquad \forall i \tag{19}$$

4.2 The demand for labor

When studying staggered wage contracts we assume that the goods market has flexible prices. In that case each firm determines its optimal price, employment and production by solving the following program:

$$\text{Max}\, P_{ikt}Y_{ikt} - W_{ikt}N_{ikt} \qquad \text{s.t.}$$

$$Y_{ikt} = Z_t N_{ikt}^{\alpha} \text{ and } Y_{ikt} = \frac{P_t Y_t}{P_{it}}\left(\frac{P_{ikt}}{P_{it}}\right)^{-1/(1-\theta)}$$

This yields the first order condition:

$$W_{ikt}N_{ikt} = \alpha\theta P_{ikt}Y_{ikt} \tag{20}$$

Let us combine this relation with the production function (3) and the demand question (18). We obtain notably the demand for labor:

$$N_{ikt} = \left[\frac{\alpha\theta P_{it}^{\theta} Z_t^{\theta}(P_t Y_t)^{1-\theta}}{W_{ikt}}\right]^{1/(1-\alpha\theta)} \tag{21}$$

5 Optimal wage contracts

We shall now derive the optimal wage contracts. They are characterized through the following proposition:

Proposition 1: *The wage contract X_t signed in t is given by:*

$$X_t^{\nu} = \frac{\xi}{(\alpha\theta)^2}\left[\frac{\alpha\theta(1-\beta)}{\omega}\right]^{\nu}(1-\beta\gamma)\sum_{s\geq t}\beta^{s-t}\gamma^{s-t}E_t(M_s)^{\nu} \tag{22}$$

Proof: Household (i, k) maximizes his discounted expected utility. We shall consider here only the terms corresponding to the wage contracts signed at time t and still in effect at time s. Since contracts have a probability γ to survive each period, the contract signed in t has a probability γ^{s-t} to be still in effect in period s, and the household (i, k) will thus maximize the following expected utility:

$$E_t\sum_{s\geq t}\beta^{s-t}\gamma^{s-t}\left[\text{Log}\, C_{iks} + \omega\text{Log}\frac{M_{iks}}{P_s} - \frac{\xi N_{iks}^{\nu}}{\nu}\right] \tag{23}$$

subject to the budget constraints in each period:

$$C_{iks} + \frac{M_{iks}}{P_s} = \frac{X_{ikt}}{P_s}N_{iks} + \frac{\mu_s M_{iks-1}}{P_s} + \Pi_s \tag{24}$$

and the equation giving the demand for labor:

$$N_{iks} = \left[\frac{\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}}{X_{ikt}}\right]^{1/(1-\alpha\theta)} \tag{25}$$

Inspecting this maximization problem we first see that all households with the same index i face exactly the same circumstances, so that in equilibrium:

$$X_{ikt} = X_{it} \qquad \forall k \tag{26}$$

Moreover from equations (19) and (20):

$$W_{is}N_{is} = \alpha\theta P_{is}Y_{is} = \alpha\theta P_s Y_s \tag{27}$$

So all consumers have the same income, and therefore the same consumption and money holdings (but they differ, of course, in their wages and employment levels):

$$C_{iks} = C_s \qquad M_{iks} = M_s \qquad \forall i, k \tag{28}$$

Households indexed by (i, k) maximize (23) subject to (24) and (25). Let us insert the value of N_{iks} (equation (25)) into (23) and (24). Taking into account (28) the corresponding Lagrangean is written (we omit the terms that are not used):

$$E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t}\left\{\text{Log}\, C_s - \frac{\xi}{\nu}\left[\frac{\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}}{X_{ikt}}\right]^{\nu/(1-\alpha\theta)}\right\}$$

$$+ E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t}\lambda_{iks}\left\{[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{1/(1-\alpha\theta)} \frac{X_{ikt}^{-\alpha\theta/(1-\alpha\theta)}}{P_s} - C_s\right\} \tag{29}$$

Maximization in C_s yields:

$$\lambda_{iks} = \frac{1}{C_s} \tag{30}$$

so that the term in X_{ikt} is, suppressing unimportant constants:

$$X_{ikt}^{-\alpha\theta/(1-\alpha\theta)} E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t} \frac{1}{P_s C_s}[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{1/(1-\alpha\theta)}$$

$$-\frac{\xi}{\nu} X_{ikt}^{-\nu/(1-\alpha\theta)} E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t}[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{\nu/(1-\alpha\theta)} \tag{31}$$

The first order condition in X_{ikt} is:

$$\alpha\theta X_{ikt}^{-1/(1-\alpha\theta)} E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t} \frac{1}{P_s C_s}[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{1/(1-\alpha\theta)}$$

$$= \xi X_{ikt}^{-\nu/(1-\alpha\theta)-1} E_t \sum_{s\geq t} \beta^{s-t}\gamma^{s-t}[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{\nu/(1-\alpha\theta)} \tag{32}$$

Now we know from above (26) that in equilibrium $X_{ikt} = X_{it}$ for all k. So all firms who share the same index i will choose the same price $P_{iks} = P_{is}$. From formulas (3), (19) and (20) we compute it as:

$$P_{is} = \frac{(P_s Y_s)^{1-\alpha}}{Z_s}\left(\frac{X_{it}}{\alpha\theta}\right)^{\alpha} \tag{33}$$

From this we deduce:

$$[\alpha\theta P_{is}^{\theta} Z_s^{\theta}(P_s Y_s)^{1-\theta}]^{1/(1-\alpha\theta)} = \alpha\theta P_s Y_s X_{it}^{\alpha\theta/(1-\alpha\theta)} \tag{34}$$

Inserting this into the first order condition (32) we obtain:

$$\frac{(\alpha\theta)^2}{1-\beta\gamma} X_{it}^{\nu} = \xi \sum_{s\geq t} \beta^{s-t}\gamma^{s-t} E_t(\alpha\theta P_s Y_s)^{\nu} \tag{35}$$

We see that the solution in X_{it} is the same for all agents i and we denote it as X_t:

$$X_t^{\nu} = \frac{\xi}{(\alpha\theta)^2}(1-\beta\gamma)\sum_{s\geq t} \beta^{s-t}\gamma^{s-t} E_t(\alpha\theta P_s Y_s)^{\nu} \tag{36}$$

from which, using equation (10), we immediately obtain equation (22).

Q.E.D.

6 Optimal price contracts

We now assume that the labor market clears competitively, but that firms make price contracts. The optimal price contracts are characterized through the following proposition:

Proposition 2: *The price contract Q_t signed in period t is given by:*

$$Q_t^{1/\alpha} = \frac{\xi}{\alpha\theta}\left(\frac{1-\beta}{\omega}\right)^{\nu/\alpha}(1-\beta\phi)\sum_{s\geq t} \beta^{s-t}\phi^{s-t} E_t\left[\left(\frac{M_s}{P_s Z_s}\right)^{\nu/\alpha} P_s^{1/\alpha}\right] \tag{37}$$

Proof: Firm (i, k) maximizes its discounted expected profits weighed by the marginal utility of goods. We shall consider here only the terms corresponding to the price contracts signed at time t and still in effect at time s. Since contracts have a probability ϕ to survive each period, the contract

signed in t has a probability ϕ^{s-t} to be still in effect in period s, and the household maximizes the following expected profit:

$$E_t \sum_{s=t}^{\infty} \beta^{s-t} \phi^{s-t} \frac{1}{P_s C_s} (P_{iks} Y_{iks} - W_s N_{iks}) \tag{38}$$

subject to the equation giving the demand for goods:

$$Y_{iks} = \frac{P_s Y_s}{Q_{it}} \left(\frac{Q_{ikt}}{Q_{it}} \right)^{-1/(1-\theta)} \tag{39}$$

where Q_{ikt} is the new price contract to be determined and Q_{it} the average price contract of all other firms indexed by i. In view of (17), this is equal to:

$$Q_{it} = \left(\int_0^1 Q_{ikt}^{-\theta/(1-\theta)} dk \right)^{-(1-\theta)/\theta} \tag{40}$$

Inspecting this maximization problem we first see that all firms with the same index i face exactly the same circumstances, so that at equilibrium we will have:

$$Q_{ikt} = Q_{it} \qquad \forall k \tag{41}$$

Firms indexed by (i, k) maximize (38) subject to (39). Let us insert the value of Y_{iks} (equation (39)) into (38). Taking into account $C_s = Y_s$ and $N_{iks} = (Y_{iks}/Z_s)^{1/\alpha}$, the maximand is written:

$$E_t \sum_{s \geq t} \beta^{s-t} \phi^{s-t} \left[\left(\frac{Q_{ikt}}{Q_{it}} \right)^{1-1/(1-\theta)} - \frac{W_s}{P_s Y_s} \left(\frac{P_s Y_s}{Z_s Q_{it}} \right)^{1/\alpha} \left(\frac{Q_{ikt}}{Q_{it}} \right)^{-1/\alpha(1-\theta)} \right] \tag{42}$$

The first order condition in Q_{ikt} is:

$$E_t \sum_{s \geq t} \beta^{s-t} \phi^{s-t} \left[\theta \left(\frac{Q_{ikt}}{Q_{it}} \right)^{-1/(1-\theta)} - \frac{W_s}{\alpha P_s Y_s} \left(\frac{P_s Y_s}{Z_s Q_{it}} \right)^{1/\alpha} \left(\frac{Q_{ikt}}{Q_{it}} \right)^{-1/\alpha(1-\theta)-1} \right] = 0 \tag{43}$$

Now we know from above (41) that in equilibrium $Q_{ikt} = Q_{it}$ for all k, so this becomes:

$$\alpha\theta Q_{it}^{1/\alpha} = (1 - \beta\phi) E_t \sum_{s \geq t} \beta^{s-t} \phi^{s-t} \left[\frac{W_s}{P_s Y_s} \left(\frac{P_s Y_s}{Z_s} \right)^{1/\alpha} \right] \tag{44}$$

We see that Q_{it} is the same for all sectors i, and we denote it Q_t:

$$Q_t^{1/\alpha} = \frac{1 - \beta\phi}{\alpha\theta} \sum_{s \geq t} \beta^{s-t} \phi^{s-t} E_t \left[\frac{W_s}{P_s Y_s} \left(\frac{P_s Y_s}{Z_s} \right)^{1/\alpha} \right] \tag{45}$$

which, using equation (10), yields:

$$Q_t^{1/\alpha} = \frac{1-\beta\phi}{\alpha\theta}\left(\frac{1-\beta}{\omega}\right)^{(1-\alpha)/\alpha}\sum_{s\geq t}\beta^{s-t}\phi^{s-t}E_t[W_s M_s^{(1-\alpha)/\alpha}Z_s^{-1/\alpha}] \tag{46}$$

Now combining equations (3), (9) and (10) we obtain:

$$W_s = \frac{\xi(1-\beta)}{\omega}\left(\frac{1-\beta}{\omega}\right)^{(\nu-1)/\alpha}\left(\frac{M_s}{P_s Z_s}\right)^{(\nu-1)/\alpha}M_s \tag{47}$$

Inserting (47) into (46) yields equation (37). Q.E.D.

7 Average prices and wages

We shall now show how the average price and wage P_t and W_t are deduced from the optimal contracts Q_t and X_t that we computed above.

7.1 The average price

Let us start with the preset prices case. We shall derive the average price index P_t as a function of all contracts signed in the past, Q_s, $s \leq t$. We first have the general formula (15) giving P_t:

$$\text{Log}\, P_t = \int_0^1 \text{Log}\, P_{it}\, di \tag{48}$$

Because of the law of large numbers, and since price contracts survive with probability ϕ, a proportion $1-\phi$ of the price contracts comes from period t, a proportion $\phi(1-\phi)$ from period $t-1, \ldots,$ a proportion $\phi^{t-s}(1-\phi)$ from period s, and so on. As a result formula (48) is rewritten:

$$\text{Log}\, P_t = (1-\phi)\sum_{s=-\infty}^{t}\phi^{t-s}\text{Log}\, Q_s \tag{49}$$

7.2 The average wage

For the preset wages case we shall similarly define an aggregate wage index W_t as:

$$\text{Log}\, W_t = \int_0^1 \text{Log}\, W_{it}\, di \tag{50}$$

Because of the law of large numbers, and in view of the survival rate γ for wage contracts, a proportion $1-\gamma$ of the wage contracts comes from

period t, a proportion $\gamma(1-\gamma)$ from period $t-1, \ldots,$ a proportion $\gamma^{t-s}(1-\gamma)$ from period s, and so on. As a result formula (50) is rewritten:

$$\text{Log}\, W_t = (1-\gamma)\sum_{s=-\infty}^{t} \gamma^{t-s} \text{Log}\, X_s \tag{51}$$

8 Macroeconomic dynamics with preset wages

Let us first assume that wages are determined under a system of Calvo contracts with probability γ. We characterize the dynamics of wages and employment in the following proposition:

Proposition 3: *Assume that the monetary process is:*

$$m_t - m_{t-1} = \frac{u_t}{1-\rho L} \tag{52}$$

where u_t is a white noise. Under this monetary process wages and employment are, up to a constant, given by:

$$w_t = m_t - \frac{\gamma(1-\beta\rho)u_t}{(1-\beta\gamma\rho)(1-\gamma L)(1-\rho L)} \tag{53}$$

$$n_t = \frac{\gamma(1-\beta\rho)u_t}{(1-\beta\gamma\rho)(1-\gamma L)(1-\rho L)} \tag{54}$$

Proof: From formula (51) the average wage is an average of the past new contracts:

$$w_t = \frac{1-\gamma}{1-\gamma L} x_t \tag{55}$$

So let us compute the new contracts x_t. Loglinearizing formula (22) we obtain:

$$x_t = (1-\beta\gamma)\sum_{s=t}^{\infty} \beta^{s-t}\gamma^{s-t} E_t m_s = (1-\beta\gamma)\sum_{j=0}^{\infty} \beta^j \gamma^j E_t m_{t+j} \tag{56}$$

Now let us rewrite m_{t+j}:

$$m_{t+j} = m_t + \frac{u_{t+1}}{1-\rho L} + \ldots + \frac{u_{t+j}}{1-\rho L} \tag{57}$$

so that:

$$E_t m_{t+j} = m_t + \frac{\rho u_t}{1-\rho L} + \ldots + \frac{\rho^j u_t}{1-\rho L} = m_t + \frac{\rho(1-\rho^j)u_t}{(1-\rho)(1-\rho L)} \tag{58}$$

Inserting (58) into (56) we find:

$$x_t = (1-\beta\gamma)\sum_{j=0}^{\infty}\beta^j\gamma^j\left[m_t + \frac{\rho(1-\rho^j)u_t}{(1-\rho)(1-\rho L)}\right]$$

$$= m_t + \frac{\beta\gamma\rho u_t}{(1-\beta\gamma\rho)(1-\rho L)} \tag{59}$$

Now let us return to formula (55) and rewrite it:

$$w_t - m_t = \frac{1-\gamma}{1-\gamma L}(x_t - m_t) + \frac{1-\gamma}{1-\gamma L}m_t - m_t$$

$$= \frac{1-\gamma}{1-\gamma L}(x_t - m_t) - \frac{\gamma(1-L)}{1-\gamma L}m_t \tag{60}$$

Inserting the value of x_t we found (equation (59)) into (60) we obtain:

$$w_t = m_t + \frac{1-\gamma}{1-\gamma L}\frac{\beta\gamma\rho u_t}{(1-\beta\gamma\rho)(1-\rho L)} - \frac{\gamma u_t}{(1-\gamma L)(1-\rho L)}$$

$$= m_t - \frac{\gamma(1-\beta\rho)u_t}{(1-\beta\gamma\rho)(1-\gamma L)(1-\rho L)} \tag{61}$$

$$n_t = m_t - w_t = \frac{\gamma(1-\beta\rho)u_t}{(1-\beta\gamma\rho)(1-\gamma L)(1-\rho L)} \tag{62}$$

These are formulas (53) and (54). Q.E.D.

9 Macroeconomic dynamics with preset prices

Let us now move to the case of staggered Calvo prices with probability ϕ. Loglinearizing equation (37) giving the new price contracts we obtain:

$$q_t = \frac{1-\beta\phi}{1-\beta\phi L^{-1}}[\nu(m_t - p_t - z_t) + p_t] \tag{63}$$

On the other hand we also have equation (49) giving the price as a function of past new contracts:

$$p_t = \frac{1-\phi}{1-\phi L}q_t \tag{64}$$

We shall concentrate here on the effects of monetary shocks, and thus set $z_t = 0$. With the help of equations (63) and (64) we can now characterize the dynamics of prices and output:

Proposition 4: *Assume again that the monetary process is:*

$$m_t - m_{t-1} = \frac{u_t}{1 - \rho L} \tag{65}$$

Prices and output are given by:

$$m_t - p_t = y_t = \frac{(1 - \beta\rho)\psi u_t}{(1 - \beta\rho\psi)(1 - \rho L)(1 - \psi L)} \tag{66}$$

where ψ is the smallest root solution of:

$$\Phi(\psi) = \beta\phi\psi^2 - (1 + \beta\phi^2)\psi - (1 - \phi)(\nu - 1)(1 - \beta\phi)\psi + \phi = 0 \tag{67}$$

Proof: Let us make the hypothesis:

$$m_t - p_t = \frac{au_t}{1 - \psi L} + \frac{bu_t}{1 - \rho L} \tag{68}$$

where a, b and ψ are to be determined below. From formula (63):

$$\begin{aligned} q_t - m_t &= \frac{(\nu - 1)(1 - \beta\phi)}{1 - \beta\phi L^{-1}}(m_t - p_t) + \frac{1 - \beta\phi}{1 - \beta\phi L^{-1}} m_t - m_t \\ &= \frac{(\nu - 1)(1 - \beta\phi)}{1 - \beta\phi L^{-1}}\left[\frac{au_t}{1 - \psi L} + \frac{bu_t}{1 - \rho L}\right] + \frac{\beta\phi L^{-1}(1 - L)m_t}{1 - \beta\phi L^{-1}} \\ &= (\nu - 1)(1 - \beta\phi)\left[\frac{au_t}{(1 - \beta\phi\psi)(1 - \psi L)} + \frac{bu_t}{(1 - \beta\phi\rho)(1 - \rho L)}\right] \\ &\quad + \frac{\beta\phi\rho u_t}{(1 - \beta\phi\rho)(1 - \rho L)} \end{aligned} \tag{69}$$

Now we have:

$$\begin{aligned} p_t - m_t &= \frac{1 - \phi}{1 - \phi L}(q_t - m_t) + \frac{1 - \phi}{1 - \phi L} m_t - m_t \\ &= \frac{1 - \phi}{1 - \phi L}(q_t - m_t) - \frac{\phi u_t}{(1 - \phi L)(1 - \rho L)} \end{aligned} \tag{70}$$

Combining (69) and (70), we obtain:

$$\begin{aligned} m_t - p_t = & \frac{\phi u_t}{(1-\phi L)(1-\rho L)} - \frac{\beta\phi\rho(1-\phi)u_t}{(1-\beta\phi\rho)(1-\phi L)(1-\rho L)} \\ & - \frac{(1-\phi)(\nu-1)(1-\beta\phi)}{1-\phi L}\left[\frac{a u_t}{(1-\beta\phi\psi)(1-\psi L)}\right. \\ & \left. + \frac{b u_t}{(1-\beta\phi\rho)(1-\rho L)}\right] \\ = & \frac{\phi(1-\beta\rho) - b(1-\phi)(\nu-1)(1-\beta\phi)}{(1-\beta\phi\rho)(1-\phi L)(1-\rho L)} \\ & - \frac{a(1-\phi)(\nu-1)(1-\beta\phi)}{(1-\beta o\psi)(1-\phi L)(1-\psi L)} \end{aligned} \tag{71}$$

This is of the form:

$$m_t - p_t = \frac{A u_t}{(1-\phi L)(1-\psi L)} + \frac{B u_t}{(1-\phi L)(1-\rho L)} \tag{72}$$

with:

$$A = -\frac{a(1-\phi)(\nu-1)(1-\beta\phi)}{1-\beta\phi\psi} \tag{73}$$

$$B = \frac{\phi(1-\beta\rho) - b(1-\phi)(\nu-1)(1-\beta\phi)}{1-\beta\phi\rho} \tag{74}$$

We make the identification:

$$\frac{A u_t}{(1-\phi L)(1-\psi L)} + \frac{B u_t}{(1-\phi L)(1-\rho L)} = \frac{a u_t}{1-\psi L} + \frac{b u_t}{1-\rho L} \tag{75}$$

which leads to the three conditions:

$$A\psi = a(\psi - \phi) \tag{76}$$

$$B\rho = b(\rho - \phi) \tag{77}$$

$$A(\rho - \phi) + B(\psi - \phi) = 0 \tag{78}$$

Combining these three equations yields:

$$\frac{\alpha}{\psi}+\frac{b}{\rho}=0 \tag{79}$$

Now inserting the values of A and B in (73) and (74) into (76), (77) and (79) yields respectively:

$$(\psi-\phi)(1-\beta\phi\psi)+(1-\phi)(\nu-1)(1-\beta\phi)\psi=0 \tag{80}$$

$$b=\frac{\phi\rho(1-\beta\rho)}{(\rho-\phi)(1-\beta\phi\rho)+(1-\phi)(\nu-1)(1-\beta\phi)\rho} \tag{81}$$

$$a=-\frac{\phi\psi(1-\beta\rho)}{(\rho-\phi)(1-\beta\phi\rho)+(1-\phi)(\nu-1)(1-\beta\phi)\rho} \tag{82}$$

The first equation, which yields ψ, is rewritten:

$$\Phi(\psi)=\beta\phi\psi^2-(1+\beta\phi^2)\psi-(1-\phi)(\nu-1)(1-\beta\phi)\psi+\phi=0 \tag{83}$$

which is equation (67) above. Now using (83) the values of a and b (equations 81 and 82) become:

$$a=\frac{(1-\beta\rho)\psi^2}{(\psi-\rho)(1-\beta\psi\rho)} \tag{84}$$

$$b=-\frac{(1-\beta\rho)\psi\rho}{(\psi-\rho)(1-\beta\psi\rho)} \tag{85}$$

Inserting these into equation (68), we find the final expression for output:

$$y_t=\frac{(1-\beta\rho)\psi}{(1-\beta\psi\rho)(\psi-\rho)}\left[\frac{\psi u_t}{1-\psi L}-\frac{\rho u_t}{1-\rho L}\right] \tag{86}$$

$$y_t=\frac{(1-\beta\rho)\psi u_t}{(1-\beta\psi\rho)(1-\psi L)(1-\rho L)} \tag{87}$$

which is equation (66). Q.E.D.

10 Comparing price and wage contracts

10.1 A summary

To compare price and wage staggering, let us first summarize briefly our results. We ignore technology shocks and concentrate on monetary shocks of the form:

$$m_t - m_{t-1} = \frac{u_t}{1 - \rho L} \tag{88}$$

If wage contracts are characterized by the parameter γ, corresponding to an average contract duration of $\gamma/(1-\gamma)$ periods, output dynamics is given by (we eliminate the constants):

$$y_t = \frac{\alpha\gamma(1-\beta\rho)u_t}{(1-\beta\gamma\rho)(1-\gamma L)(1-\rho L)} \tag{89}$$

Similarly, if price contracts are characterized by the parameter ϕ, corresponding to an average contract duration of $\phi/(1-\phi)$ periods, output dynamics is given by:

$$y_t = \frac{\psi(1-\beta\rho)u_t}{(1-\beta\psi\rho)(1-\psi L)(1-\rho L)} \tag{90}$$

where ψ is the smallest root solution of:

$$\Phi(\psi) = \beta\phi\psi^2 - (1+\beta\phi^2)\psi - (1-\phi)(\nu-1)(1-\beta\phi)\psi + \phi = 0 \tag{91}$$

10.2 The comparison

If we want to compare the two contracts for equivalent durations, i.e. for $\gamma = \phi$, we see that a central element of the comparison will be the size of the root ψ, which intervenes in the price contracts dynamics, as compared to the root γ which intervenes in wage contracts dynamics.

In particular a small root ψ has two effects in the preset price dynamics: (a) It reduces the impact of a monetary shock on output. (b) It reduces the persistence of the effects of the shock. So it is important to see what are the factors that affect ψ. We shall use for that the characteristic polynomial $\Phi(\psi)$ giving ψ (equation (91)). We may first compute:

$$\Phi(0) = \phi > 0 \tag{92}$$

$$\Phi(\phi) = -(1-\phi)(\nu-1)(1-\beta\phi)\phi \leq 0 \tag{93}$$

We thus have one positive root between 0 and ϕ. Since the product of the roots is $1/\beta$, the other root is bigger than one. So $0 < \psi < \phi$.

10.3 Labor disutility and dampening

We shall now show that a sufficiently convex disutility of labor can reduce this root ψ as much as one wants. More precisely we want to show that by

making ν big enough, one can make ψ as small as wanted. This is made precise in the following proposition:

Proposition 5: *The root ψ is characterized by:*

$$0 \leq \psi \leq \frac{\phi}{1 + (\nu - 1)(1 - \phi)(1 - \beta\phi)} \tag{94}$$

Proof: Let us recall the equation giving ψ:

$$\Phi(\psi) = \beta\phi\psi^2 - (1 + \beta\phi^2)\psi - (1 - \phi)(\nu - 1)(1 - \beta\phi)\psi + \phi = 0 \tag{95}$$

We want to prove that the root ψ is smaller than $\epsilon\phi$ with:

$$\epsilon = \frac{1}{1 + (\nu - 1)(1 - \phi)(1 - \beta\phi)} \leq 1 \tag{96}$$

We can compute indeed that:

$$\Phi(\epsilon\phi) = \beta\phi^3\epsilon(\epsilon - 1) \leq 0 \tag{97}$$

Q.E.D.

11 Conclusion

The objective of this chapter was to compare prices and wage staggering from the point of view of output persistence. For that purpose we constructed a rigorous dynamic stochastic general equilibrium model, and studied its dynamics under two alternative assumptions: (a) Calvo wage contracts, (b) Calvo price contracts. We found that usually, for the same average duration, price contracts deliver less persistence than wage contracts. We saw in propositions 4 and 5 that this was due to a dampening of the effects of demand shocks in the case of price contracts. This dampening is stronger, the less elastic the labor supply. The mechanism, which had been uncovered by Andersen ([1]), can be briefly summarized as follows: If a positive demand shock hits the economy, firms will demand more labor to satisfy it. If labor supply is inelastic, this will lead to a large wage increase on the fraction of labor markets which clear. This wage increase will itself substantially dampen the effect of the initial shock. This effect appears particularly clearly in our model since, except for the difference between wage and price contracts, the two versions of our dynamic model are exactly the same.

Notes

* I wish to thank Torben Andersen for his comments on an earlier version. Of course I am responsible for all remaining deficiencies.

1 The initial contributions are Gray ([6]), Fischer ([5]), Phelps–Taylor ([9]), Phelps ([8]), Taylor ([10], [11]) and Calvo ([3]).

2 Other comparisons of price and wage staggering are found in Ascari–Garcia ([2]) and Edge ([4]).

3 We should use the more precise notation k_i instead of k, but this would complicate the notation without much gain in understanding.

References

[1] Andersen, Torben M. (1998) 'Persistency in Sticky Price Models', *European Economic Review* 42: 593–603.

[2] Ascari, Guido and Juan Angel Garcia (1999) *Price/Wage Staggering and Persistence*, European University Institute, Firenze.

[3] Calvo, Guillermo (1983) 'Staggered Prices in a Utility-Maximizing Framework', *Journal of Monetary Economics* 12: 383–98.

[4] Edge, Rochelle M. (2002) 'The Equivalence of Wage and Price Staggering in Monetary Business Cycles' Models', *Review of Economic Dynamics* 5: 559–85.

[5] Fischer, Stanley (1977) 'Long Term Contracts, Rational Expectations and the Optimal Money Supply Rule', *Journal of Political Economy* 85: 191–205.

[6] Gray, Jo-Anna (1976) 'Wage Indexation: A Macroeconomic Approach', *Journal of Monetary Economics* 2: 211–35.

[7] Lucas, Robert E. Jnr (1972) 'Expectations and the Neutrality of Money', *Journal of Economic Theory* 4: 103–24.

[8] Phelps, Edmund S. (1978) 'Disinflation Without Recession: Adaptive Guideposts and Monetary Policy', *Weltwirtschaftliches Archiv* 114: 783–809.

[9] Phelps, Edmund S. and John B. Taylor (1977) 'Stabilizing Powers of Monetary Policy Under Rational Expectations', *Journal of Political Economy* 85: 163–90.

[10] Taylor, John B. (1979) 'Staggered Wage Setting in a Macro Model', *American Economic Review* 69: 108–13.

[11] Taylor, John B. (1980) 'Aggregate Dynamics and Staggered Contracts', *Journal of Political Economy* 88: 1–23.

4 Contours of employment protection reform

*Olivier Blanchard and Jean Tirole**

Introduction

There may be no labor market institution more controversial than employment protection regulation – the set of laws and procedures regulating separations between firms and workers.

- Firms complain not only about the direct cost, but also about the complexity and the uncertainty introduced by such regulation. They argue that it makes it difficult for them to adjust to changes in technology and product demand, and that this in turn decreases efficiency, increases cost and, in so doing, deters job creation.
- Workers, on the other hand, focus on the pain of unemployment, and argue that such pain should be taken into account by firms when they consider closing a plant, or laying off a worker. That workers protected by employment protection would favor it is no great surprise. But evidence from surveys shows that support for employment protection is more general, more broad based.
- Many economists and international economic organizations, from the OECD to the IMF, tend to side with firms. There is, they argue, a trade-off between insurance and efficiency. The current system impedes reallocation, and, by implication, reduces efficiency. It leads to higher costs, and thus lower employment. At a minimum, it could and should be made more efficient. More likely, overall employment protection should be reduced.
- Faced with conflicting advice and demands, the governments of Western Europe have been prudent (or timid, depending on one's point of view). They have learned, often the hard way, that workers covered by employment protection are not eager to see it reduced, and that these workers represent the majority of the labor force, and a large part of the electorate. So, most if not all of the recent employment protection reforms have worked at the margin, through the introduction and extension of the scope for fixed duration contracts – contracts subject to more limited employment protection and simpler

administrative rules. For the most part, employment protection for regular contracts has remained unchanged. The evidence so far is that this dual system has led to an increasingly dual labor market, with mixed efficiency and distributional effects.

Despite the heat and the rhetoric, we are struck by how little work has gone into the question of how "good employment protection regulation" should look. Starting from the status quo, firms and international organizations have argued for less protection. Workers and unions have fought to keep the protection they had. Governments have looked for politically feasible incremental reforms. But the ultimate goal, the shape of optimal employment protection, has been left undefined.

Consider for example the following questions:

- Should there be any state mandated employment protection, or should "employment at will" remain the principle, leaving any additional protection to voluntary agreements between workers and firms?
- If there is an argument for state mandated employment protection, should this protection simply take the form of a schedule of payments by firms in case of layoffs, with the layoff decision then left to the firm? Or should there be, in place or in addition to such a schedule, other non-price restrictions? In that context, what should be the role of the judicial process, if any?
- How large should payments by firms either to workers or to the state be? Should firms pay workers directly, or should they pay the state? Should the payments cover, in expected value or in realization, the unemployment benefits and other payments received by laid-off workers? Should the payments be made by firms at the time the layoffs take place, or should they be paid over time, as in the case in experience rated systems?

This chapter's primary purpose is to answer these questions, and to apply the answers to draw the contours of institutional reform in the French context.

We organize our discussion by starting from a simple benchmark. In that benchmark, firms and the state are risk neutral. Workers are risk averse and cannot fully self-insure against unemployment. In that context, characterizing optimal employment protection is straightforward. Firms must be made to internalize the cost of unemployment. So, if for example, unemployment benefits are administered by an unemployment agency, firms must pay contributions to the agency equal to the present value of unemployment benefits paid by the agency to the workers they lay off. Put another way, the contribution rate of firms, defined as the ratio of contributions paid by the firm to the benefits received by the worker, must be equal to one. In that sense, unemployment insurance and employment

protection are both integral parts of the optimal set of labor market institutions.

This benchmark, like all benchmarks, is both useful and too simple. The labor market suffers from many imperfections, and most of them impinge on the nature of optimal labor market institutions. These imperfections range from the need to give the unemployed incentives to search for a job, to the scope for ex-post renegotiations of wages, to liquidity constraints faced by firms. Each of these imperfections affects the optimal contribution rate, and thus the optimal degree of employment protection. But the general principle remains, that of making firms internalize, to the extent possible, the social cost of unemployment.

Turning to actual institutions, it is clear that this principle is at odds with the French system of employment protection in at least two main dimensions.

First, in the French system, contributions by firms to the unemployment insurance fund take the form of payroll taxes. A firm with a higher layoff rate does not pay higher contributions. Severance payments, as set in the law, are low. In other words, the contribution rate is (close to) zero. Second, the layoff process is subject to heavy administrative and judicial control. Firms have to prove either fault by the worker in the case of an individual layoff, or economic need in the case of collective layoffs. Judges can and often do disagree with the firms' decisions, leading to substantial time and financial penalties on firms.

This diagnosis naturally leads to our two main recommendations for reform:

- First, to increase the contribution rate of firms (that is introduce a layoff tax, and decrease the corresponding payroll tax) so firms internalize the cost of unemployment.
- Second, to limit the role of the judicial system. To the extent that firms are willing to incur the financial costs associated with laying off workers (and we are arguing that these costs should be higher at the margin than they are today), judges should not be allowed to second guess their decisions.

Our chapter is organized in nine sections.

Section 1 surveys of what is known – and not known – about the effects of existing systems of employment protection on the nature of labor markets – from the flows of separations, to the duration of unemployment, to the level of unemployment, to the response of the economy to shocks.

Sections 2 to 6 then focus on the optimal design of employment protection. Section 2 characterizes optimal employment protection in the benchmark economy. The next three sections explore three main deviations from the benchmark, and their implications for employment protection. Section 3 introduces limits on unemployment insurance, coming from

search or shirking incentive constraints. Section 4 explores the implications of liquidity constraints on firms. Section 5 explores the implications of alternative forms of wage setting. Building on these extensions, Section 6 discusses quits versus layoffs, and the role of the judicial process.

Section 7 attempts to put all these elements together, and draws conclusions about the contours of optimal employment protection. (Those readers who have no particular love for economic theory can go directly to Section 7 and get, we hope, the basic logic behind our conclusions.)

Section 8 then returns to the employment protection system in place in France today. Our purpose here is not to give an exhaustive description of the system, but rather to examine it in the light of our earlier analysis of how an optimal system might look, and to point out the major differences.

Having done so, we sketch in Section 9 the contours of what employment protection reform might look like in the case of France.

1 Employment protection and the labor market. Some empirical evidence

Most theories of the labor market suggest that employment protection – that is, either legal and administrative restrictions on layoffs, or tax payments to the state in case of layoffs, or severance payments to laid-off workers – is likely to have the following effects:

- It is likely to lead to lower layoff rates, and thus to smaller flows of workers through the labor market.
- It is likely to increase unemployment duration.[1]
- It is therefore likely to have a strong effect on the nature of the labor market, making it more stagnant, more "sclerotic" (smaller flows in and out, and higher duration of unemployment). But its effect on the unemployment rate itself, the product of duration and flows, is ambiguous.

1.1 Cross country evidence

The empirical cross-country evidence is indeed largely consistent with these implications:[2]

- Based on the indexes of employment protection constructed by the OECD and various other authors, there appears to be a strong negative correlation across countries between employment protection and flows of workers in and out of employment, or in and out of unemployment.[3]
- Using the same indexes, there appears to be a positive correlation between employment protection and unemployment duration. Countries with high employment protection tend to have higher individual average unemployment duration.

- The unemployment rate is the product of unemployment duration and the flows of workers in (or out of) unemployment.[4] Employment protection increases duration and decreases flows. The result of the two effects turns out to be a nearly zero cross-country correlation between protection and the unemployment rate.

The comparison between Portugal and the United States is revealing here (see Blanchard and Portugal (2001) for more details). Despite the fact that the degree of employment protection is much higher in Portugal than in the United States, the two countries have had roughly the same average unemployment rate over the past 30 years. These two unemployment rates hide however a very different reality. Unemployment duration has been three times higher on average in Portugal than in the United States; flows (relative to employment) three times lower.

These points are illustrated in the three panels of Figure 4.1, taken from Blanchard and Portugal (2001), that plots flows, duration, and the unemployment rate against measures of employment protection for 19 OECD countries.

Monthly flows into unemployment are constructed as the average number of workers unemployed for less than one month, for the period 1985–1994, divided by the average labor force during the same period, for each OECD country.

Unemployment duration is constructed as the ratio of the average unemployment rate for the period 1985–1994 to the flow into unemployment constructed above.

The employment protection index, "EPL", is the overall index constructed by the OECD for the late 1980s (OECD 1999b: Table 2.5); this index is a rank index for 19 countries, going from low to high protection. (The index is based solely on institutional aspects of employment protection, not on labor market outcomes.) The value of the index goes from 1 for the United States to 19 for Portugal (10 for France).

The top part of the figure shows a clear negative relation between the flow into unemployment (as a ratio to the labor force) and employment protection. The middle part shows a clear positive relation between unemployment duration and employment protection. The bottom part shows roughly no relation between the unemployment rate and employment protection. Regressions of the log flow, log duration, and the log unemployment rate on the employment protection index give:

log flow	=	0.50	-0.078 EPL (sd = 0.020)	$\overline{R}^2 = 0.46$
log duration	=	1.65	$+0.073$ EPL (sd = 0.033)	$\overline{R}^2 = 0.19$
log u rate	=	2.16	-0.005 EPL (sd = 0.019)	$\overline{R}^2 = -0.06$

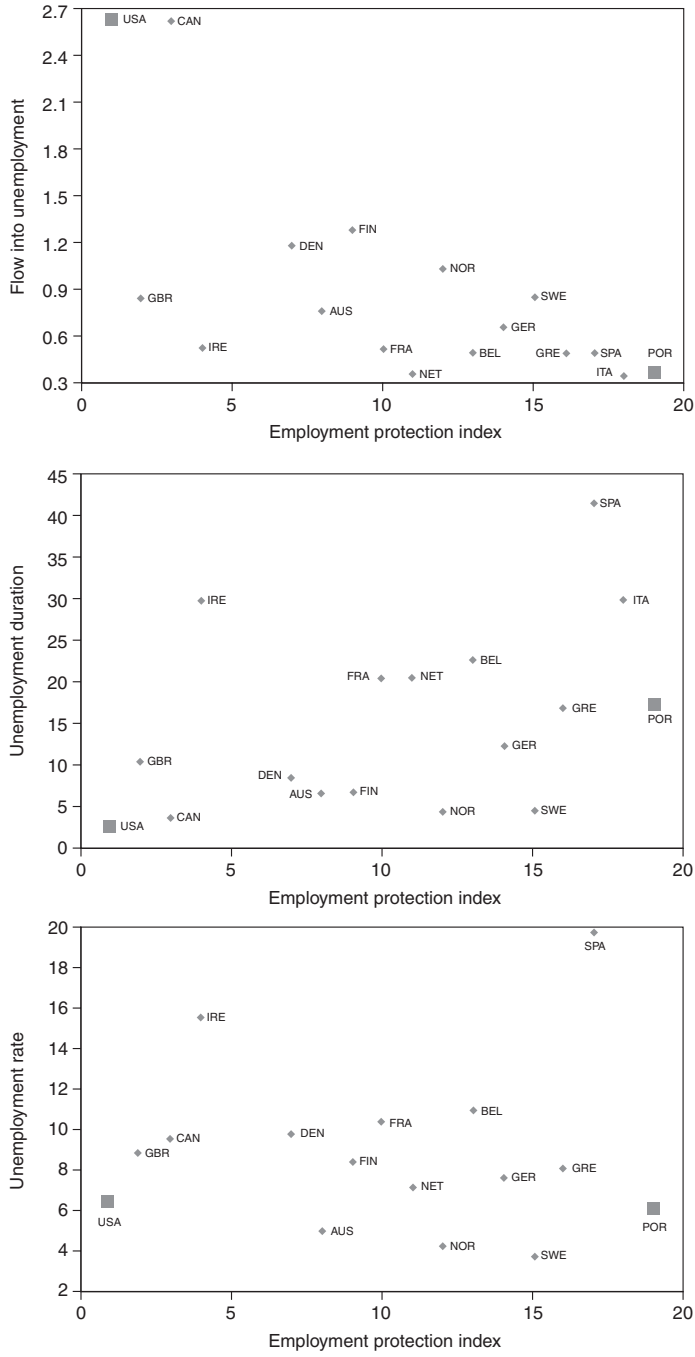

Figure 4.1 Flows, duration, unemployment and employment protection.

Thus, an increase in employment protection leads to a decrease in flows, and an increase in unemployment duration. But the two effects roughly cancel each other when looking at unemployment.

Based on these cross-country findings, the effects of employment protection appear quite bad. Employment protection decreases flows, and thus presumably decreases reallocation and efficiency. And, because it increases unemployment duration, not only does it not decrease unemployment, but it makes individual unemployment experiences more painful.

In addition, research on the evolution of unemployment over time and over countries, shows that countries with the more sclerotic labor markets (lower flows, higher duration) are also the countries which have suffered the largest and/or the more persistent increases in unemployment over the last 30 years (see for example Blanchard and Wolfers (2000)). This suggests that, to the extent that employment protection leads to more sclerotic labor markets, it also leads to larger and longer lasting effects of shocks on unemployment.

1.2 Two puzzles

The case is in fact not as tight as it looks. First, there are a few disturbing puzzles, facts that do not quite fit the general picture. Second – and to state the obvious – correlation does not imply causality. Let us take both points in turn.

In looking at reallocation in the labor market, economists have constructed two different sets of measures:

- The first, called "worker flows", are measures of the number of workers who change employment status over a given interval of time; for example, monthly flows from employment to unemployment, divided by total employment (as was used in Figure 4.1 above).
- The second, called "job flows", are measures of changes in employment levels of firms. Two standard measures here are "job creation", defined as the sum of changes in employment levels over a given interval of time, at all firms with a net increase in employment, divided by total employment; and "job destruction" defined as the sum of employment changes over a given interval at all firms with a net increase in employment.

The empirical puzzle is the following. As we saw above, measures of worker flows – for example flows out of employment – are lower in countries with higher employment protection. Measures of job flows – for example, measures of job destruction – appear however rather similar across countries. The puzzle is an important one to resolve: Worker flows suggest a strong adverse impact of employment protection on reallocation; job flows do not.

There are conceptually three reasons why the series on flows out of employment and the series of job creation may differ:

- Measurement errors. Worker and job flows are typically constructed from different sources; one source may be more reliable than the other. Also, because of differences in data construction, comparisons across countries may be misleading. While, indeed, comparisons across countries are often difficult, this line does not appear to be the key to solving the puzzle.
- Quits by workers. Such quits will show in worker flows, but if firms quickly replace the workers who have quit, those quits will not show up in changes in employment level of firms, and thus will not show up in job destruction. Based on a comparison of Portugal and the United States (a pair of countries chosen both for the difference in their labor markets, and the quality and comparability of their data), this appears to be relevant. One hypothesis is that to the extent that employment protection leads to long unemployment duration, it also makes employed workers reluctant to quit and look for another job, leading to lower quits.
- Differences in time intervals. Measures of job creation are typically constructed by looking at employment changes over a year. Measures of worker flows are often constructed at quarterly or monthly frequencies. Thus, transitory movements in firms' employment levels, movements reversed over the course of the year will show up in (say, monthly) worker flows, not in (say, annual) job flows. This indeed seems to be also part of the explanation. Firms in countries with high employment protection appear to smooth employment more, to reduce expected transitory movements in employment.

  If this is true, this has an important implication. It suggests that employment protection reduces transitory movements in employment. But it may not stand in the way of low-frequency reallocation, the kind of reallocation required by the process of structural change associated with growth.

These hypotheses are still tentative, and the subject of current research (see for example Bertola and Rogerson (1997) and Boeri (1999)). But they indicate that the link between employment protection and reallocation is more complex than it looks at first glance.

The other puzzle is a macroeconomic one. One would expect higher employment protection to lead to a slower and weaker response of aggregate employment to fluctuations in aggregate output. While this relation seems to hold roughly across countries, there is a number of exceptions. One such striking exception is Spain, one of the countries with the highest indexes of employment protection, where the response of aggregate employment to aggregate output is both strong and fast, stronger and

faster for example than in the United States. (This is true even for the time period when temporary contracts played a much smaller role in Spain than they do today (Bentolila and Blanchard 1990).) We know of no good explanation for this puzzle.

1.3 Correlations versus causality

The cross-country relation between employment protection and worker flows or unemployment duration, is suggestive of causality, but is hardly conclusive.

It is easy to think of other labor market institutions that may be correlated with employment protection and also affect flows and duration, yielding spurious correlations between protection and either flows or duration. In that respect, it is reassuring that, while we looked earlier at simple correlations and bivariate regressions, the empirical evidence suggests that the same results apply to partial correlations and multivariate regressions: Controlling for a number of other labor market institutions, such as the generosity of unemployment insurance systems or the nature of collective bargaining, higher employment protection still appears to affect flows negatively, and unemployment duration positively.[5] But even this evidence can easily be challenged. The other relevant institutions may be poorly measured, or simply not included in the regressions.

And labor market institutions, including employment protection, are not exogenous. It is also easy to think of factors which might lead to both higher employment protection and low worker flows, without implying a causal relation from protection to flows. (Think for example of a poorly developed mortgage market or high turnover taxes on housing, leading to low turnover in the housing market, low geographical mobility, and a political demand for employment protection.)

More conclusive evidence can only be obtained by observing the effects of changes in employment protection over time and space. Here, and somewhat ironically, most of the available evidence comes from the United States. Ironically, because the United States is often thought to be the country with no employment protection. But, while, indeed "employment at will" remains largely the rule and administrative restrictions on layoffs are minimal (for further discussion, see Autor *et al.* (2002)), the "experience rating" system implies that firms pay a large part of the cost of the unemployment benefits received by the workers they lay off.[6] And because the design of the system is left to each state, there are substantial variations both across states and across time.[7] These variations have been exploited by a number of researchers to obtain estimates of the effects of changes in the contribution rate on various dimensions of the labor market.[8]

One must be careful about the lessons one can draw from these empirical studies for the design of employment protection at the level of a

country. Increasing the contribution rate in one U.S. state but not in others, under conditions of high labor mobility across states, is likely to have very different implications for wages, flows, and unemployment than would increasing the contribution rate in all U.S. states at once (the type of change we want to think about when thinking about reform in a country). Nevertheless, some results come out relatively clearly, and are of direct relevance for us. In particular, a higher contribution rate clearly leads to a decrease in layoff rates. For example, estimates from Anderson and Meyer (1998), based on the 1984 change for the state of Washington, imply that an increase in the contribution rate from zero to one would have decreased layoffs by about 20 percent. Other relevant findings are that a higher contribution rate decreases seasonal fluctuations in employment (for a review, see Baicker *et al.* (1997)), and that a higher contribution rate increases the use of temporary help services by firms (Autor 2001).

1.4 The effect of recent reforms

Over the past 20 years, many European governments have attempted to reduce employment protection at the margin, by allowing firms, under specific conditions, to offer contracts with more limited employment protection. These contracts are typically of short maximum duration, with restrictions on renewals, on what type of worker or what type of job they can be used for. In France, these contracts are known as CDD ("contrats a durée determinée", fixed duration contracts), in contrast to the regular contracts, known as CDI ("contrats a durée indeterminée", contracts of indeterminate duration), and now account for 70 percent of new hires and 11 percent of employment (46 percent of employment for the 20 to 24 year olds).

The evidence is that the introduction of these temporary contracts considerably modifies the nature of the labor market.[9] It leads to higher turnover for those eligible for such contracts. The effects on both unemployment and estimates of welfare however are far from obvious. Much of the turnover appears to reflect a succession of low productivity, often dead end, jobs, and an increasingly dual labor market. Perhaps the main problem is that firms are very reluctant to keep workers at the end of their CDD, even if the worker proves to be a good match, it may be more attractive for firms to let him or her go and hire a new worker on a CDD, rather than keeping the existing worker under a CDI. Based on the evidence to date, the main effect of CDDs appears to be the emergence of an increasingly dual labor market.

2 Designing employment protection. A benchmark

In thinking about the issues, it is useful to start from a simple benchmark, which shows most clearly the relation between unemployment insurance and employment protection.

2.1 The benchmark economy

Think of the following economy:[10]

- Firms hire workers.
- After a worker has been hired by a firm, the firm learns about the productivity of the worker (that productivity may depend on the quality of the match between the worker and the firm, or on the demand for the firm's product, and so on).
- The firm may then want to keep the worker and produce, or lay the worker off. If the worker is laid off, he becomes unemployed.
- Absent any additional income, the utility of the worker when unemployed, is low. Put another way, and the terminology will be useful below, absent additional income, the "wage equivalent of being unemployed" is low.
- Workers are risk averse. Firms are risk neutral.
- There are no information problems, so everything is observable and contractible.

Under these conditions, firms will offer the following contract to workers:

- They will fully insure workers. They will do so by paying a constant wage to the workers they keep, and a severance payment to the workers they lay off.

  The severance payment will be such that the severance pay is equal to the wage, minus the wage equivalent of being unemployed. Workers will therefore have the same level of utility, whether or not they are employed or unemployed.
- They will lay workers off when productivity is lower than the wage equivalent of being unemployed.

  This is clearly the socially efficient rule for layoffs. From an efficiency point of view, workers should be kept on only if their productivity is higher than the wage equivalent of being unemployed.

  And firms do not need to commit to do so, because they fully internalize the cost of a layoff for workers. Given the wage and the severance pay, ex-post profit maximization leads them to lay a worker off only if productivity is less than the net labor cost, that is less than the wage minus severance pay, equivalently if productivity is less than the wage equivalent of being unemployed. But this is exactly the same as the efficiency condition above.

  In short, severance pay will be used to fully insure workers. And its presence will lead firms to take efficient separation decisions.
- Knowing that they will receive severance payments if they are laid off will lead workers to accept a lower wage in exchange. And because workers are risk averse, the provision of insurance by firms will

decrease their overall expected labor cost. Thus, firms will be eager to offer severance payments: this increases their expected profit.

In that economy, there will be substantial employment protection. It will take the form of severance payments by firms to laid off workers, sufficient to insure them against the loss of utility if unemployed. But, in that economy also, there will be no need for the government to mandate employment protection, firms will provide it willingly, and in the right amount.

2.2 Introducing an unemployment agency

To fully insure workers, firms must be able to assess the utility loss from unemployment. This is not easy for them to do:

- As of the time of the layoff, this loss is a random variable. The outcome of search is uncertain, and the worker does not know how long he is going to be unemployed. If the firm were to make a one-time severance payment to offset that loss, this one-time payment would do a poor job of insuring the laid-off worker.
- If the firm decides instead to pay the laid-off worker over time, contingent on his unemployment status, many other issues arise. The difficulty for the firm to actually track the worker, and determine whether he is still unemployed or has found another job; the difficulty in monitoring his search effort and making sure that he is indeed looking for another job.
- Rather obviously, individual firms cannot monitor laid-off workers well enough to provide them with adequate insurance. The role of monitoring unemployment status and search intensity must be therefore delegated to an agency, private or public.

  The state, given its existing administrative structure, is likely to be in the best position to do the monitoring, and to administer the payment of unemployment benefits, either alone or in conjunction with the private sector.[11]

So, go back to our benchmark, but now suppose that an agency is put in charge of monitoring and distributing unemployment benefits to the unemployed. Suppose further that the agency can perfectly monitor and thus provide unemployment insurance at no cost in terms of search intensity of the unemployed. How will this change the outcome relative to the benchmark?

The answer is: not much.

- Firms, when they lay a worker off, will make unemployment contributions to the unemployment agency – payments equal to the expected

value of the unemployment benefits to be paid to the laid-off worker, or payments over time corresponding to the unemployment benefits actually paid to the laid-off worker.

- The unemployment agency in turn will monitor and give unemployment benefits to the laid-off workers for as long as they are unemployed.
- There will be a sharper institutional distinction between unemployment contributions (paid by firms to the agency) and unemployment benefits (paid by the agency to workers). But, in this benchmark, the two will still be equal. The *contribution rate* – defined as the ratio of the contribution paid by the firm to the value of the unemployment benefits received by the laid-off worker – will be equal to one. (As we shall see, this will no longer be the case when we introduce other labor market distortions.)
- The allocation will be the same. Workers will be fully insured. Firms, because of the unemployment contributions they have to make to the agency in case of layoff, will fully internalize the social cost of unemployment and choose an efficient level of separations. There will be no trade-off between insurance and efficiency.

2.3 Unemployment contributions or severance payments?

We have so far interpreted unemployment contributions as a form of employment protection. A system in which all payments are made from firms to the state rather than to workers indeed provides employment protection to the worker. It makes it more expensive for firms to lay workers off, and thus reduces layoffs. But it may not look and feel like employment protection to the workers, who do not see the unemployment contributions paid by firms to the state, and do not receive payments directly from the firm.

It is therefore worth asking whether some payments could or should be made by firms to workers directly at the time of layoff. To do so, it is useful to distinguish between the costs of "becoming unemployed" and the costs of "being unemployed".

- The cost of becoming unemployed is the cost associated with losing a job, not with unemployment per se. It is a psychic cost, and while it is often ignored by economists, it plays an important role in public discussions of employment protection,[12] and its relevance has been well documented by social psychologists.[13] The loss of a long held job can and often does lead to a loss of a network of workplace friends, health deterioration, a loss of self esteem.
- The cost of being unemployed is the financial and psychic cost from remaining unemployed until one has found another job.

For our purposes, the main difference between these two costs is that the first is incurred at the time of the separation, and thus can be offset (in

terms of utility) by a one-time payment from the firm to the worker. The second, instead, is random at the time of the layoff. This suggests a natural division of tasks. Severance payments from firms to workers, at the time of the layoff, to compensate them for the cost of becoming unemployed. And unemployment benefits from the agency to workers, paid over time, and financed by payments from firms to the agency, to compensate workers for the cost of being unemployed.

In that light, how should the schedule of severance payments look like? The psychic loss appears to be primarily a function of time in the job of seniority. It is likely to be low for workers with low seniority, and to become high only with high seniority.[14] This suggests an increasing and convex schedule of severance payments as a function of seniority.[15]

Having established a framework, we consider in the next three sections three major deviations from the benchmark, and discuss, in each case, how they modify our conclusions.

3 Limited unemployment insurance

In our benchmark, the unemployment agency fully insured laid off workers. But, to the extent that the agency cannot fully monitor the search behavior of the unemployed, it can only offer limited insurance. Offering anything close to full insurance would lead the unemployed to stop searching and remain unemployed.

This issue has been studied at length in the theoretical and empirical literature on unemployment insurance.[16] And a number of recent reforms of the unemployment system in Europe, such as the PARE in France, have indeed had as their goal to combine more generous and longer lasting unemployment benefits with stronger incentives for the unemployed to accept jobs if offered by the unemployment agency. These reforms clearly go in the right direction. They potentially offer better tailored insurance. If truly no jobs are available, then the unemployed continue to receive unemployment benefits. And they remove, at least in principle, some of the problems associated with the open ended unemployment benefits of the past. But realistically, even the best designed systems cannot fully eliminate monitoring problems, and so, less than full insurance is optimal. There has to be some utility cost to unemployment to motivate search.

In such a context, the optimal employment protection/unemployment benefits system is more complex to characterize. The general architecture remains the same, but the details are different:

- The unemployment agency pays unemployment benefits to workers, providing as much partial insurance as is consistent with search incentives.
- The lower the feasible level of insurance, the higher the utility costs that layoffs impose on laid-off workers. So, to lead firms to take these

costs into account, unemployment contributions by firms to the agency must now exceed the unemployment benefits paid by the agency to workers. *The optional contribution rate is now greater than one.* And the layoff rate is smaller than in the benchmark.

- So, the more stringent the constraints on the amount of insurance the agency can provide, the higher the contribution rate relative to the benchmark, the lower the layoff rate, and, in this sense, the higher the optimal degree of employment protection.

Three remarks before moving on.

- Under this deviation from the benchmark, unemployment insurance and employment protection are substitutes, not complements. The poorer the insurance, the higher the optimal degree of employment protection. While the result is normative, this negative relation appears to be present in the data across Continental European countries.[17] The countries with the highest degree of employment protection (using the OECD index) are also the countries where unemployment insurance coverage has been relatively limited.

  Here, the political economy explanation may actually follow the same logic as our normative argument. To the extent that unemployment insurance was historically limited, employment protection probably served as a partial substitute.

  It is however a potentially poor substitute, leading to too few layoffs from the point of view of allocation. Thus, reforms of the unemployment system which introduce better monitoring and thus allow for better insurance have the added advantage of potentially allowing for a decrease in employment protection towards the benchmark, and thus a smaller cost in terms of reallocation.
- The results above bear a close relation to the results obtained in the "implicit contract literature" of the 1970s and early 1980s (in particular Baily (1974), Azariadis (1975), and Akerlof and Miyazak (1980)). That literature looked at the optimal contract between risk neutral and risk averse workers. Under the assumption that there were neither severance payments nor unemployment benefits, one of the conclusions was that there would be overemployment, that firms would layoff too little relative to the efficient outcome. One of the criticisms addressed to those papers was the question of why firms did not offer unemployment insurance or severance payments. In the discussion here, the limits come from monitoring problems, and the solution takes the form of a layoff tax rate imposed by the state. But the logic is very much the same.
- Returning to the discussion of unemployment benefits versus severance payments discussed in the previous section, it has sometimes been argued that severance payments are preferable because they do

not lead to the search incentive problems discussed here. This is correct, but fixed payments in the face of random unemployment duration deliver very poor unemployment insurance. Even if search considerations imply declining unemployment benefits over time, it is unlikely that the optimal schedule consists of a lump sum payment at the start, and another thereafter.

Only if the administrative costs of setting up an unemployment agency appear prohibitive, does a system based on severance payments make sense. This may be the case for some low and middle income countries; it is surely not the case for France.

4 Risk aversion and shallow pockets

A second assumption of our benchmark was that firms were risk neutral, had deep pockets, and could therefore fully insure workers (with the help of an agency to run the unemployment insurance system, and subject to the discussion we just had about incentives to search when unemployed).

This assumption is also clearly too strong. It is based on the idea that, if firms are widely held, most of the risk faced by a firm is diversifiable. But while most of the variations faced by firms are idiosyncratic, some are not. And most small firms are not widely held. Many are privately held, and their owners' wealth is not much diversified. So, the assumption of risk neutrality, especially for small firms, is too strong.

And, even if we were to assume that firms are risk neutral, the assumption that they have deep pockets, and thus can pay workers in bad states, is also too strong. Clearly a firm that has gone bankrupt may not be able to pay its unemployment contributions or make severance payments. But short of this extreme case, corporate finance suggests that the shadow price of internal funds to firms is likely to be a decreasing function of the state. The shadow price of severance payments to workers or payments to the state in bad states, even if feasible, may be high; the funds could be better used for other purposes.

Now, to state the obvious, layoffs are more likely to take place in bad states, when the shadow price of internal funds is high, than in good states. And so, a higher layoff tax may potentially make things worse for firms, imposing a high utility cost on the small entrepreneur, or preventing the larger firm from taking other measures required to improve its situation. One may hope that, in response to an increase in layoff taxes, financial markets will partly adjust to alleviate the problem, providing more funds to the firms in bad times to allow them to pay the now higher layoff taxes. But, more likely than not, the adjustment is likely to be only partial.

What should the state then do?

- Separate the timing of layoffs from the timing of unemployment contributions. Ideally, the state should collect layoff taxes from firms in

good states rather than in bad states. And tax payments to the agency should depend on the probability of layoffs. Firms which, for some reason (a different distribution of productivity or demand shocks for example), have a higher probability of laying workers off should make higher contributions.

The problem is how to design such a system, or an approximation to such a system, in practice. One possibility, and that adopted for example in the United States, is to introduce a bonus-malus, or an experience rated system. We return to it below.

- Even the best designed experience-rated systems are unlikely to fully eliminate the additional liquidity problems created by layoff taxes. If so, it may then be optimal for the government to choose a lower layoff tax rate, and thus a *contribution rate lower than one*. This decreases the tax burden on firms in bad times. Obviously, it does so at the cost of raising another distortion, a contribution rate below one leads firms not to fully internalize the costs of layoffs, and thus leads to too high a layoff rate.

In the rest of the section, we take up two related issues. First, the design of experience-rated systems; second, the issues raised by limited liability and the possibility of bankruptcy, issues we have left aside up to now.

4.1 Bonus malus, and experience-rated systems

As we have just seen, an ideal collection system for layoff taxes is one in which the state (to the extent that it has deeper pockets than the firms) collects layoff taxes in good states rather than in bad states, and where the tax rate to be based on the firm specific probability of layoffs.

Two obvious problems in practice are that both the state faced by the firm and the probability of layoff by the firm are likely to be unobservable by the agency. A natural solution is then to base the payments of firms on their past behavior (as in bonus-malus system), and to allow them to pay the taxes over time. This is the rationale for so called "experience-rated" systems of unemployment contributions used in particular in the United States (a useful description of the U.S. experience is given in Fougère and Margolis (2000)).

The systems vary across U.S. states. It is useful to describe the most commonly used system, called the "reserve ratio" system of unemployment contributions. Leaving aside the many complicated details, its principle is simple. Each firm has a running balance with the state unemployment agency, with contributions by the firm to the fund on one side, and benefits paid by the agency to the workers laid off by the firm on the other. Once a year, the state computes the net outstanding balance, and requires the firm to pay some proportion of this outstanding balance over the following year. The factor of proportionality depends both on the net

balance of the firm, and the net balance of the state fund as a whole. This system has two implications:

- Ignoring discounting, and assuming that firms do not go bankrupt and do not hit the various ceilings that limit contributions (all considerations being relevant in practice), firms eventually pay the full cost of unemployment benefits for the workers they lay off – the contribution rate is equal to one.
- The factor of proportionality determines how the timing of payments depends on current and past layoffs. If the factor of proportionality is equal to one, so firms are asked to return to zero balance each year, then payments are closely related to current (or more precisely last year's) layoffs. The lower the factor of proportionality, the more contributions depend on past layoffs.

How should one then think about the choice of the factor of proportionality? If firms are operating in a stable, ergodic, environment, going sometimes through good times, sometimes through bad times, then letting the factor of proportionality be small will make the firm's contributions depend on its mean observed layoff rate in the past, which is also equal to the probability of a layoff in the future. If, however, as is more likely, the underlying probability changes over time, then a higher factor of proportionality, giving more weight to recent layoffs, will be closer to the underlying current probability. But it will impose higher liquidity costs on firms.

4.2 Bankruptcy, and the role of firm guarantees

The possibility for firms to pay layoff taxes over time rather than at the time the layoffs take place raises an issue that we have avoided so far. This is the possibility for firms to evade taxes by going bankrupt. Absent firm guarantees, or other commitments, a firm that lays off its workers at the same time it declares bankruptcy may be able to avoid paying most if not all layoff taxes. And the problem is likely to be worse under an experience-rated system. The longer the lag between layoffs and tax payments, the larger the proportion of layoff taxes a firm will be able to avoid through bankruptcy.[18]

This is likely to lead some firms to reorganize so as to make it easier to avoid paying taxes. Ways of doing so include isolating high risk divisions and transforming them into separate legal units with little collateral, so that, in case of bankruptcy, there are few or no assets left to the agency or any other creditor to recover (leaving the unit "judgment proof"). Such a behavior has been well documented in the case of environmental protection (see for example Ringleb and Wiggins (1990)). And, as recent examples such as the bankruptcy of MetalEurop show, some European firms are already moving systematically in that direction.[19]

What should the state then do?

- It should, as is already the case for other legal obligations, have senior creditor status. But there may not be enough assets left even for the senior creditors to collect.
- It can extend responsibility for payments of these taxes to third parties.

  This is for example the approach taken by the law on environmental liability passed in the United States in 1980 (a law called the Comprehensive Environmental Response, Compensation, and Liability Act, or CERCLA). Under that law, if a bankrupt firm cannot pay for decontamination of a site, the state can go after any "potentially responsible party". What this means is unclear and evolving; it may include suppliers, investors, or new site owners.
- It can ask for more guarantees, in the form of collateral, physical or financial, or of bank guarantees to cover potential layoff taxes.

Guarantees, collateral, and extension of liability to third parties all have costs. Collateral may be better used for other activities. Third parties may prefer not to deal with a firm if this exposes them to potential tax liabilities. We have no set view, and the evidence from environmental protection is still unclear. Nevertheless, any proposal to increase the contribution rate of firms must confront the issue.

5 Insurance, employment protection, and wage determination

A third assumption of our benchmark was that wages were set at the time of hiring. So, to the extent that firms (or the overall system of unemployment insurance cum employment protection) offered insurance in case of layoff, risk averse workers were willing to accept a lower wage on the job, and willing to accept lower expected income overall.

This may not be the right view of wage setting. True, initial wages are set at the time of hire. But these are only set for a short period of time, at which point they may be renegotiated. At that time, wages will reflect the bargaining position of each side. This has important implications.

5.1 Ex-post wage setting, and bonding

Consider two firms:

- One offers severance payments to its workers, makes unemployment contributions to the state, and the workers it lays off receive unemployment benefits. The other does not offer severance, does not make unemployment contributions, and the workers it lays off do not receive unemployment benefits.

If wages are set at the time of hiring, the first firm will be able to offer lower wages, and indeed, because workers are risk averse and value the insurance the firm provides, it will have lower expected labor costs than the second.

- Now suppose that workers can renegotiate wages after hiring. Then, workers in the first firm will be in a much stronger bargaining position than in the second. If they find themselves unemployed, they will receive unemployment benefits. And, if the firm wants to lay them off, the firm will have to pay severance and contributions to the state. Thus, the firm that provides insurance will now have higher wages and by implication expected labor costs than the second.
- Given the choice, firms will therefore not be eager to offer insurance. And, if the state puts in place a schedule of severance payments, of unemployment contributions, and of unemployment benefits, along the lines we described in the benchmark, all three components will lead to higher wages, and thus to higher expected costs for firms.

This view of wage setting may itself be too extreme. The central issue here is known in labor economics as "bonding".

Suppose firms could extract "bonds" from workers – that is payments from workers at the time of hiring to compensate for the increase in wages they know will take place after hiring. Firms could then eliminate the effects of ex-post bargaining on cost.[20] If bonding was prevalent, we would be much closer to the benchmark model, or the extensions we saw earlier. Firms would be willing to pay severance to the workers or make payments to the state as under our benchmark. Whatever increase in cost this implied, they could recoup through the receipt of a sufficiently large bond at the time of hiring.

The obvious remark at this point is that we just do not observe "naked" bonding. Workers do not pay firms at the time of hiring. Bonding however exists in more disguised forms. Some workers accept to be paid a low initial wage, in effect paying a bond early in their job tenure, to partly compensate the firm for the higher wages later in their job tenure. Yet, in practice, the room for bonding is limited, and so the conclusion must be that insurance cum employment protection is more likely to increase than to decrease labor costs.[21]

What should the state then do? It clearly faces a trade-off:

- Choosing a contribution rate equal to unity, i.e. a system that makes firms pay for the full cost of an additional layoff, will lead firms to take the right decision at the destruction margin. Layoffs will take place only when the productivity of a job is less than the wage equivalent of being unemployed.

  But this high contribution rate will also increase the bargaining power of workers, and thus increase the wage. This will increase the

overall cost of labor, both directly and indirectly, and will adversely affect job creation. How much will depend on the amount of effective bonding.

- Choosing a contribution rate less than one, will lead firms, in contrast, to destroy too many jobs, and lead to too many layoffs. It will however lead to a smaller increase in the overall cost of labor, both directly and indirectly (through the effect on wages), and thus have a smaller adverse effect on employment creation.[22]

Parallel arguments apply to the direct severance payments part, but for one difference, in the case of the contribution rate, we were looking at the effects of varying firm contributions, keeping unemployment benefits the same. Here, by the very nature of direct severance payments, we are changing both the firm's contributions, and the benefits received by workers. The closer these payments are to fully compensating for the psychic costs to workers of being laid off, the less distorted the destruction decision. But the higher is then the cost of labor, both directly and indirectly, and so the more distorted is the creation margin.

In short, the more firms are made to pay for the expected cost of unemployment benefits, the smaller the distortion will be at the destruction margin, but the larger the distortion will be at the creation margin.[23] Because of these distortions, there is now a trade off between insurance and efficiency. Even if it were feasible (if there was no problem in monitoring the search behavior of the unemployed), it will no longer be optimal to provide full insurance to laid off workers. And the *optimal contribution rate will be less than one*. It will be closer to one,

- the higher the scope for bonding, and so, the smaller the adverse effect on layoff taxes on creation;
- the lower the bargaining power of workers, or the higher the commitment ability of firms;
- the more elastic job destruction to the layoff tax;
- the more inelastic job creation.[24]

5.2 Heterogeneity of firms and workers

Not all firms nor all workers are alike. Some firms operate in more volatile goods markets, and so are more likely to have a high layoff rate. Some workers, because of their characteristics, are more uncertain and more likely to be laid off. If laid off, some workers are likely to find a job more easily than others. What does a positive contribution rate imply for their respective fortunes?

To think about this question, first go back to our benchmark case, in which wages are set at the time of hiring.

- In that benchmark, firms must offer the same level of utility to a given worker, otherwise the worker will not accept the job offer. Thus, firms that face more volatile demand, and thus higher layoff rates will have to make higher overall unemployment contributions and will not be able to pass those costs on to workers through lower wages. They will therefore have higher costs. This is indeed as should be, given that they impose larger social costs.
- In that benchmark, if workers are substitutes in production, a worker with higher expected unemployment duration if laid off will be hired by a firm only if the total cost he imposes on the firm is the same as that for other workers. Thus, workers with worse labor market prospects, will have to accept lower wages. At those lower wages, firms will be willing to employ them.

Now turn to the case where, instead of being set ex-ante, wages are set ex-post through bargaining, and the contribution rate is positive. Then:

- As wages are now likely to increase rather than decrease in response to a positive contribution rate, all firms will face higher costs. But, to a first approximation, the increase in the wage will be the same across firms, so the increase in costs (relative to the benchmark) is the same at all firms. There is therefore no obvious reason why the contribution rate should thus be modulated across firms, for example, why it should be smaller for firms with high turnover rate.
- An issue arises however with respect to firms operating in isolated labor markets. Take for example the case of a firm operating in a depressed region. If the firm is the only one around, and closes its plant, it may be very difficult for workers to find other jobs. The layoffs will have high social costs. This suggests imposing larger contributions on the firm that is laying off. But, with such large contributions, which firm will ever want to open another plant in that labor market? To the extent that the state wants to maintain employment in the region, the solution is not to modulate the contribution rate, but rather to use job creation subsidies.
- The situation is now different for workers. Workers who are perceived by firms to be more risky, in the sense of either a higher probability that the worker will have to be laid off, or a higher expected unemployment duration if he becomes unemployed, will cost more to firms. Once they have been hired, they will be able to renegotiate the wage, and thus increase the firm's costs. The increase in costs will be larger, the higher the probability that the worker may be laid off, or the longer his or her expected unemployment duration. Knowing this however, firms will not want to hire these workers in the first place.

  Thus, a positive contribution rate (in general any employment protection) will lead to increased discrimination by firms in the labor

market. Workers with a short labor market history, workers with poor skills, older workers may have a hard time finding jobs.

What should the state then do?

To reduce the problem of different ex-ante probabilities of layoffs for different workers, a natural, if partial, solution is to give time to both parties before the usual rules of employment protection and unemployment insurance apply. This may take two, not mutually exclusive, forms:

- A trial period, during which any of the two parties can separate at no cost. This period must be long enough to allow the firm to learn about the worker, but short enough to make it unattractive for firms to fill jobs through rotations of trial period workers.
- A transition period during which, in case of separation, both the payments by the firm to the agency, and the unemployment benefits received by the laid off worker are less than under standard rules, and increase with seniority.

To the extent that this may be not enough, one may think of other types of solutions, for example:

- Targeted hiring subsidies, but this is likely to add yet another layer of complexity and arbitrariness in the employment protection system. The French experience is not particularly encouraging in this respect.
- Or/and a contribution rate by firms which depends on the *number of layoffs*, rather than on the expected or actual total unemployment benefits paid by the agency to the workers laid off by the firm. This second solution does not eliminate the problem raised by different ex-ante probabilities of layoffs for different workers, but it eliminates the problem raised by different ex-ante unemployment durations for different workers.

6 Quits and layoffs

We have focused so far on adverse shocks to productivity, shocks which lead the firm to lay a worker off even if the worker does not have another job opportunity. The reason was that these are obviously the shocks where unemployment insurance and employment protection may have a role to play. But these are not the only shocks triggering separations. Workers leave for other reasons, often because they have a more attractive job opportunity elsewhere. In France today, leaving aside the separations that take place at the end of fixed duration contracts (CDD), layoffs account only for about one third of separations, quits for the remaining two thirds (Goux and Maurin 2000).

The presence of both layoffs and quits introduces a number of issues in

the design of employment protection, and these are the issues we discuss in this section.

6.1 Introducing quits and layoffs

Go back to our benchmark model. Assume now that there are two shocks that take place after a worker has been hired. First, as before, productivity is realized. Second (and simultaneously), with positive probability, the worker receives an outside job offer. Suppose, for simplicity, that if the outside job offer comes, it dominates any offer the firm can make to the worker.

There are now two reasons why there may be a separation. Productivity may be low, and the worker becomes unemployed. Let's call this a layoff; it is initiated by the firm. Or, the worker may receive an outside job offer, in which case he will leave. Let's call this a quit.

If the level of productivity, and the existence of a job offer, are both observable, and if required, verifiable in court, the conclusions we reached earlier extend straightforwardly to this case. If a separation comes from low productivity, and is therefore a layoff, firms make contributions to the state, and pay severance payments to the worker. If a separation comes from an outside job offer and is therefore a quit, it triggers neither severance payments, nor unemployment contributions by the firm, nor unemployment benefits to the worker.

The problems arise when the reason behind the separation is unobservable, or not verifiable, or worse, manipulable by firms or by workers. This generates two types of potential games, first between firms/workers and the state, second between firms and workers themselves.

6.2 Games between the firms/workers and the state

Focus first on the payments from the firm to the state to finance unemployment benefits. And assume, for reasons we discussed in the previous section, that firms support only a proportion of these costs: the contribution rate is less than one. This opens the possibility that, for the firm and the worker taken together, each layoff may be associated with a net subsidy from the state (the firm pays less to the state than the payment of the state to the laid off worker). Thus, to the extent that firms and workers collude, they may have an incentive to call every separation a layoff.

Is this likely to be a serious issue in practice? Probably not. The parties may have neither the ability nor the incentive to collude:

- It may not be easy for the firm and the worker to collude. Collusion implies a payment from the worker to the firm, so as to offset the payment from the firm to the state. To the extent that the payment comes from future unemployment benefits or future wages received

by the worker, the ability of the firm to make sure that such payments actually take place may be limited.

- The exact nature of contributions by firms to the agency matters here. If contributions by firms depend on actual unemployment benefits paid to the workers who were laid off, then indeed firms and workers together benefit from calling a quit a layoff. Suppose instead that contributions depend (for the reasons discussed at the end of the previous section) only on the number of layoffs, or, equivalently, on the number of layoffs times the average duration of unemployment benefits. In this case, it is much less obvious that the firm and the worker together will benefit from declaring a quit to be a layoff. Workers who quit are likely to have a shorter unemployment duration than average, and thus receive smaller unemployment benefits, than average. In particular, many of the quits are directly to another job, in which case the worker receives no unemployment benefits at all. In that case, there is no gain to the worker, and a loss to the firm in declaring it a layoff. There is no incentive for them to call the quit a layoff.

6.3 Games between workers and the firm

Assume now that the contribution rate is one, so that we can ignore the previous game between firms/workers and the state. There is another game we have to consider. Other things equal, firms would rather have a separation be called a quit and save on severance payments and unemployment contributions. Symmetrically, workers would rather have a separation be called a layoff, and receive both severance payments and unemployment benefits.

If the worker could not affect the productivity of the match, and the firm could not affect the relative attractiveness of the outside option of the worker, then there would still be no problem.[25] Firms with a low productivity shock could not force the worker to quit. Workers with an outside job offer could not force the firm to lay them off. But, in fact, workers can affect the productivity of a match, and firms can affect the relative attractiveness of the outside option of the worker.

A worker who wants to quit but also wants to receive severance payments and unemployment benefits, can shirk and decrease the productivity of the match, leaving no choice to the firm than to lay him off. A firm that wants to lay a worker off but would rather have him quit so as to save on severance payments and unemployment contributions, can harass the worker into quitting. The stronger the stakes, that is the higher the contribution rate and the higher the unemployment benefits or the severance payments, the higher the incentives to harass or to shirk.[26]

As in the game between firms/workers and the state we discussed earlier, there is a relevant difference between severance payments and unemployment benefits. If a worker has an outside job offer, it makes

sense to shirk, so as to be laid off, and receive severance pay. But, if the worker intends to take the other job right away, shirking so as to be laid off and receiving unemployment benefits is of no value to that worker: he will not be unemployed.

This has two implications. Unemployment benefits are, in that respect, less likely to lead to gaming, than severance payments. Shirking by workers may be less of an issue than harassment by firms.

Until now, in our argument, there was no reason to have courts involved in the process of separation (except for the usual reasons, making sure that existing rules – payment of severance, advance notice, no discrimination on the basis of sex, age, physical appearance, no layoff because of union activity, and so on – that are in place are not violated). But the issues we just discussed now create such a role. Let's turn to this.

6.4 The role of courts

Under the logic of our arguments, what courts have to do is conceptually clear (if not necessarily easy to do in practice):

- If a separation has been reported as a layoff, look, if requested by the firm, at evidence of shirking by the worker. (This can take different forms, with different ways of allocating the burden of proof. A firm that does not want to pay severance payments to a worker and the payment to the state, may state that the separation is the result of misbehavior by the worker, and, if challenged by the worker, has to prove it in court.)
- If a separation has been reported as a quit, look, if requested by workers, at evidence of harassment by the firm.

An important remark at this point, to which we shall return after having described the role of courts in the current French employment protection system: the role of courts described above is very different from their role in France today. In particular, in our framework, if a firm is willing to call a separation a layoff and make the associated payments to the state and to the worker, there is no justification for the court to second guess the decision of the firm, no justification for the court to intervene at all. This is not the case today.

7 The contours of optimal employment protection

The purpose of this section is simply to summarize the main conclusions reached in the previous five sections.

- Employment protection is a natural counterpart to unemployment insurance.

A full discussion of unemployment insurance falls outside the scope of this chapter. Nevertheless, let us make a few remarks here.

Individual self insurance is not sufficient to insure workers against the risk of job loss and unemployment. Perhaps more could and should be done here (for example along the lines of the unemployment accounts proposal presented in Feldstein and Altman (1998); see also Kugler (2002), for an analysis of severance payment savings accounts in Colombia). In any case, we take as given in this chapter that such private accounts cannot simply replace traditional unemployment systems.

This implies the need for an agency to administer unemployment insurance. This agency can be either a public agency, or a public–private partnership. Only the state has the required administrative infrastructure, to follow the unemployed, to distribute benefits, and to collect contributions for firms.

A public agency may however not have all the right incentives. We see this in some of the problems emerging in the implementation of the PARE in France. The PARE represents an attempt to provide more generous (in time) unemployment benefits, in exchange for stronger incentives for the unemployed to accept jobs if such jobs are available. Agency employees do not however have strong incentives to force the unemployed to take jobs, and the preliminary evidence suggests an increase in benefits has not come with much stronger inducements for the unemployed to take jobs. A public–private agency would have stronger incentives to place the unemployed into jobs.[27]

- The general principle should be that firms make payments to the unemployment agency equal to the expected or actual unemployment benefits paid to the laid off workers. In other words, the *contribution rate* of firms, defined as the ratio of contributions paid by the firm to the (expected or actual) unemployment benefits paid to the worker, should be equal to one. Such a rate leads firms to fully internalize the social cost of layoffs and take an appropriate layoff decision.
- The principle is important. But a number of other imperfections in the labor and other markets require however a number of qualifications:
    - To the extent that unemployment insurance is necessarily incomplete (for example to maintain incentives to search), it is then optimal to choose a contribution rate larger than one, therefore decreasing layoffs below the efficient level, but in doing so, providing more insurance to workers.
    - To the extent that firms have liquidity problems, a high contribution rate and the payment of unemployment contributions may create serious problems for firms already in financial trouble. In this case, it is better to separate the timing of unemployment

contributions by firms and the payment of unemployment benefits, according for example to a bonus-malus or an experience-rated system. It may also be optimal to choose a contribution rate less than one, so as to decrease the burden on firms in financial trouble, again at some cost in efficiency.

- To the extent that wages do not fully reflect the provision of insurance, a contribution rate equal to one will avoid distortions at the destruction margin, but it will also increase labor costs, decrease profits, and thus create distortions at the creation margin. In this case, it is again optimal to balance the two distortions by choosing a contribution rate less than one.

On net, given our state of knowledge – theoretical and empirical – no one can state with much confidence what the optimal contribution rate should be. Our guess, but it is hardly more than a guess, is that the last two factors dominate the first, and the contribution rate should be positive, but somewhat below one.

- A related question is whether the contribution rate should be modulated across firms or across sectors.

Some sectors and some firms have a much higher turnover than others. This turnover will decrease as the contribution rate is increased. But it is likely that some sectors will continue to have higher turnover and thus higher layoff tax costs. This is however as it should be. These sectors impose higher costs on society, and this should be reflected in higher costs for firms in those sectors.

The contribution rate may however have to be modulated across workers. Some workers are more uncertain and thus more likely to be laid off than others; some workers have higher expected unemployment duration than others, this may be because of age, of skill, or other characteristics. If these workers accepted sufficiently lower wages, firms would be willing to hire them. But, in the presence of wage floors, or ex-post wage setting, wages are unlikely to adjust enough and these workers are likely to cost more to employ. This in turn will lead firms to discriminate against workers who are, or are perceived as, more likely to be laid off, or more likely to remain unemployed for a long time.

A partial solution is to have a trial period during which separation can happen at no cost to either party, or/and a transition period during which unemployment benefits and employment contributions are lower than under standard rules, and increase with seniority.

Other partial solutions include targeted hiring subsidies (but experience suggests that there are many pitfalls with such targeted subsidies), and unemployment contributions by firms that depend not on actual or expected unemployment benefits paid to laid off workers, but on the number of layoffs. In this case, firms have no incentives to

discriminate against workers with longer expected duration of unemployment.

- In case of bankruptcy, unpaid outstanding layoff tax balances should be counted as a liability of the firm, and the state should be a senior creditor.

  As the experience with environmental liabilities has shown, this may not be enough. Firms may systematically reorganize and spin off risky units so as to leave empty shells in the event of bankruptcy. In this case, it may be desirable to have the outstanding liabilities to the unemployment agency be backed by collateral or by bank guarantees. This will transfer the monitoring of the balance sheet of firms to banks or other creditors.
- The previous points have concentrated on contributions by firms to the state. But there is also a potential role for severance payments, payments made directly to workers.

  Their role should not be to help workers finance unemployment. This is better done through unemployment benefits. Their role should be to compensate, at least in part, for the costs of becoming (as opposed to being) unemployed. These payments should be a (non linear) function of seniority, with low payments until high seniority has been achieved.

  Thus, on the financial side, employment protection could take two forms. Unemployment contributions to the state; while these are not directly visible to workers, they protect unemployment in the sense of making layoffs more expensive for firms. And severance payments directly to workers.
- The role of the judicial system should then be, in addition to making sure that administrative steps are followed, to assess whether declared layoffs are indeed layoffs, and declared quits are quits.

  To avoid having to pay unemployment contributions and severance payments, firms may harass workers into quitting. In order to qualify for unemployment benefits and receive severance payments, workers may shirk so as to be laid off.

  The role of the judicial system should then be twofold. If asked by workers, to look for evidence of harassment of workers if a separation has been called a quit. If asked by firms, to look for evidence of shirking by workers if the separation has been called a layoff.

  The role of the judicial system should not however be to second guess the layoff decisions of firms. If a firm is willing to call a separation a layoff, follow the relevant administrative steps, and pay the associated financial costs, this decision should not be subject to judicial challenge (except on usual grounds such as discrimination based on race or sex).

8 The French employment protection system

Our purpose here is not to give an exhaustive presentation of the French employment protection system but rather to present it in such a way as to facilitate the comparison with the conclusions of the previous section.[28]

Much of the evolution of employment protection has been organic, the result of jurisprudential decisions, codified once in a while by new laws. While the Napoleonic code was based on the notion of "employment at will", the law has evolved towards the notion of "social responsibility" of firms (what this means, we shall argue below, is far from clear).

This evolution has been a slow, steady, one. It accelerated, in the direction of stronger employment protection, in the 1970s and early 1980s, probably in response to the steady increase in unemployment during the period. In 1973 for example, the burden of the proof that a layoff is justified was shifted to the firm. In 1975, the state introduced the requirement of prior administrative approval for layoffs; this requirement was eliminated in 1986. Except for the extension of the scope for fixed duration contracts, not much has happened since to employment protection for regular contracts. Indeed, and here France is an outlier in Europe, the most recent law, the "Loi de Modernisation Sociale" passed in 2002, has reinforced employment protection for regular contracts.

8.1 The need for motive

The general principle today is the need for motive. The firm must have and show "real and serious cause". Only if such a cause exists can the firm layoff a worker.

The law distinguishes between two types of layoffs:

- "Personal" (that is related to the behavior of the employee). The firm must show that the layoff is the result of a "serious misdemeanor" (faute serieuse).

  What "serious" means is not clearly defined (one definition, found in the reference labor law text (Pélissier *et al.* 2002) is: "serious: sufficient to justify the layoff"). It does not require malicious intent, but it must be more than a "light misdemeanor" (faute légère), which does not justify a layoff.
- "Economic" (that is related to the situation of the firm). The firm must show that the layoff (or layoffs) are the result of "real transformation or elimination of job(s)".

  What this exactly means is even more unclear. The ambiguity, and why this is an issue, is best shown in the recent case of layoffs at Michelin–Wolber. In July 1999, Michelin decided to layoff 451 workers at its Wolber plant, at the same as it was announcing large benefits for the group as a whole. In February 2002, the labor tribunal

concluded that the layoffs were not justified, and asked Michelin to pay a total of 10 million euros to the 162 laid off employees who had contested the decision, or about 60,000 euros per employee. The tribunal argued that "layoffs for economic reasons cannot be justified on the basis of improving the competitiveness or the profits of the firm, but only on the basis of maintaining its competitiveness. In the case of Michelin, the purpose was to improve competitiveness, and thus the layoffs were not justified." (The decision is being appealed.)

A charitable interpretation of the court's opinion is that the firm should exercise more restraint with regards to layoffs when it is not liquidity constrained (such a conditioning would make economic sense). We doubt, though, that the courts have the ability and the information to make such business judgments.

Lest one think that this is an isolated case, very much the same thinking was embodied in the 2002 law, which stated that only when other avenues had been exhausted, were layoffs justified. Two of the provisions of the law were subsequently thrown out by the French Supreme Court (the Conseil Constitutionel) on the grounds that the law had moved from the principle that layoffs were justified if they were required to maintain competitiveness to the principle that layoffs were justified if they were required to ensure the survival of the firm – a much more stringent criterion.

In short, the principles that the courts must use in assessing whether layoffs are justified are extremely unclear. The fact that the firm decided that such layoffs were necessary is clearly not by itself sufficient proof for the courts.

8.2 Limited financial costs

If layoffs are not contested, or are found by the court to be justified, the direct financial costs to firms are relatively limited:

- Contributions by firms to unemployment benefits are collected through payroll taxes. The rate is independent of the history of layoffs by the firm – in other words, the contribution rate is zero.

  One exception is the "contribution Delalande", introduced in 1987, and mandating additional payments to the unemployment agency in case of layoffs of older workers. For large firms (50 employees or more), the contribution is equal to two months for a 50-year-old, increasing to 12 months for a 56-year-old, and decreasing back to 6 months for a 59-year-old or older (the number of months is halved for firms with less than 50 employees).
- The severance payments mandated by the law are relatively low, and non linear in seniority: 2/10 months per year of seniority, plus, for workers with more than 10 years, 2/15 months per year above 10

years. This gives 2 months for a worker with 10 years' seniority, 8.3 months for a worker with 30 years' seniority.

Some other obligations of firms, such as the obligation for large firms to pay a "congé reclassement", described below, are however equivalent to severance payments (plus a training component). And severance payments set by sectoral agreements ("conventions collectives") are often higher than those set in the law. Estimates by Abowd and Kramarz (1997) for 1992 give a marginal cost of a layoff to a firm of 5 to 7 months of average labor costs per worker.

8.3 A long procedure

Firms that decide to lay workers off for personal or economic reasons must follow an often long series of administrative steps. These steps have two separate purposes.

- The first is to give time to the workers to prepare themselves for the layoff and to facilitate their reemployment. Depending on seniority, workers get an advance notice of up to three months. Workers in large firms (1,000 employees or more) are entitled to a retraining period ("congé reclassement") of 4 to 9 months. For the part of the period that coincides with the advance notice period, workers get 100 percent of their salary; for the rest of the period, they get 65 percent of their salary, paid by the firm. Under the new unemployment insurance system, workers in smaller firms are eligible for training and help in finding jobs from the start of their advance notice (the "PARE anticipé"), not the moment they become unemployed.
- The other purpose is, officially, to make sure that alternatives to the layoffs have been fully explored. The steps (which must take place before workers are notified of the layoff) grow more numerous with the size of the firm, and the size of the layoffs. For layoffs for personal reasons, the steps are typically minimal – an interview and the sending of an official letter. For layoffs for economic reasons, and for firms with more than 100 workers, the process can take up to half a year. The steps involve a number of meetings with the representatives of the workers, the presentation by the firm of a detailed "plan to save jobs" ("plan de sauvegarde de l'emploi"), the approval of the labor inspection office; they may also involve the nomination of an auditor if requested by worker representatives, and the recourse to an arbitrator if the workers' representatives disagree with the firm's plan.

At the end of this process, the firm can start the advance notice period, and then proceed with the layoffs. But the workers, if they disagree, can go to court. Different courts have different jurisdictions. In the case of collective layoffs, workers or firms go to regular tribunals, either "Tribunaux

d'instance" or "Tribunaux de grande instance". For individual layoffs, and most labor contract disputes, the standard court is the labor tribunal known as the "Prud'hommes", an institution created in 1806. Each such tribunal has two elected union representatives and two elected representatives from business organizations. In the case of a tie, the decisive vote is cast by a professional judge.

When a case is taken to the Prud'hommes, the first step is an attempt at arbitration ("audience de conciliation"). The second is a judgment (audience de jugement), which can decide that layoffs were not justified, and impose fines and payments to the firm (98 percent of the cases are brought by workers, only 2 percent by firms; 80 percent of the cases are decided in favor of workers). The judgment can then be appealed, going first to the appeals court ("Cour d'appel"), then possibly to the highest court ("Cour de cassation"); 50 percent of the cases are appealed, 70 percent are decided in favor of workers.

The number of cases taken by the Prud'hommes has increased rapidly in the recent past, reaching close to 200,000 new cases (half of those related to layoffs) per year at the end of the 1990s. Both at regular tribunals, and at the Prud'hommes, the delays in reaching a decision can be substantial (the mean time to the first judgment at Prud'hommes is now around 10 months. The Michelin case, now on appeal, is now more than 3 years old).

If layoffs are found not to be justified, the firm has to pay additional severance payments. These payments can be substantial. If for example the firm has more than 11 employees, and the worker has more than two years' seniority, severance payments must be at least equal to six months.

8.4 The sharp distinction between CDD versus CDI

Since the late 1970s, successive governments have introduced fixed-term contracts, called "contrats à durée determinée", or CDDs. These contracts still require a severance payment, but eliminate the recourse to courts when termination takes place at the end of the contract.[29]

A brief history of CDDs goes as follows: CDDs were introduced in 1979. With the election of a socialist government in 1981 and the passage of another law in 1982, their scope was reduced. A list of 12 conditions was drawn, and only under those conditions could firms use fixed-term contracts. In 1986, the 12 conditions were replaced by a general rule: CDDs should not be used to fill a permanent position in the firm. The current architecture dates for the most part to an agreement signed in March 1990.

Under this agreement, CDDs can be offered by firms for only one of four reasons: (1) the replacement of an employee on leave; (2) temporary increases in activity; (3) seasonal activities; (4) special contracts, aimed at facilitating employment for targeted groups, from the young to the long-term unemployed. The list of special contracts has grown in the 1990s, as

each government has tried to improve labor market outcomes for one group or another; some of these contracts require the firm to provide training, and many come with subsidies to firms.

CDDs are subject to a very short trial period, typically one month. They have a fixed duration, from 6 to 18 months depending on the specific contract type. Mean duration is roughly one year. They typically cannot be renewed, and, in any case, cannot be renewed beyond 24 months. If the worker is kept, he or she must then be hired on a regular contract (CDI). If the worker is not kept, he or she receives a severance payment equal to 10 percent of the total salary received during the life of the contract. (Note that this is a much higher percentage of salary than is the case for severance on regular contracts. But workers on CDDs cannot go to the Prud'hommes to contest the end of employment on the CDD.)

As we indicated earlier, these CDDs have been very popular with firms, and now represent 70 percent of the flow of hires, and a bit above 10 percent of total employment.

9 Contours of employment protection reform in France

When we compare the existing French system of employment protection to the structure that emerges from our analysis, we believe that there is a strong case for reform along two main lines:

- An increase in the marginal financial cost of layoffs for firms.
- A decrease in the role of courts in case of layoffs, leading to a less costly and less uncertain process for firms.

Or in more detail:

- An increase in the contribution rate of firms to the financing of unemployment insurance.

  Firms at this stage finance a large part of the unemployment insurance system; but they do so through a fixed-rate payroll tax, so the marginal contribution rate (through that tax) is equal to zero. Our analysis suggests the following conclusions:

  - This contribution rate should be positive, although probably less than one.
  - Starting from the current legislation, this implies a reduction in the payroll tax (on employers and employees), and the introduction of unemployment contributions by firms related to their layoff behavior. The contributions need not be made at the time of layoffs. Some form of experience rating (for example along the lines of the reserve ratio system described in Section 3) may be appropriate.

  - To avoid discrimination by firms against workers with different labor market prospects, the system should include both a trial period and a transition period during which contributions and benefits are less than and converge to the standard regime. Other measures may include basing the contributions by firms on the number of layoffs, rather than on expected or actual unemployment benefits.

- While shifting to a positive contribution rate will lead firms to reduce layoffs, this increase in employment protection (with payments from firms to the unemployment agency, rather than directly to workers) will be less visible to workers than some of the other forms of employment protection. But it is nevertheless an increase in employment protection. It leads firms to take into account the social costs of unemployment, and decrease their layoff rate.
- Is there a role left for severance payments, direct payments to workers? We think so, but their role should be only to offset the costs of job loss (as separate from unemployment). This should be their only and limited purpose; unemployment insurance is better provided through unemployment benefits.

  Given that the costs of job loss appears to be increasing and convex in seniority, this suggests the use of a schedule which is increasing and convex in seniority, with low payments until high seniority is achieved. We do not have a view as to whether the schedule currently in place has the right level of curvature. As noted earlier in the discussion of the Delalande contribution, there are constraints on how steep the schedule can be at high seniority. If it is too steep, it runs the risk of generating discrimination against middle age workers.
- In the case of bankruptcy, firms should be liable for contributions and severance payments due to their workers, and the state should be a senior creditor.

  As we know however from recent cases, firms have an incentive to escape those liabilities by designing complex structures of ownership so as to benefit from limited liability. The problem will only grow more serious, if, as we argue should happen, contribution rates are increased. The problem is an old one and has given rise to an intense debate in the area of environmental liability on how to reach solvent principals. Something can be learned from this debate both about the difficulties involved in assigning liability and about the techniques that can achieve this.

  If liability tracing cannot be achieved without creating large administrative costs or creating perverse incentives for the private sector, a simple alternative could be the deposit of some form of collateral: firm's assets (but state agencies have little expertise in assessing the value of this collateral and in monitoring that it is maintained ade-

quately), bank guarantees, or financial assets (interest, dividends, and possibly the principal would be adjusted over time so as to maintain the value per worker).

- Advance notice periods combined with retraining, and other measures designed to help laid off workers find another job (such as the congé reclassement, the PARE anticipé, and the PAP – an individualized "action plan" for training and job search), are highly desirable and justify delays between the layoff decision and its implementation.
- The heavy hand of the judicial process, as it now exists, seems however largely unjustified.

  We do not see why an arbitrator, the Prud'hommes or the other tribunals, the appeals court, and the Cour de cassation, should be asked to second guess the decision of the firm, if the firm goes through the proper administrative steps and is willing to pay both contributions to the state and severance payments to its workers.
- The role of the tribunals should therefore be much more limited than it is today. In particular, if a firm is willing to declare a separation a layoff and to pay the associated costs and severance payments, either the Prud'hommes or the Tribunaux d'instance should not second guess the firm's decision, and should not intervene. In the case of collective layoffs for example, the role of the courts should be to check that proper administrative steps have been taken, contributions and severance paid, not to assess whether the firm was justified in laying the workers off.
- The sharp contrast between the CDI and CDD regimes that exists today should be eliminated.

  At the short end of the seniority scale, there is however an important role for a trial period at the start of a contract, and termination during that period should not trigger either the payment of unemployment contributions or severance. The period should be long enough to allow both sides to learn about the match, but short enough to make it unprofitable for firms to use the trial period to rotate workers into a given job.

  At the long end of the seniority scale, we have argued that severance payments should be non linear in seniority, so that it is more expensive for the firm to lay off workers with high seniority. (The non linearity should however come in at much higher seniority levels than the one-year or two-year duration of fixed-term contracts.)

  The elimination of this two-contract regime should reduce the dual nature of the labor market, which we see as a major and perverse effect of recent reforms.
- The increase in the financial marginal cost of laying off a worker, compensated by a decrease in the complexity and the uncertainty of the layoff process, might well be more attractive both to firms and to workers. The example of CDDs, which combine a higher severance

pay than CDIs, with a much simpler process of termination, suggests that firms would be eager to accept such a trade off. But we believe that this need not come with a decrease in the welfare of workers, both those on CDDs, and those on CDIs. Given its goals, the current system is inefficient. Efficiency gains can make both sides better off.

Notes

* This chapter is an English adaptation of a report written for the French Conseil d'Analyse Economique. We thank Daron Acemoglu, David Autor, Olympia Bover, Pierre Cahuc, Daniel Cohen, Mathias Dewatripont, Francis Kramarz, Fiorella Padoa-Schioppa, Gilles Saint-Paul, Robert Solow, Nicolas Véron, Robert Wagner, and especially Denis Fougère, Jacques Freyssinet for discussions and comments. We thank Jean Christophe Bureau, Jean Yves Kerbourc'h and Frédéric Lerais for their help.

1 The mechanism is the following: the increase in employment protection is likely to increase firms' costs. This increase in costs leads in turn to lower job creation and so to lower hiring, until the resulting increase in unemployment duration, which makes unemployment more painful, reduces wages and returns costs to a level consistent with the required rate of return on capital.

2 For recent surveys, see OECD (1999a), and Addison and Teixera (2001).

3 Whether such indexes can successfully capture the many dimensions of employment protection is open to discussion. To a first approximation, we believe the ranking of countries implied by these indexes is not misleading.

4 For example, if the average duration of unemployment is six months, and the flow of workers who become unemployed each month is equal to 2 percent of the labor force, the unemployment rate is equal to 6×2 percent, or 12 percent.

5 See for example Nickell (1997).

6 For a review of the U.S. experience rating system and its potential implications for France, see Margolis and Fougere (1999).

7 Variations over states: in 1996, contribution rates (the ratio of tax payments by firms to benefits paid by the state to laid off workers) ranged from 8 percent in North Carolina to 86 percent in the state of New York. Variations over time: in 1984, the state of Washington moved from a zero contribution rate to a contribution rate of 50 percent.

8 In another ironic twist (relative to the spirit of the debate on employment protection in Europe), this line of research often starts from the presumption that increasing the contribution rate to one – in other words, increasing employment protection – would be desirable. The argument is that this would lead firms to fully take into account the social costs of unemployment. As we shall see, this argument is incomplete and the optimal contribution rate is probably less than unity.

9 See the symposium on this topic in the *Economic Journal*, 2002, including Dolado *et al.* (2002) for Spain (the country with the highest proportion of workers on fixed duration contracts), and Blanchard and Landier (2002) for France.

10 A formal model underlying the arguments (or to be honest, most of the arguments, as we have sometimes taken educated guesses beyond what our model could answer) in this and the next four sections is presented in Blanchard and Tirole (2003).

11 This is indeed the French solution, with the combination of the state run ANPE and the tripartite (government, business organizations, and unions) UNEDIC–ASSEDIC.

12 Two quotes from judges at the Prud'hommes, the French labor courts (translated from French): "Employers have a hard time understanding that the main issue is not the financial loss. Psychologically, a layoff is very tough. For the family. For your health. It puts your whole life into question." And "People put their lives in their jobs. And, at once, everything is taken away from them." (*Liberation* December 2, 2002.)
13 For example, Price (1992), and studies at the Michigan Prevention Research Center.
14 A fact consistent with this hypothesis (but also with a number of others): in France, there are four times as many quits as layoffs for workers with 2 to 9 years of seniority, but four times as many layoffs than quits for workers with 10 or more years of seniority (Goux and Maurin 2000: Table A1).
15 For usual incentive constraint reasons, the schedule cannot however be too steep; otherwise it would give incentives for firms to layoff workers at mid-career, i.e. when the severance payments associated with laying them off are still relatively low.
16 For a recent survey, see for example Frederiksson and Holmlund (2003).
17 See for example Boeri (2002).
18 Our understanding from Table 4 in Margolis and Fougère (1999) is that the proportion of contributions due but not paid because of bankruptcy is under 10 percent in most U.S. states, with some exceptions (for example, California with 13 percent).
19 Interestingly, the correlation between the stated contribution rate and the bankruptcy rate across U.S. states appears to be small (Margolis and Fougère 1999). If interpreted as a causal relation from the contribution rate to the probability of bankruptcy, this would imply that increasing the contribution rate may not have much impact on bankruptcy rates. But there are reasons to be skeptical of this causal interpretation. For example, firms with high risk are more likely to incorporate or move operations to states with a low contribution rate.
20 For an early discussion of the role and the scope of bonding in the context of employment protection, see Lazear (1990).
21 Think for example of public employees. Given the high degree of employment protection and the typically generous retirement benefits, many are willing to become public employees even if wages are lower than in the private sector. But, because they cannot be laid off, except at great cost, public employees are in a very strong bargaining position, and sometimes use it to extract higher wages or other advantages from the state.
22 There is a set of taxes and transfers which can achieve both efficient destruction and efficient creation. A contribution rate of one, so there is no distortion at the destruction margin. A subsidy to new jobs to eliminate the adverse effects of the increase in cost on job creation (see for example Mortensen and Pissarides (2001). But this raises in turn the issue of how these job subsidies themselves are financed (they may have to be very large). So our discussion here is predicated on the absence of, or at least on limits on, job subsidies.
23 This is why the line of argument used in the context of experience rating to argue that the contribution rate should be equal to one is misleading. Such a rate removes distortions at the destruction margin, but can have a large adverse effect on creation.
24 A case often analyzed in the research on labor market equilibrium is the case of zero bonding and a fully elastic supply of capital (see, for example, Pissarides (2000)). In that case, a strong – and depressing – result emerges. The "pain of unemployment", more specifically the difference in the value of being employed over the value of being unemployed, remains constant. What the

unemployed gain relative to the employed through, for example, higher unemployment benefits when unemployed, they must lose in equilibrium through higher unemployment duration. (Otherwise, wages would be too high, profits too low, and firms would not create jobs.) The result is extreme, but an important warning nevertheless that general equilibrium effects can lead to effects quite different for those intended by the policy maker.

25 This statement may be too strong, as there might still be some room for gaming. If for example, the worker receives an outside offer and the firm simultaneously receives a bad productivity shock, both have an incentive to having the other side take the decision to separate.

26 Anderson and Meyer (1998) show that the 1985 increase in the contribution rate in the state of Washington led to a substantial increase in the number of denial of benefit cases brought up by firms.

27 For a similar discussion in a different context (Who should run prisons?), see Hart *et al.* (1997).

28 Three useful sources on French institutions are Pélissier *et al.* (2002) (which presents the legal structure), CFDT (2003) (which gives a user guide for workers) and JurisClasseur Groupe Lexis-Nexis (2002) (which gives the text and interpretation of the 2002 law, called "Loi de Modernisation Sociale").

29 Poulain (1994) gives a detailed description of the rules governing CDDs.

References

Abowd, J. and Kramarz, F. (1997) *The Costs of Hiring and Separations*, NBER WP 6110.

Addison, J. and Teixera, P. (2001) *The Economics of Employment Protection*, IZA DP 381.

Akerlof, G. and Miyazaki, H. (1980) "The Implicit Contract Theory of Unemployment Meets the Wage Bill Argument", *Review of Economic Studies* 48: 321–38.

Anderson, P. and Meyer, B. (1998) *Using a Natural Experiment to Estimate the Effects of the Unemployment Insurance Payroll Tax on Layoffs, Employment, and Wages*, mimeo Dartmouth College.

Autor, D. (2001) *Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing*, MIT.

Autor, D., Donohue, J., and Schwab, S. (2002) *The Costs of Wrongful-Discharge Laws*, MIT WP 02-41.

Azariadis, C. (1975) "Implicit Contracts and Underemployment Equilibria", *Journal of Political Economy* 83: 1183–202.

Baicker, K., Goldin, C., and Katz, L. (1997) *A Distinctive System: Origins and Impact of U.S. Unemployment Compensation*, NBER WP 5889.

Baily, M. (1974) "Wages and Employment Under Uncertain Demand", *Review of Economic Studies* 41: 37–50.

Bentolila, S. and Blanchard, O. (1990) "Spanish Unemployment", *Economic Policy* 234–81.

Bertola, G. and Rogerson, R. (1997) "Institutions and Labor Reallocation", *European Economic Review* 41 (6): 1147–61.

Blanchard, O. and Landier, A. (2002) "The Perverse Effects of Partial Labor Market Reform: Fixed Duration Contracts in France", *Economic Journal* 112: F214–44.

Blanchard, O. and Portugal, P. (2001) "What Hides Behind an Unemployment

Rate. Comparing Portuguese and U.S. Unemployment", *American Economic Review* 91 (1): 187–207.
Blanchard, O. and Tirole, J. (2003) *The Design of Optimal Labor Market Institutions*.
Blanchard, O. and Wolfers, J. (2000) "Shocks and Institutions and the Rise of European Unemployment. The Aggregate Evidence", *Economic Journal* 110 (1): 1–33.
Boeri, T. (1999) "Enforcement of Employment Security Regulations, On-the-job Search, and Unemployment Duration", *European Economic Review* 43 (1): 65–90.
Boeri, T. (2002) *Making Social Europe Compete*, mimeo, Universita Bocconi.
CFDT (2003) *Salariés: Guide de vos droits 2003*, La Decouverte; Guides.
Dolado, J., Garcia Serrano, C., and Jimeno, J. (2002) "Drawing Lessons from the Boom of Temporary Jobs in Spain", *Economic Journal* 112: F270–95.
Feldstein, M. and Altman, D. (1998) *Unemployment Insurance Savings Accounts*, NBER WP 6860.
Fougère, D. and Margolis, D. (2000) "Moduler les cotisations employeurs à l'assurance chômage: Les expériences de bonus malus aux état unis", *Revue Francaise d'Economie* (2): 3–76.
Frederiksson, P. and Holmlund, B. (2003) "Improving Incentives in Unemployment Insurance: A Review of Recent Research", *CES* IFO working paper (922).
Goux, D. and Maurin, E. (2000) *Labor Market Institutions and Job Stability: A Firm Level Analysis of Layoff Risk for High and Low Seniority Workers*, WP Crest, 29.
Hart, O., Shleifer, A., and Vishny, R. (1997) "The Proper Scope of Government: Theory and an Application to Prisons", *Quarterly Journal of Economics* 112 (4): 1127–62.
JurisClasseur Groupe Lexis-Nexis (2002) *Le licenciement pour motif economique après la loi de modernisation sociale*, under the direction de Christophe Willmann et Jean-Yves Kerbough.
Kugler, A. (2002) *From Severance Pay to Self Insurance: Effects of Severance Payment Savings Accounts in Colombia*, CEPR DP 3197.
Lazear, E. (1990) "Job Security Provisions and Employment", *Quarterly Journal of Economics* 105 (3): 699–726.
Margolis, D. and Fougere, D. (1999) "Moduler les cotisations employeurs à l'assurance chomage: Les experiences de bonus-malus aux Etats Unis", *Rapport sur l'architecture des prelevements en France*, Conseil d'Analyse Economique.
Mortensen, D. and Pissarides, C. (2001) *Taxes, Subsidies, and Equilibrium Labour Market Outcomes*, CEPR DP 2989.
Nickell, S. (1997) "Unemployment and Labor Market Rigidities: Europe Versus North America", *Journal of Economic Perspectives* 11 (3): 55–74.
OECD (1999a) *Employment Outlook*, OECD, Chapter 2.
OECD (1999b) *OECD Employment Outlook*, OECD.
Pélissier, J., Supiot, A., and Jeammaud, A. (2002) *Droit du Travail*, Dalloz: Paris.
Pissarides, C. (2000) *Equilibrium Unemployment Theory* (2nd edn), MIT Press.
Poulain, G. (1994) *Les contrats de travail à durée déterminée*, Litec: Paris.
Price, R. (1992) "Psychosocial Impact of Job Loss on Individuals and Families", *Current Directions in Psychological Science* 1-1.
Ringleb, A. and Wiggins, S. (1990) "Liability and Large Scale, Long-term Hazards", *Journal of Political Economy* 98: 574–95.

5 Trashing J.B. Say

The story of a mare's nest*

Robert W. Clower

> If it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic.
>
> Tweedledum to Tweedledee, Lewis Carroll[1]

J.B. Say (1767–1832) merits notice in the twenty-first century as the foremost early interpreter of Adam Smith, the first professor of political economy in France, the innovative and influential nineteenth-century author of textbooks that dominated academic teaching of 'political economy' through the first half of that century[2] and, in my view, the best all round economist of the 'classic' generation that included Ricardo and Malthus. But Say is best known today for various twentieth-century propositions called 'Say's Law' by Keynes, Lange, Patinkin and other notable moderns who, perhaps inadvertently, have made a mockery[3] of their monumentally worthy and distinguished predecessor.

1 What Say said

The trashing of J.B. Say derives mainly from unscholarly treatment by later writers of Say's short chapter 'Des Débouchés' (Market Openings) in the first edition of his *Traité D'Économie Politique* (1803: Bk. I, ch. xxii, pp. 152–4) where Say summarizes Smith's transition in *The Wealth of Nations* from chapter II (on the division of labour) to chapter III (on limits imposed by the extent of the market). Say effectively collapsed the central ideas in Smith's two chapters (see especially pages 16 and 22 of Smith (1776)) into a single paragraph, as follows [*my translation*]:

> A single industrious man in a group of people who produce nothing beyond coarse food, can do nothing with produce beyond what is required to purchase coarse food for himself. The same individual in a group of one hundred thousand similarly industrious persons will confront numerous markets where he can advantageously sell his own produce to pay for that of others. Money plays only an intermediary

> role in this process. When the exchange is completed, it is discovered that *products pay for products* [emphasis mine].

A century later Charles Gide (1904: 148) summarized Say's chapter xxii – which he called the 'theory of markets' – as stating:

> Every commodity will find a sale more readily with every increase in the variety and abundance of other commodities.

About the same time (1909), in an effort to strengthen elementary economics teaching at the University of Michigan, F.M. Taylor called Say's *théorie* 'Say's Law': 'the principle that products constitute at once the demand for goods and the supply of goods, ... and so, if we assume production to be directed in accord with individual wants, supply and demand must necessarily be equal'.

Some may be puzzled by the apparent addition over the years of 'excess baggage' to Say's original thoughts about markets – interested readers may discover plausible answers in Sowell's (1972) book on Say's Law. Personally, I am puzzled by the lack of awareness by so many writers of numerous early antecedents (starting with Aristotle) of Say's so-called 'theory of markets'.[4]

About thirty years before Say published the first edition of his *Traité*, his countryman Turgot (1770: 3) wrote:

> If ... all the inhabitants of a country ... each had precisely the quantity of [land] necessary for his support and nothing more it is evident that ... [no one] would possess anything with which to pay for the labour of another.

And a few pages later [p. 28] continued as follows:

> Reciprocal want has led to the exchange of what people want for what they have not. People exchange one kind of produce for another, or produce for labour.

Earlier, David Hume (1752) had noted:

> A single man can scarcely be industrious, where all his fellow-citizens are idle. The riches of the several members of a community contribute to encrease my riches.... They consume the produce of my industry, and afford me the produce of theirs in return.

Earlier still, Aristotle in *The Nicomachean Ethics* (Welldon translation as excerpted in Monroe (1924: 28)), observed:

> Demand ... is ... a principle of unity binding society together ... because, if there is no mutual demand on the part of two persons, if neither of them or one only needs the services of the other, they do not effect an exchange, whereas, if somebody wants what somebody else has, e.g., wine, they effect an exchange, giving the wine in return for the right of importing corn.

Given these and related early statements of what Gide and Rist (1930: 115) later called the:

> vague, self-evident notion [that] 'products are given in exchange for products'

the real puzzle is why Say's *Des Débouchés* attracted so much later attention. What Say had to say on the subject could more appropriately be named *Say's Platitude* than 'Say's Law'. That being said, I have only to add that to link *Say's Platitude* with any proposition parading as a 'Law' would be fatuous.

2 Keynes the phrasemaker

Say still figured as a generally admired classical economist into the early 1930s. Thus Alexander Gray, in his classic *Development of Economic Doctrine* (1931), wrote:

> The shadow of Say [lies] heavily across the text-books of the nineteenth (and perhaps the twentieth) century. [...] Apart from [his] ... work of vulgarization, certain ideas were peculiarly Say's own. Of these, the one which had in its time most renown ... and which was indeed for long regarded as Say's passport into the company of the immortals, was the once celebrated 'theory of markets' (*la théorie des débouchés*). This theory ... is to the effect that goods and services are only superficially bought with money; they are, in fact, bought with other goods and services. The money is merely the 'carriage' which, having effected the exchange of two commodities, will forthwith proceed to exchange others. But in reality, products are always exchanged against other products. Consequently, he tells us one ought not to say: 'Sale does not take place because money is scarce, but because other products are so'. It is likewise an integral part of the theory that a product, when created, offers from that very moment a market for other products. General over-production is thus an impossibility. If certain products are in excess, it is because there is a deficiency elsewhere, and the cure for 'over-production' in one direction is therefore more production elsewhere to serve as a *débouché* for this excess.

So much for the J.B. Say of *unslanted* doctrine history. The publication of Keynes's *General Theory* in 1936 brought a sudden end to Say's 'century in the sun'. At a time when Say's theory of markets had largely disappeared as a topic of contemporary interest, Keynes poured burning coals on the long-extinguished 'fire' by oracularly asserting (JMK VII: 18):

> From the time of Say and Ricardo the classical economists have taught that *supply creates its own demand* [italics mine]; – meaning by this some significant, but not clearly defined, sense that the whole of the costs of production must necessarily be spent, in the aggregate, directly or indirectly, on purchasing the product.

And a few pages later (JMK VII: 21–2):

> Supply creates its own demand in the sense that the aggregate demand price is equal to the aggregate supply price for all levels of output and employment.

Keynes cited no source for the phrase *supply creates its own demand*. However, in a letter to Kahn (JMK XIII: 422–3) of April 1934, Keynes wrote:

> I have been making rather extensive changes in the early chapters of my book, ... consequential on a ... precise definition of ... effective demand:
>
> Let W be the marginal prime cost ... when output is O.
>
> Let P be the expected selling price of this output.
>
> Then OP is effective demand.
>
> The fundamental assumption of classical theory, 'supply creates its own demand', is that OW = OP *whatever* the level of O, so that effective demand is incapable of setting a limit to employment.... On *my* theory OW ≠ OP for *all* values of O, and entrepreneurs have to choose a value of O for which it *is* equal; – otherwise the equality of price and marginal prime cost is infringed. This is the starting point of everything.

To this day, the source of the Keynes phrase 'supply creates its own demand' remains a mystery. I suspect that, before 1936, the phrase was an 'oral tradition' in Cambridge among Joan Robinson and some of her friends, derived perhaps from J.S. Mill's book on *Unsettled Questions of Political Economy* (1844: 73) where Mill wrote:

> Nothing is more true than that it is produce which constitutes the market for produce, and that every increase of *production*, if distributed without miscalculation among all kinds of produce in the

> proportion which private interest would dictate, *creates*, or rather constitutes, *its own demand* [italics mine].

Mill here refers to *production* instead of *supply*, but nowhere else in pre-Keynesian literature can I find a remark that output (in some sense) *creates its own demand*. Whatever its original source, the phrase 'supply creates its own demand' is a far cry from Say's phrase that 'products pay for products'.

After much pondering of Keynes's remarks on Say and classical economics in the *General Theory* and earlier (see, e.g., Keynes's 1934 'The Listener' article as reprinted in (JMK XIII: 485–92)), I now conjecture that Keynes confused Mill's interpretation of Say's 'theory of markets' with the problematic presumption that seems to run through the writings of most pre-Keynesian economists, to the effect that there are unspecified forces (invisible fingers or hands?) operating 'behind the scenes' to ensure short-run *viability* of private-ownership economies. On this basis, I am inclined to regard Keynes's somewhat grotesque restatement of Say as no more than an *unintentional* trashing – perhaps a put down aimed more at Mill or Ricardo than at Say.

3 Lange's nonlogical 'laws'

Unfortunately, Keynes's rendition of 'Say's Law' as *supply creates its own demand*, following several years of discussion as to the basis of Keynes's indictment of 'classical economics', became the central target of later detractors of Say, starting with Oscar Lange and his puzzlingly specious and logically flawed, but curiously influential, paper 'Say's Law: a restatement and criticism' (Lange 1942: 49). Lange opened his paper with an outright burlesque of Say's 'law of markets':

> Say's Law is the proposition that there can be no excess of total supply of commodities (general oversupply) because the total supply of all commodities is *identically* equal to the total demand for all commodities [. . .]. Associated with it is the proposition that there cannot be such a shortage of total entrepreneurial receipts relative to total entrepreneurial cost as to cause losses throughout the whole economy (general overproduction).

As a basis for the 'restatement' of Say, Lange considers

> a [Hicksian, by Lange's fn. 1, p. 49] closed system in which *n* commodities are exchanged, one of them – say, the *n*th commodity – functioning as medium of exchange as well as *numéraire*, i.e., as money.

Continuing,

> Denote by p_i the price of the ith commodity. We have $p_n \equiv 1$. Let $D_i = D_i\ (p_1, p_2, \ldots, p_{n-1})$ and $S_i = S_i\ (p_1, p_2, \ldots, p_{n-1})$ be the demand function and the supply function, respectively, of the ith commodity.

Lange then departs from all previous theory (and good sense) by asserting,

> There are only $n-1$ independent demand functions and $n-1$ independent supply functions, the demand and the supply function for the commodity which functions as money being deducible from the other ones. We have

$$\sum_{i=1}^{n-1} p_i D_i \equiv S_n \tag{2.3}$$

and

$$\sum_{i=1}^{n-1} p_i S_i \equiv D_n \tag{2.4}$$

As a matter of logic, the Lange identities (2.3) and (2.4) *redefine* S_n and D_n in terms that ignore (and are inconsistent with) the 'Hicksian' functional definitions with which Lange's argument begins.[5] Instead of S_i in (2.3), Lange should have introduced a **new** variable, say $\mathbf{V_d}$, the *aggregate numéraire value of planned purchases of non-numéraire goods*, and similarly for (2.4), a **new** variable, say, $\mathbf{V_s}$, to represent the *aggregate numéraire value of planned sales of non-numéraire goods*.[6]

Now using the valuation notation (**V**) introduced above, Lange's equations (2.5) and (2.6) can be rewritten as in Lange's (2.7):

$$\sum_{i=1}^{n-1} p_i D_i + D_n \equiv \sum_{i=1}^{n} p_i D_i + \mathbf{V_d} \equiv \sum_{i=1}^{n-1} p_i S_i + S_n \equiv \sum_{i=1}^{n} p_i S_i + \mathbf{V_s}$$

or as,

$$\sum_{i=1}^{n} p_i D_i \equiv \sum_{i=1}^{n} p_i S_i - (\mathbf{V_d} - \mathbf{V_s}) \tag{LL}$$

Then *if and only if* the term $(\mathbf{V_d} - \mathbf{V_s})$ is *identically* zero (i.e., if and only if the *aggregate numéraire value of non-numéraire goods demanded is identically equal to the aggregate numéraire value of non-numéraire goods supplied*) can we write (**LL**) – **Lange's Law** – as

$$\sum_{i=1}^{n} p_i D_i = \sum_{i=1}^{n} p_i S_i \tag{WL}$$

i.e., 'the identity' that Lange [p. 50] named ***Walras' Law***. Evidently if we further assume

$$S_n = D_n$$

i.e., that the supply of the *numéraire* good, S_n (Lange's 'money'!), is equal to the demand for it, we obtain Lange's Identity (4.1) [p. 52] which he named ***Say's Law***, described by Lange on this occasion as stating that 'the total demand for commodities (exclusive of money) is *identically* equal to their total supply'. It is obvious from the foregoing analysis that what Lange calls ***Say's Law*** is exactly the same mathematical proposition that Lange called ***Walras' Law*** (**WL**) above, the only difference being that ***Say's Law***, by contrast with (**WL**), applies to a 'system' with just $n-1$ rather than n commodities!

Now, Lange's formal *statement* of **WL** in his paper corresponds exactly to the linear dependence identity proved to hold for *contemplated* trades of individual traders by Walras (1926: 128) and by Hicks (1939: 314). But that Lange's purported 'proof', which nowhere discusses individual traders, is invalid may be inferred from Patinkin's opening assertion in his *New Palgrave* (1987: 863) article on Walras' Law:

> Walras' Law (so named by Lange, 1942) is an expression of the interdependence among excess-demand equations of a general equilibrium system that stems from the budget constraint.

Patinkin could not have drawn this thought from Lange's paper; Patinkin referred only to the initial step in all valid proofs of Walras' Law – introduction of budget constraints for all individual 'agents', as occurs in the books of Walras and Hicks, and in the work of all later writers – with the *exception* of Lange! Compounding errors in the body of his paper, Lange offers a second erroneous proof of **WL** in fn. 2 [pp. 50–1], where he works with a double subscript notation and fails to recognize that the symmetry conditions required for his 'proof' to be valid are true *only if* trade within and across 'trading posts' (Walras 1926: 119) is so perfectly organized and rivalrously competitive that there is 'no occasion for advantageous arbitrage transactions'. Only on this restrictive assumption (compare Cournot 1838: 38–40) of 'no advantageous arbitrage operations available' does Lange's 'proof' go through.

One is naturally inclined, on the basis of the present account of flaws in Lange's 'celebrated' paper, to question Lange's honesty or his competence, or both. But charity suggests a kinder explanation for the farrago of nonsense[7] in Lange's paper: the paper was *not* refereed by a critical reader before it was published (Lange was one of the three editors of the volume in which the paper was first published).

4 Concluding observations

Despite its logical errors, Lange's 1942 paper on Say's Law profoundly affected subsequent treatments of both macro- and microeconomic theory. By stating 'Say's Law' in terms reminiscent of the phrase 'supply creates its own demand', Lange effectively converted Keynes's problematic epigram into a 'scientific' formula (Sowell 1972: 1–6). Further consequences ensued, for although both 'proofs' of the identity that Lange named 'Walras' Law' were erroneous, the proposition itself (**WL**) was thereafter treated as established doctrine – initially because the formally valid 'linear dependence' proofs of Walras (1954) and Hicks (1939) were deployed by well-known authors of a series of papers on 'classical' monetary theory (see references in Brunner (1951)), later through heavy emphasis on 'Walras' Law' by Patinkin in his best-selling *Money, Interest and Prices* (1956, 1965, 1989) which, in company with the magisterial volumes of Hicks (1939) and Samuelson (1947), dominated graduate economic theory courses throughout the world for more than two decades.

Surprisingly, some authors (even in the twenty-first century) may believe that 'Lange's Laws' are in the nature of 'accounting identities' (cf. Becker and Baumol 1952: 600), failing to recognize that the demand and supply variables of Walras and later theorists all concern *dispositions á l'enchére* – 'planned', 'notional', 'target', 'virtual' – never '*actual*', '*realized*', or '*active*' variables[8] such as might be collected in real-world studies. Walras' focus – as should be the case for all equilibrium theorists – is not on *trade* but only on *notionally contemplated trades*.

Little wonder, therefore, that with its wealth of accumulated analytical propositions and dearth of well-confirmed empirical results, modern economic theory should be widely regarded as a 'dismal' excuse for serious science. However that may be, there can be little doubt that common opinion of present-day economics is adversely affected by the proclivity of many theorists to confuse *formal models* of observed phenomena with descriptive accounts of actual events. The proclivity has never been more prominent than in modern treatments and trashings of the inspired work of Jean Baptiste Say.

Notes

* I first met Professor Fitoussi in 1987 at a meeting in Aalborg, Denmark hosted by K. Velupillai, who had earlier introduced me to Fitoussi's *Modern Macroeconomic Theory* (1983), used as a graduate text at UCLA. It is now my pleasure to welcome Professor Fitoussi to the cohort of the contented 'over 60s'. As a preliminary, I am pleased to acknowledge the always helpful advice and editing assistance of my wife Georgene. Finally, I must mention the contribution of my ex-typist and long-time research assistant, Q.Y. Zuxzzs. He takes personal responsibility for all mistakes.

1 Carroll's whimsical reflection is aptly complemented by a related observation of Manin (1977: 34): 'Logic does not concern itself with the external world, but only with systems for trying to understand it.'

2 See Palmer (1997: 1–5, 67–89); Palmer is perhaps the world's leading historian of

the Napoleonic era, and is most certainly the leading biographer of J.B. Say [the full title of Palmer's book is *J.B. Say: An Economist in Troubled Times*]. See also the article on Say in the (old) Palgrave *Dictionary of Political Economy* (1925: 357–8). Say's extensive writings spawned numerous texts by others, e.g., Thomas Cooper's *Lectures on the Elements of Political Economy (2nd edn 1829)*. Cooper (later President of the University of South Carolina), based his lectures on the 1821 Prinsep (English) translation of the fourth edition of Say's *Treatise*.

3 A veritable 'mare's nest', as suggested by my title – described in the 1971 *Compact Edition of the Oxford English Dictionary* as: 'Originally … to have discovered something wonderful, which in fact has no existence. Hence, *an illusory discovery, esp. one that is much vaunted and displays foolish credulity*' [emphasis mine].

4 For further details see Paul Lambert's paper on Say in the IEA Translation series Paper No. 6 (1956).

5 On this, compare Hicks (1939: 58): 'We shall find it convenient, when dealing with multiple exchange, always to take some particular commodity as a standard of value (*numéraire*). [. . .] then it must be clearly understood that it has not yet been given any more of the qualities of money than these – that it is an object of desire, and that it is used as a standard of value.'

6 The equations (2.3) and (2.4) remind me (in 2003) of my discussion (Clower (1967: 4–5)) of a *pure money economy* where I introduced a dichotomized budget constraint and stated the aphorism *money buys goods and goods buy money, but goods do not buy goods*. I do not believe the thoughts set out then were in any way suggested by Lange's 1942 paper.

7 On this see Clower and Leijonhufvud (1973) and Johnsson (1997).

8 See Arrow and Hahn (1971) for explanations of this terminology.

References

Aristotle, *The Nicomachean Ethics* (Welldon translation, see Monroe (1924)).

Arrow, K. and Hahn, F. (1971) *General Competitive Analysis*, San Francisco, CA: Holden-Day.

Becker, G.S. and Baumol, W.J. (1952) 'The Classical Monetary Theory: The Outcome of the Discussion', *Economica* November: 355–76 (Reprinted in 1976, 'considerably modified and enlarged', but with all original errors repeated), in *Selected Economic Writings of William J. Baumol* No. 31.

Brunner, Karl (1951) 'Inconsistency and Indeterminacy in Classical Economics', *Econometrica* 19, April.

Carroll, Lewis (C.L. Dodgson) (no date) *Alice in Wonderland*, New York, NY: Modern Library.

Clower, Robert W. (1967) 'A Reconsideration of the Microfoundations of Monetary Theory', *Western Economic Journal* 6, December: 1–8.

Clower, Robert W. and Leijonhufvud, A. (1973) 'Say's Principle, What it Means and Doesn't Mean', *Intermountain Economic Review* 4, Fall: 1–16.

Cooper, Thomas (1829) *Lectures on The Elements of Political Economy* (second edition), Columbia, SC.

Cournot, A.A. (1838) *Researches sur les principes mathématique de la théorie des richesses*, Paris: Hachette.

Fitoussi, J.-P. (1983) *Modern Macroeconomic Theory*, ed., J.P. Fitoussi. Oxford: Blackwell.

Gide, Charles (1904) *Principles of Political Economy* (second American edition), Boston, MA: D.C. Heath.
Gide, Charles and Rist, Charles (1930) *A History of Economic Doctrines* (translation of second (1912) French edition), Boston, MA: D.C. Heath.
Gray, Alexander (1931) *The Development of Economic Doctrine*, London: Longmans Green and Co. Ltd.
Hicks, J.R. (1939) *Value and Capital*, Oxford: The Clarendon Press.
Hume, David (1752) 'Of Money', in *David Hume, Writings on Economics*, edited by E. Rotwein. Madison, WI: University of Wisconsin Press, 1970.
Jonsson, Petur O. (1997) 'On Gluts, Effective Demand, And The True Meaning of Say's Law', *Eastern Economic Journal* 23, Spring: 203–18.
Keynes, J.M. (cited as JMK #), *The Collected Writings of John Maynard Keynes* (various dates), London, Macmillan.
Lambert, Paul (1956) 'The Law of Markets Prior to J.B. Say and the Say–Malthus Debate', *International Economic Papers* No. 6, London, Macmillan.
Lange, Oscar (1942) 'Say's Law: A Restatement and Criticism', in *Studies in Mathematical Economics and Econometrics*, edited by Oscar Lange, Francis McIntyre and Theodore O. Yntema. Chicago, IL: The University of Chicago Press, pp. 49–68.
Manin, Iu I. (1977) *A Course in Mathematical Logic*, New York, NY: Springer-Verlag.
Mill, John Stuart (1844) *Essays on Some Unsettled Questions of Political Economy*, London: John W. Parker.
Monroe, A.E. (1924) *Early Economic Thought*, Cambridge, MA: Harvard University Press.
Palgrave, R.H.I. (1925) *Dictionary of Political Economy* 3 vols.
Palmer, R.R. (1997) *J.B. Say: An Economist in Troubled Times*, Princeton, NJ: Princeton University Press.
Patinkin, Don (1956, 1965, 1989) *Money Interest and Prices*, 1st edn, Evanston: Row Peterson; 2nd edn, Harper & Row; abridged 2nd edn, MIT Press.
Patinkin, Don (1987) 'Walras's Law', *The New Palgrave* 4, New York, NY: Stockton Press.
Samuelson, Paul A. (1947) *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.
Say, J.B. (1803) *Traité D'Économie Politique*, 1st edn, Paris: Deterville.
Say, J.B. (1821) *A Treatise on Political Economy*, Prinsep edn, reprinted by Kelley, New York, 1964.
Smith, Adam (1776) *An Inquiry into the Nature and the Causes of the Wealth of Nations*, Oxford: Clarendon Press, edited by R.H. Campbell and A.S. Skinner (1976).
Sowell, Thomas (1972) *Say's Law: An Historical Approach*, Princeton, NJ: Princeton University Press.
Taylor, F.M. (1909) 'Methods of Teaching Elementary Economics at The University of Michigan', *Journal of Political Economy*: 688–703.
Turgot, A.R.J. (1770) *Reflections on the Formation and the Distribution of Riches*, New York, NY: Macmillan, 1922.
Walras, L. (1954) *Élements D'Économie Politique Pure* (*edition definitive*), Paris: Pichon, 1926. Translated by W. Jaffe as *Elements of Pure Economics*, London: Allen & Unwin, 1954. All page citations in the text refer to the Jaffe translation, and so appear in the form 'Walras (1954, p. #)'.

6 Assets, debts and interest in the EU and the US

The slump in Europe revisited[1]

Jacques Le Cacheux

Introduction

Is "Old Europe" in decline and if so why? Over the past two decades at least, the growth performance of the US economy has systematically outpaced that of the European Union (EU), and especially that of its larger members. Why is the former living through "prosperity" while the latter has been experiencing mostly "depression"? (Prescott 2002) Almost three decades of "falling behind" have followed the three decades of "catching-up" in the aftermath of the Second World War. Not really depression, but sluggish economic growth, so that average per capita output in most European countries has not kept pace with that of the US, and poor employment performance and persistent mass unemployment, in all large countries of the Continent. Back in the mid-1980s, there were already signs of this divergence, and mounting evidence of a "slump in Europe" (Fitoussi and Phelps 1988). At the time, the most common suspect was the macroeconomic "policy mix", and its consequences on interest rates, exchange rates and prices: the focus was on international transmission mechanisms.[2]

Over the last ten years, the emphasis in the majority of analyses of this lasting phenomenon has shifted towards the idea of a "structural slump" (Phelps 1994). Among the most favored candidates for the leading cause of European poor economic performance, labor market rigidities, high taxes and insufficient effort to invest in R&D and education have usually been deemed important, whereas macroeconomic policies were blamed merely for the lack of fiscal consolidation in large Euro zone member countries. But with hindsight, none of these explanations alone appears entirely convincing and it seems that the original Fitoussi and Phelps hypothesis of a significant role for asset prices and exchange rates movements in a negative international transmission mechanism is worth investigating again. The story has to be different though, insofar as the context has changed, especially with financial globalization, which may explain some features of the current situation that were not so apparent back then. Without denying the existence of "structural" differences between the US

and European economies, this chapter explores an alternative, and rather heuristic, hypothesis emphasizing the paramount role of asset prices and debt accumulation and valuation in the processes of economic growth and of international transmission of macroeconomic shocks. Financial markets and macroeconomic policies play a major part in these processes because they contribute, in ways to be spelled out below, to the dynamics of assets and debts and their induced effects on individual behaviors, of firms and households.

The chapter is organized as follows. The first section briefly reviews a number of popular explanations of the relatively poor economic performance of Europe, especially of larger member countries of the Euro zone,[3] compared to that of the US in particular; it concludes that explanations in terms of "structural" weaknesses of Europe cannot be sufficient. In section 2, we focus on the processes of asset and debt accumulation and introduce demographic factors, along the lines of the life-cycle hypothesis, in order to characterize an important and often neglected element of difference between the US and the European Union (EU). Section 3 focuses on financial markets, and the way they price assets and debt instruments, in order to stress the importance, for macroeconomic analyses, of looking at individual agents' balance sheets, rather than simply considering net assets as in existing, representative agent macroeconomic models. Section 4 offers an analysis of international transmission mechanisms and emphasizes the strong financial links across the Atlantic and the striking co-movements in asset price and nominal returns that have been observed in the 1990s and in the first years of the new century. In section 5, we then show that this convergence, along with markedly different orientations in monetary policies, has distinctly different implications for debts, hence on the decisions to save and spend, in each of the two economies. Finally, section 6 offers some tentative conclusions and sketches directions for further research.

1 Tales of two continents

In the mid-1980s, at a time when it was becoming clear that the US economy had been recovering quite remarkably from the stagflation of the second half of the 1970s and from the deep recession following the fierce fight against inflation led by the Fed under the chairmanship of Paul Volcker, Europe was still struggling with low growth, high unemployment and relatively high inflation. To explain the "slump in Europe", Fitoussi and Phelps (1988) emphasized the leading role of macroeconomic policy mixes in the US and in Europe, giving rise to a negative, international transmission mechanism through the channels of asset prices and exchange rates changes. More specifically, the US policy mix and its macroeconomic consequences were shown to induce a rise in real interest rates, both in the US and worldwide, as well as an appreciation of the US

dollar vis-à-vis European currencies. The conjunction of these two developments worsened the output–inflation trade-off in Europe, essentially through higher mark-ups being chosen by firms in their pricing decisions on European markets. Thus, disinflation in Europe was made more costly, both in terms of unemployment, and because it led, through a standard Sargent–Wallace (1981) "unpleasant arithmetic", to an accumulation of public debts. The European situation was further aggravated by the perverse functioning of the European Exchange Rate Mechanism (ERM) in a context fraught with asymmetries among European economies.[4]

The recovery in Europe in the late 1980s somehow muted the debate and the consequences of German unification and of the progress towards monetary union in Europe tended to shift emphasis on other issues and mechanisms. However, more recently, the persistent and widening gap in economic performance of the two sides of the Atlantic Ocean has attracted attention from many economists, even among the "best and brightest". For the most, their explanations have tended to focus on "structural" causes.[5] In many European analyses, the single most important structural problem in Europe appears to be labor market rigidities: insufficient wage flexibility and very low labor mobility have often been singled out as impeding necessary adjustments, especially in an era of innovation and structural change induced by the combination of technical progress and economic globalization. Thus, due to their obsolete and excessively rigid labor market institutions and regulations, as well as their overly generous welfare systems, weighing too heavily on labor costs and inducing too strong disincentive effects on individual labor supply and effort, the larger European countries are said to have been unable to seize the benefits from the wave of innovations and market dynamism that characterized the 1990s.[6] However, a number of authors[7] have convincingly shown that labor market institutions and labor relations are so varied that one cannot associate a specific institutional configuration to a particular performance.

Other analyses also stress supply-side difficulties and the negative effects of public policies in Europe. Thus, for instance, excessive tax burdens are seen as "the" single most important explanation of the poor economic performance of most European countries compared to the US in a recent paper by Prescott (2002) who blames the French "depression" on the large tax wedges existing on consumption, labor and capital incomes, and their disincentive effects on supply decisions. But once again, available empirical evidence[8] on the supply effects of tax wedges suggests that they are not large enough to explain observed differences in aggregate performance.

In a somewhat related vein, though from a less traditional perspective, Phelps (1994) has developed an approach that may be termed an "asset view" of macroeconomics, and applied it to the different circumstances observed in the US and in Europe. He argued that decisions by firms to

accumulate various types of assets – physical, productive capital, human capital and customers (or market shares)[9] – are the major determinants of their supply and pricing decisions, which in turn shape the macroeconomic consequences of the various policy mixes. Such a hypothesis appears more appropriate for characterizing the European situation over the last decades, with relatively low investment in productive capital, in R&D and in higher education, as well as relatively less competitive goods markets. But although Phelps' framework also allows to explicitly investigate the interaction between "structural" features of the economy and macroeconomic policies, the choice of simple theoretical frameworks with representative, infinitely lived agents implies that macroeconomic policies mostly have "non-Keynesian" effects, and public debts crowd out private asset accumulation.

2 Ageing and capital accumulation in an international perspective

Although the emphasis on the central role of asset accumulation in Phelps' analysis appears a welcome addition to the Fitoussi–Phelps' framework to understand the persistence of a "slump" in Europe, the assumptions made with regard to saving and asset accumulation behaviors of private agents (individuals and firms) in the former do not seem appropriate to render the observed differences between the US and the Euro zone. Two additional channels of divergence may be worth exploring, and if possible combining, namely those arising from the existence of financial markets and the accumulation of private debts, on the one hand, and those arising from different demographic trends, on the other.

2.1 Asset accumulation and debts in the closed economy

As is well known, the life-cycle hypothesis of consumption and saving (Modigliani 1986) suggests that individual decisions to consume or save out current income depend on the life-cycle profile of earnings, the objective of the rational individual being to smooth her consumption over the life-cycle. In contemporary, developed economies, where earnings have a very distinct, rather hump-shaped, life-cycle profile, in part due to age-related wages during the years of labor-market participation – whether this upward trend is due to growing productivity as a result of accumulated experience and human capital, or whether it simply reflects dominant conventions such as seniority compensation rules is not important for the purpose of this reasoning – and in part as a result of retirement, implying zero or low earnings passed a certain age, a major implication of the life-cycle hypothesis is the necessity for individuals to save and accumulate wealth while active in the labor market in order to finance old-age consumption and hence to dissave by selling accumulated assets to younger

individuals, as well as, possibly, the desire of younger individuals to somewhat anticipate future earnings and consume more than current income by incurring debt.

Unlike in financially underdeveloped economies, where wealth has to be accumulated in the form of real assets – land, precious metals, possibly machinery, provided it does not depreciate too fast – and where borrowing is hardly possible, contemporary economies have, with the passage of time, progressively equipped themselves with more and more sophisticated devices for transferring wealth and resources through time, either anticipating their use by borrowing, or postponing it by accumulating them and investing, not so much in real assets as in securities, be they debt or shares, i.e. property rights on future production and profits. Of course, these vast opportunities to lend and invest or, on the other hand, to borrow resources, have been a major engine of economic growth over centuries, and especially in recent decades, when financial innovations, deregulation and globalization have boosted the development of financial markets and made financial transactions easier and more pervasive than ever before probably. But they have also altered the nature and mechanics of intergenerational relations and transactions: in order for older individuals to finance their current consumption out of their accumulated stock of – mostly financial – assets, they have to earn sufficient real (consumer-goods) income or to sell them for consumer goods to the younger, active generations; hence, the relative well-being of the successive generations essentially depends on the real – goods – yields and prices of assets.[10]

In parallel, borrowing goes on, either directly from older individuals, who then accumulate – private or public – bonds, or indirectly, through banks and other financial intermediaries, or via non-financial firms, who borrow to invest in real, productive assets, or in shares of other firms. For public, infinitely-lived borrowers, the well-known debt-sustainability constraint dwells on the "critical" difference between the nominal, apparent, interest rate on the stock of accumulated debt and the nominal growth rate of the economy: if the gap is negative, any initial stock of debt is easily sustainable; in the opposite case, any non-zero debt is unsustainable, due to the "snowballing" effect, and indebted agents have to have primary surpluses in order to stabilize, or reduce, their debt to GDP ratio. For private borrowers, the solvency conditions are, in theory, just as simple, but in practice much less so, essentially because their time horizon is likely to be finite, hence difficult to predict, and because their future incomes are also less predictable and have to be anticipated on the basis of current information. Due to these informational difficulties and asymmetries concerning debtors, lenders, be they banks or market investors, have developed a battery of indicators, based on accounting conventions, with ratings and thresholds, for financial analysis, as well as all sorts of derivative financial instruments to cover some of the individual risks to be incurred.

Why should investors want to hold debt instruments, and especially

public debt, in their portfolios, rather than exclusively accumulating their wealth in real assets or in property rights over productive capital, i.e. in common stock? Diversification in the face of risk, as well as various degrees of risk aversion among asset holders, is the most obvious explanation: just as firms in a non-Modigliani–Miller world will usually want to diversify their sources of external financing, therefore issuing not only stocks, but bonds and also taking bank credits, individual investors demand fixed-income securities, alongside shares and other property rights over real assets, which, historically, have had quite volatile values and yields.[11]

2.2 Ageing and international financial flows

In a financially globalized world economy, where labor is relatively immobile, and there are significant differences in the age profiles of the domestic populations in the various countries of the world, there is no reason why domestic saving should match domestic investment. Indeed, the degree of correlation between the two aggregates has been interpreted as an indicator of financial integration and international capital mobility (Feldstein and Horioka 1980; Obstfeld and Rogoff 1996). And even though empirical studies have usually found it to be relatively high (hence international financial integration to be incomplete), it is probably higher for transatlantic capital flows than in many other parts of the world so that capital mobility may be said to be high and increasing over recent years.

As a first approximation, it may be assumed that all domestic populations of the world go through the same process of ageing, generating the typical evolutions that demographers call the "demographic transition": with economic development and progress in welfare, nutrition, hygiene and the like, mortality declines and there is a progressive lengthening of life; later, probably as adults realize their children on average survive longer, so they do not need to give birth to many, fertility in turn declines. This process of demographic ageing, which may be characterized simply by a gradual increase in the average age of the population, is actually dominated, at least in the most economically developed countries, by the progressive postponement of death and lengthening of old age in good health and physical condition. The major difference between countries, in this respect, stems from lags in economic development and hence from the fact that the richer countries started their transition earlier, so that there are lags in the domestic demographic evolutions. A simple way of characterizing these differences, from the point of view of saving and wealth accumulation, is to look at the ratio of "high-savers" – i.e. according to the life-cycle hypothesis, individuals aged between about 45 and 65, corresponding, in economies were earning profiles have the classical hump shape, where children are born to couples aged between 25 and 35, and

labor market activity is, on average, between 20 and 65 – in total population, which gives an indication of the aggregate volume of saving in the domestic economy.[12]

From a worldwide perspective, the succession of peaks in the ratio of high savers to total population in the various countries shows Japan ageing first, already at the end of the twentieth century, closely followed by the EU, and indeed Europe at large; then comes North America (US and Canada),[13] then Russia, then China, and, progressively all other parts of the world, ending with sub-Saharan Africa. In the absence of major changes in institutions or other, national savings should therefore successively peak in one country after the other.

Focusing on Western industrialized countries, the differences in demographics, although less marked than between developed and less developed regions, are quite striking too (Figure 6.1). In particular, comparing the time profiles of the Western European and Northern American demographic structures with a view at their consequences on aggregate domestic savings reveals significant differences: looking at our conventional synthetic indicator of the influence of demographics on aggregate savings – i.e. the proportion of individuals between 45 and 65 years old in total population – clearly shows a time lag, as well as a substantial discrepancy in levels between the two regions, which may in part explain the lower US savings rate over the last two decades at least.[14]

On the other hand, there is little or no reason why domestic investment in productive capital should be so closely related to demographic profiles, except to the extent that discrepancies in the evolution of the working-age population, combined with similar per-worker capital stock, would lead to more investment in the demographically more dynamic region. Not withstanding this effect, peaks in national savings should be reflected in parallel evolutions of current account surpluses, hence, at least initially, in capital exports to other parts of the world, unless domestic investment rates differ for other reasons, such as the occurrence of a technological

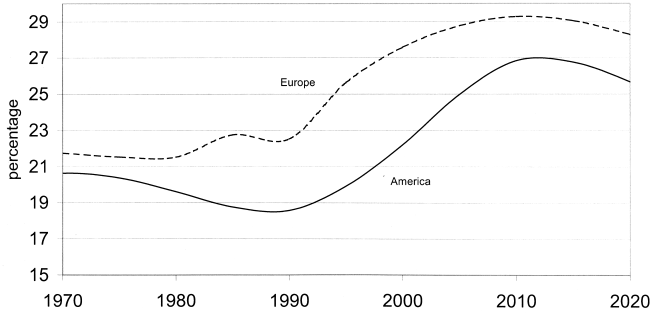

Figure 6.1 High savers (45 to 65) as a percentage of total population (North America and Western Europe).

shock in one region. In all such circumstances, the current account of the region with a lower domestic saving rate and/or a higher investment rate would be in deficit and capital can be predicted to flow from the high-saving/low-investment regions to the low-saving/high-investment one. Indeed it may even be argued that financial globalization is, at least to some extent, an answer to ageing in the developed economies.[15] But equally important is to recall that the external current account is the change in the net external foreign asset holding of the considered country: in the contemporary context of financially sophisticated economies and global finance, the underlying gross flows are usually very large and diverse, so that their valuation process may have significant effects, both on the current account balance itself and on macroeconomic developments.

3 Debt and capital accumulation and market valuation

In sharp contrast with the first decades of the after-war period, the 1980s and 1990s have witnessed the accumulation of ever larger stocks of debts, an increasing share of which in the form of marketable debt instruments, be they sovereign debts in the context of international financial markets, or public debts, or also private debts, issued by firms directly on financial markets or issued by banks[16] who have granted loans to firms or households. The possibility of borrowing and accumulating debt is related to the perceived creditworthiness of each agent or category of agents, itself resulting from conventional decision procedures and acceptance of collaterals, usually valued at market prices. Hence, in contemporary economies, unlike in simple macroeconomic models with representative agents, productive assets are not directly valued on asset markets: financial markets actually price debt instruments by passing judgment on the likely future repayment by the debtor, and price shares, that are ownership titles on complex bundles of productive (physical, organizational and human) assets and debts.[17]

Because outstanding stocks of debt and shares are held in the same portfolios, it is likely that their pricing mechanisms are closely related and that price changes on one or the other will have induced effects on all financial markets. The high development of financial instruments and markets, while permitting easier and cheaper diversification of investors' portfolios, have also generated greater interdependencies in asset price variations, in part through the mechanism of expectation formation, mimetic behavior and contagion,[18] and in part through the linkages created by valuation conventions and ratios.

Thus, in contemporary economies, both the financial wealth of individual investors (mostly older households) and the financing possibilities available to firms deciding to accumulate productive assets, no matter what their exact nature is, are highly dependent on market valuation

processes, that heavily rely on expectations of future returns and perceived risks, as well as those balance-sheet ratios that consensus financial analysis deem appropriate. And, as shown in the recent episodes of sharp stock market downward adjustments and financial scandals on both continents, these conventions may not rest on very solid analytical or informational grounds, and may change along with the dominant mood of the markets, a statement that already was an essential ingredient in Keynes' analysis of business cycles. And the circular and self-validating nature of the valuation processes is quite pervasive in a world where almost all balance-sheet items of firms are "marked to market".

4 International financial transmission

As in the original Fitoussi and Phelps' (1988) analysis, the international transmission mechanisms are seen as a major ingredient in the determination of macroeconomic performance on the two sides of the Atlantic. But the channels emphasized here are different, partly because the magnitude of dollar–euro exchange rate changes, although significant, are dwarfed by the huge exchange-rate fluctuation of the 1970s and 1980s, and essentially because, in a world of global finance, the channels of international transmission of macroeconomic shocks and fluctuations are likely to be dominated by financial asset price changes, hence by the interplay of debt and asset valuation on financial markets.

Two major evolutions seem to characterize the recent decades and deserve special attention because they apparently result from similar causes and have similar macroeconomic consequences, from an international transmission perspective. One is the already noted magnitude of gross international financial flows, as well as the very large private debt issues that are in part supporting the flows of financial assets, in particular foreign direct investment flows, still larger between the US and Europe than between developed and emerging countries in spite of the spectacular expansion of the latter. These large financial flows and the resultant very large stocks of financial assets and debts accumulated in balance sheets of banks and financial institutions, firms and individuals in both regions, generate strong linkages that are likely to influence economic decisions, in particular saving and wealth accumulation, as well as portfolio allocation choices, on the part of households, and capital accumulation and financing decisions on the part of firms.

The other outstanding characteristic of the past decade or so is the almost perfect correlations between financial asset prices and returns on both sides of the Atlantic. Whatever the exact explanation for it, this correlation has been amply documented with regard to stock prices, especially the downward adjustment phase from March 2000 till early 2003, when indexes of stock prices in the US stock exchange in European markets have been divided by approximately two. But a similar, if not

Figure 6.2 Long-term nominal interest rates (US and Germany).

even closer correlation may be observed for nominal, long-term interest rates, as shown in Figure 6.2.

Contrary to what constitutes a central ingredient in most international macroeconomics, this correlation is not compatible with the standard assumption of perfect capital mobility, which is usually seen to imply an arbitrage relation on expected real rates of return, expressed in the same monetary unit, unless one is willing to assume that inflation expectations and exchange-rate changes expectations are systematically mistaken.[19] Indeed, whereas in the canonical models of international capital flows, there is some kind of world capital market where the supply and demand are confronted, determining the world real interest rate, the empirical evidence suggests that long-term real rates do differ, insofar as inflation rates are different while nominal rates have, over the last decade or so, almost always been exactly equal.

5 Macroeconomic policies and debt dynamics

When macroeconomic policies are brought into the picture, an additional striking difference between the EU and the US over the past two decades lies in the contrasted evolutions of macroeconomic policies and indicators. First, judging from its main instrument, i.e. the short-term interest rate, monetary policy has been used in a vigorous countercyclical manner in the US, at least since the beginning of the 1980s, while central banks in Europe, including the European Central Bank in the more recent years of the euro, have tended to adopt a more conservative stance, with both a more restrictive overall orientation, and much less reactivity to cyclical fluctuations, especially in the face of slowdowns or recessions in economic activity (Figure 6.3).

The management of fiscal policies has also been marked different, with the US federal governments letting budget deficits reach colossal amounts

Figure 6.3 Short-term nominal interest rates (US and Germany).

during the severe recessions of 1982–1983 and 2001–2002, but also piling up federal budget surpluses in the last years of the long-lasting boom of the 1990s, whereas national budgets in the EU have mostly been passive, at most letting automatic stabilizers operate; but on average, budget deficits have been larger, as a share of GDP, in European countries than in the US.

The evolutions of public indebtedness, measured by the general government gross debt to GDP ratio, have been completely different though. Public debts were small everywhere at the beginning of the 1980s, after several years of high inflation and sometimes negative real interest rates that had eroded outstanding stocks and prevented budget deficits from resulting in the accumulation of new debt; the disinflation phase of the 1980s, with high nominal interest rates and low growth, boosted the debt-to-GDP ratio on both sides of the Atlantic, reaching about the same level at the end of the decade. But in the following decade, the indebtedness ratios diverged, substantially declining in the US, while climbing to levels almost unprecedented in peacetime in Europe,[20] only to recede moderately towards the end of the century (Figure 6.4). In more recent years, the upward trend has resumed in the US, due to very active countercyclical fiscal policies, as well as in large European countries, but here merely as a result of automatic stabilizers and the induced effects of slow growth on public indebtedness ratios.

Why have public debts evolved so differently? Not really because of fiscal policies, but rather because of differences in the "critical gaps". One way of telling the story of the last 15 years in the US and Europe is to stress that monetary policy in the former has almost constantly kept nominal short-term rates at such levels that the corresponding critical gap was negative or small at least since the 1991 recession, while monetary policies in Europe, either before or after monetary unification, have constantly maintained the critical gap at positive, and on occasion quite

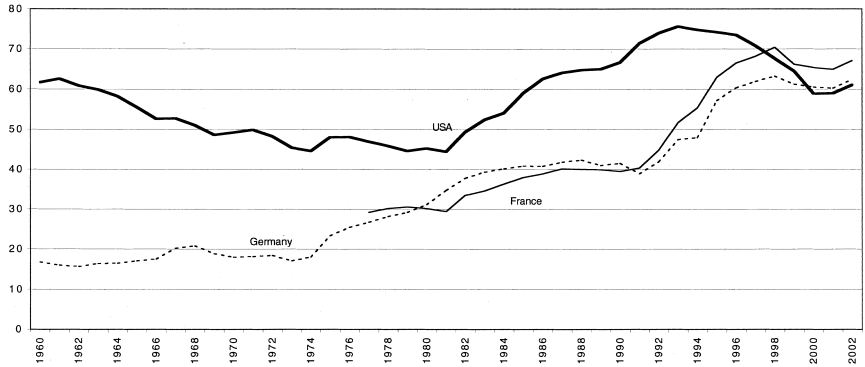

Figure 6.4 Gross public indebtedness (percentage GDP) (source: OECD).

high, values (Figure 6.5). Whatever the underlying reason for such policy orientations, this observation, combined with the outcome of the international transmission mechanism described above on the critical gap on long-term rates (Figure 6.6), implies that, for private and public debtors, sustainability was not a major issue in the US, while it was central in Europe.

The implications, both for private decisions to save and invest and for public decisions regarding fiscal policies, are quite far-reaching. First, contrary to the situation in the US, in each slow-down or recession, European debtors have been confronted with a "debt deflation"[21] problem, having to adjust spending and saving decisions to avoid bankruptcies or unsustainable debt accumulations that, anyway, would eventually be sanctioned by financial markets. Moreover, fiscal policies have been dominated by an obsession to curb the evolution of public indebtedness, which inevitably generates pro-cyclical orientations, and may lead, in bad times, to even block the working of automatic fiscal stabilizers.[22] In addition, insofar as

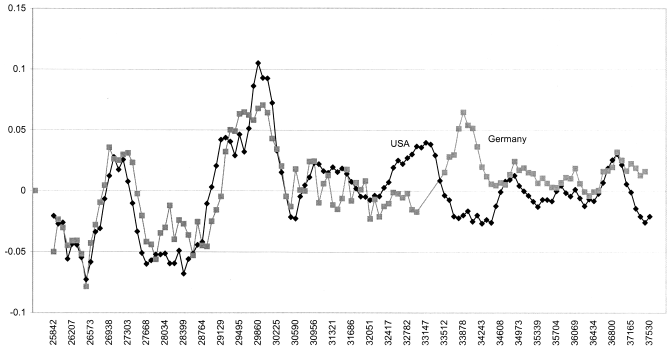

Figure 6.5 Critical gap on short-term interest rates (US and Germany).

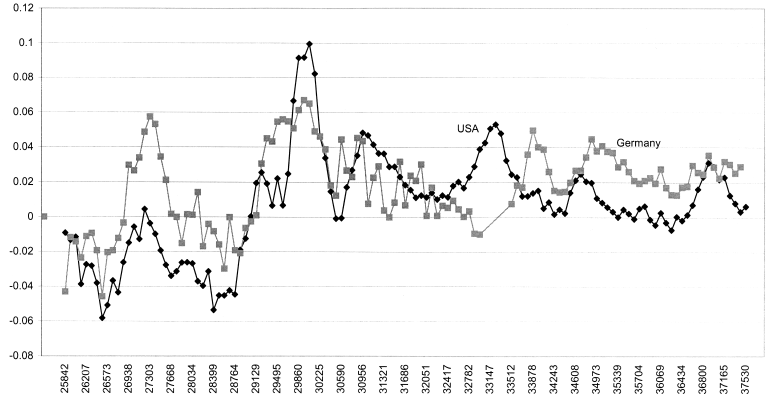

Figure 6.6 Critical gap on long-term nominal interest rates (US and Germany).

Europe is a net creditor of the US, and because in recent years European firms have tended to heavily borrow in Europe to invest in the US, the "debt deflation" mechanism and its induced effects are amplified by the asset price movements of the past three years, that have made the asset side of their balance sheets less valuable, while the liabilities were mostly unchanged, or even rising.

5 Concluding remarks: "Old Europe" as a rentier economy

Explaining the "slump in Europe" would thus seem to require more than seldom denounced labor-market rigidities and lack of structural reforms thereof. In a way, it may even be argued that wage rigidities, employment protection and relatively generous welfare provisions in most European countries are necessary features of economies in which macroeconomic and employment performances are relatively poor, while the US combines more microeconomic flexibility with some sort of macroeconomic insurance mechanisms acting at the aggregate level through the macroeconomic policy mix, which is set in such a way as to maintain sustained growth and high employment levels.

The reasoning proposed in this chapter takes inspiration from the original analysis of Fitoussi and Phelps (1988), by combining the influence of macroeconomic policies and a characterization of international transmission of aggregate fluctuations dominated by financial market outcomes. It however differs in many ways from their analysis, especially through the introduction of debts, as well by taking some "structural" elements into account, in particular the differences in demographics. With the latter aspect in mind and also in reference to the situation observed in many European countries (especially France) in the last decades of the nineteenth century, when the ageing populations were heavily investing

abroad, while many European economies were going through a long "depression" phase, when, under the gold standard, monetary stability was dominating the scene, the general orientation of macroeconomic policies may be seen as corresponding to these differences in demographic conditions. Indeed, monetary policy in the US, with its focus on financial stability and a relatively moderate price stability objective, takes debtors' interest into account and seems to respond to their need not to be forced into debt deflation, whereas the European insistence on strict monetary stability, interpreted narrowly as to mean low inflation, does seem to serve the interest of private holders of debt instruments, i.e. "rentiers", even at the cost of slow growth and high unemployment.

Notes

1 Thanks to Jérôme Creel and Eloi Laurent, for precious help and advice. And to Jean-Paul Fitoussi for inspiration. None of them should, of course, be held responsible for possible errors in this chapter.

2 This is, of course, the main line of the Fitoussi and Phelps' (1988) book. The divergences in macroeconomic policy mixes and their consequences of interest rates and exchange rates have been further documented and analyzed by Fitoussi and Le Cacheux (1988, 1989, 1994).

3 In most of the chapter, the focus will be on large EU member countries, mostly Germany and France, which have clearly had the worst macroeconomic performance over the past two decades. Many smaller European countries have outperformed them and taking them into account would somewhat lessen the contrast between the US and Europe emphasized here.

4 In Creel and Le Cacheux (2003), we show that asymmetries are still quite important in the Euro zone, four years after the completion of monetary unification. The resulting heterogeneity in macroeconomic conditions within the Euro zone makes the management of monetary policy quite problematic and may help explain the current difficulties of large European economies, especially Germany and France.

5 One recent review of many possible explanations for the relatively poor economic performance of Europe, compared to the US economy, is to be found in Gordon (2003).

6 This view is clearly dominant in the analyses and recommendations of the European Commission and of the OECD. See, for a relevant source of their common inspiration, Layard and Nickell (1999).

7 See, in particular, Fitoussi and Passet (2000) and Freeman (2000), and the discussion in Fitoussi, Jestaz, Phelps and Zoega (2001).

8 See the careful empirical work on the French tax system in Piketty (2002).

9 In his 1994 book on "structural slumps", Phelps presents three models, each one characterized by the existence of one asset, depending on the market structure assumed. His reasoning could easily be extended to take account of the accumulation of other kinds of productive assets, such as R&D for instance.

10 This is, of course, a major intuition of Samuelson's (1958) overlapping-generations model. The theoretical consequences for public debt in a growing economy have been analyzed by Diamond (1965). The implications for inflation and the real value of money have been spelled out by Tirole (1985).

11 For a vivid and insightful account of the history of financial markets' instability, see Kindleberger (1996).

11 For a vivid and insightful account of the history of financial markets' instability, see Kindleberger (1996).
12 This reasoning is spelled out in more details, and the magnitudes precisely calibrated to fit actual and projected demographics, in INGENUE Team (2002). This sub-section mostly summarizes the major conclusions of the various simulations carried out with the first version of the INGENUE, world model, itself an applied, computable version of the Obstfeld and Rogoff (1996) theoretical models. It should be stressed that our characterization of high savers holds under the assumption of unchanged institutions, especially social security, and perfect foresight. In case of uncertainty, especially regarding the future of public pensions, it is likely that "high savers" will save more, but it may also be that younger individuals will have to save more or borrow less.
13 The baby-boom/baby-bust fertility profiles that characterize most developed economies, with varying timing and magnitudes, tend to amplify the peaks of the demographic evolutions in these regions of the world, the others going through relatively smoother transitions, except for wars and epidemics, in particular AIDS in Africa. See INGENUE Team (2002).
14 This reasoning is of course simplistic, in many respects, and rests on a strong *ceteris paribus* assumption. In particular, it neglects the differences in financial structures and changes thereof. Most conspicuous in the present context is the relative ease with which households, but also to some extent firms, may borrow, which obviously has a bearing on aggregate private savings.
15 See also Le Cacheux (2002). On the historical parallel with the previous period of globalization, at the end of the nineteenth century up to the First World War, or even the late 1920s, see Kindleberger (1996); Bouët and Le Cacheux (1998); Flandreau *et al.* (1998); O'Rourke and Williamson (1999).
16 For the sake of simplicity, we will ignore the likely autonomous decisions of the latter, which would lead to an additional source of macroeconomic fluctuations and international transmission, mostly reinforcing those emphasized here through channels often referred to in recent years as "credit crunch". See, for instance, Bernanke (1995), but also, for a historical perspective, Kindleberger (1996). One would also have to take account of the ways in which financial and prudential ratios amplify these movements.
17 And indeed, with the existence of derivatives markets, they also price individual components of the various risks attached to these financial instruments.
18 On these mechanisms, which are clearly at the heart of Keynes' (1936) analysis, see also the extended and very enlightening analysis developed in Aglietta and Orléan (2003).
19 Given the already noted relatively small exchange rate fluctuations in the 1990s and early 2000s, it may be that investors actually find them negligible. It may also be the case that the possibility offered by derivatives, to insure against exchange risk, thus effectively isolating it from the other characteristics of the marketable debt instruments, generates arbitrage behaviors that mimic those in a single currency, equating nominal rates, an observation that would seem to have been relatively common under the gold standard in the late nineteenth century. See Flandreau *et al.* (1998).
20 As a matter of fact, public debt ratios had already reached similar, or even higher levels in some European countries at the end of the nineteenth century. See Flandreau *et al.* (1998).
21 This is, of course, the well-known mechanism described and analyzed by Irving Fisher (1933) in the context of the Great Depression.
22 Observers of the European scene will have recognized the well-known debate

around fiscal rules in the Euro zone, and in particular the so-called "Stability and Growth Pact". For a more detailed analysis of fiscal policies in Europe, see, for instance, Fitoussi and Le Cacheux (2003).

References

Aglietta, Michel and André Orléan (2003) *La monnaie, entre violence et confiance*, Paris, Odile Jacob.

Bernanke, Ben (1995) "The Macroeconomics of Great Depression: A Comparative Approach", *Journal of Money, Credit, and Banking* no. 1, February.

Bouët, Antoine and Jacques Le Cacheux (eds) (1998) *Globalisation et politiques économiques: les marges de manœuvre*, Paris, Economica.

Creel, Jérôme and Jacques Le Cacheux (2003) "Inflation Divergence and Public Deficits in a Monetary Union", *OFCE Working Paper* no. 5, July.

Diamond, Peter (1965) "National Debt in a Neoclassical Growth Model", *American Economic Review* vol. 55, no. 5, September.

Feldstein, Martin and C. Horioka (1980) "Domestic Saving and International Capital Flows", *Economic Journal* vol. 90, no. 2.

Fisher, Irving (1933) "Debt Deflation Theory of Great Depressions", *Econometrica* vol. 1.

Fitoussi, Jean-Paul, David Jestaz, Edmund S. Phelps and Gylfi Zoega (2001) "Roots of Recent Recoveries: Labor Reforms, Sound Public Finance or Private Sector Forces?", *Brookings Papers on Economic Activity* no. 1.

Fitoussi, Jean-Paul and Jacques Le Cacheux (1988) "Ruptures et continuité dans les politiques macroéconomiques des grands pays de l'OCDE", *Revue de l'OFCE* no. 22, January.

Fitoussi, Jean-Paul and Jacques Le Cacheux (1989) "Une théorie des années quatre-vingt", *Revue de l'OFCE* no. 29, October.

Fitoussi, Jean-Paul and Jacques Le Cacheux (1994) "Des perspectives macroéconomiques sous la tutelle des créanciers", in *La France au-delà du siècle*, Paris, DATAR/Editions de l'Aube.

Fitoussi, Jean-Paul and Jacques Le Cacheux (eds) (2003) *Rapport sur l'état de l'Union européenne 2004*, Paris, Fayard and Presses de Sciences Po.

Fitoussi, Jean-Paul and Olivier Passet (2000) "Réduction du chômage: les réussites en Europe", *Rapports du CAE* no. 23, Paris, La Documentation française.

Fitoussi, Jean-Paul and Edmund S. Phelps (1988) *The Slump in Europe*, Oxford, Oxford University Press.

Flandreau, Marc, Jacques Le Cacheux and Frédéric Zumer (1998) "Stability Without a Pact? Lessons from the European Gold Standard, 1880–1914", *Economic Policy* no. 26.

Freeman, Richard B. (2000) "Single Peaked vs. Diversified Capitalism: The Relation between Economic Institutions and Outcomes", *NBER Working Papers* no. 7556.

Gordon, Robert J. (2003) "Deux siècles de croissance économique: l'Europe à la poursuite des Etats-Unis", *Revue de l'OFCE* no. 84, January.

INGENUE Team (2002) "INGENUE: A Multi-Regional, Computable General-Equilibrium, Overlapping-Generations, World Model", in J.-O. Hairault and H. Kempf (eds), *Market Imperfections and Macroeconomic Dynamics*, Boston and Dordrecht, Kluwer Academic Publishers.

Keynes, John Maynard (1936) *The General Theory of Employment, Interest, and Money*, London, Macmillan.

Kindleberger, Charles P. (1996) *Manias, Panics and Crashes, A History of Financial Crises*, Third edition, London, Macmillan.

Layard, Richard and Steve Nickell (1999) "Labor Market Institutions and Economic Performance", in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics* vol. 3C, Amsterdam, North Holland.

Le Cacheux, Jacques (2002) "Mondialisation économique et globalisation financière: de quelques poncifs, idées fausses et vérités", *Revue de l'OFCE*, Special issue, "La mondialisation et l'Europe".

Modigliani, Franco (1986) "The Life-Cycle Hypothesis of Saving. Nobel Prize Lecture", *American Economic Review* vol. 76.

Obstfeld, Maurice and Kenneth Rogoff (1996) *Foundations of International Macroeconomics*, Cambridge, MA, MIT Press.

O'Rourke, K.H. and John G. Williamson (1999) *Globalization and History*, Cambridge, MA, MIT Press.

Phelps, Edmund S. (1994) *Structural Slumps*, *The Modern Equilibrium Theory of Unemployment, Interest and Assets*, Cambridge, MA, Harvard University Press.

Piketty, T. (2002) "Fiscalité et redistribution sociale dans la France du XXe siècle", in *L'impôt en France aux XIXe and XXe siècles, Comité pour l'histoire économique et financière de la France*, Paris.

Prescott, Edward C. (2002) "Prosperity and Depression", Richard T. Ely Lecture, *American Economic Association, Papers and Proceedings* vol. 92, no. 2, May.

Samuelson, Paul A. (1958) "An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money", *Journal of Political Economy* vol. 66, no. 3, July.

Sargent, Thomas S. and Neil Wallace (1981) "Some Unpleasant Monetarist Arithmetic", *Federal Bank Reserve of Minneapolis Quarterly Review*.

Tirole, Jean (1985) "Asset Bubbles and Overlapping Generations", *Econometrica* vol. 53, no. 6, December.

7 The long swings in economic understanding

Axel Leijonhufvud

Changing worldviews

The worldview of a representative macroeconomist has changed repeatedly and drastically over the century from which we have just escaped. With the "new classics", it may indeed seem to have come full circle. But that, on closer inspection, turns out to be largely an illusion. True enough, economists a hundred years ago generally believed that the economy always tended towards equilibrium. Today, "tendencies" do not enter into it one way or another. The modern macroeconomist believes the economy *is* in equilibrium. True enough, the main policy objective of economists a hundred years ago was to stabilize the price level and today low inflation has become the overriding goal of macropolicy. But in seeking to understand how the price level could be controlled, Wicksell and Fisher were motivated by a passionate concern with the distributive injustices arising from inflations and deflations. Distributive justice figures nowhere in the leading school of modern macroeconomics[1] and certainly not in the muddled "shoeleather" arguments marshalled in support of this exclusive concern with eliminating inflation.

If we shorten the time perspective to 50 or 40 years, the contrasts are more drastic. The Great Depression had shattered the representative economist's faith in the self-adjusting capabilities of the market system. At the height of the influence of Keynesianism, he – it was a he – believed that the private sector was unstable, riddled with market failures, and prone to fluctuations amplified by multiplier and accelerator effects. But he also believed that a benevolent and competent government, dependable in its democratic role as the agent of the electorate, could stabilize the economy and ameliorate most market failures.

Today, he – or she – believes the economy keeps to an intertemporal equilibrium path. As long as not disturbed by bad policies, markets will take care of all coordination problems. Trouble may arise in this otherwise idyllic picture on two fronts. First, if "inflexibilities" in the labor market are tolerated, the natural rate of unemployment may be quite high. More importantly, governments are seen to be prone to excessive deficits and

inflationary finance, to be constitutionally time-inconsistent and to be addicted to playing the Phillips curve.

To the older generation, macropolicy was *stabilization* policy. To the present generation, it has become the art of *constraining* governments. If the older was uncritical in trusting the power of government to do good, the present one has a distinctly cynical view of representative democracy. Jean-Paul Fitoussi's work has run counter to this trend of the times. During the 1990s, it was not easy to get a fair hearing for the view that the high interest rates with which the Bank of France defended the franc exchange rate against the DM bore a lot of responsibility for France's high rates of unemployment.[2] Indeed, even being allowed to speak was a bit of a problem.[3] And his concern with the worsening income distribution and its effect, in turn, on the acceptability of the *contrat social* raised hardly an echo in an American-dominated economics profession.

There are many strands to the story how the worldview of the economic profession came to change so radically. The strand that runs through the history of macroeconomics may be the most important of them all. Judgements on the successive episodes of that history conflict. What follows is one man's opinion.

The *General Theory* as a response to the depression

The response to the Great Depression among economists and policy makers was confused and incoherent. The challenge was met most effectively by Keynes. The *General Theory* brought a great advance in our understanding of depressions and particularly of the impediments to recovery from such extreme states. The younger generation of economists in particular received it eagerly and, while its positive contribution was not altogether clearly understood, its legacy came to dominate macroeconomics for 30 years. But the work had weaknesses as well as strengths. These flaws in the work, combined with common misunderstandings of it, eventually undermined its influence.

The core of Keynes's theoretical contribution I would (today) describe as follows.[4] The point of departure is the Marshallian tradition. Marshall's theory assumed adaptive, rather than optimizing behavior. At the cost, perhaps, of some oversimplification, consider a universe of agents obeying very simple "laws of motion". For example, consumers will increase their consumption of a good whenever their demand price exceeds the market price. Producers will increase output whenever market price exceeds their supply price. Middlemen will increase prices when faced with excess demand, etc.

Before Keynes, it was a firmly and generally held opinion that if all agents obeyed these laws of motion, the dynamics of the economy would surely carry it to full employment. Only violations of the laws of motion – such as rigid wages or other "inflexibilities" in the labor market – could

stand in the way of this result. One notes that this is the understanding to which economics has once again come around in recent years. Keynes, however, discovered that this proposition was not true. The monetary and financial institutions of a modern economy are such that saving is not an effective demand for future consumption nor is the offer of labor an effective demand for present consumption. Some of the "market forces" needed to guide the collective adaptation process may be missing therefore. A theory that recognizes the possibility of such effective demand failures has a valid claim to being a more *general theory*. This, I believe, is Keynes's main claim to lasting fame as a theorist.

The core of the theory was the saving-investment nexus, which is to say, the problem of intertemporal coordination. Years before the *General Theory* was completed, he had reached a diagnosis of the depression that he held with great conviction. The decline in fixed investment was the cause "without any doubts or reserves whatsoever – the whole of the explanation of the present state of affairs" (Keynes 1931: 351). The trouble was that the interest rate did not bring saving and investment into line: "In the past it has been usual to believe that there was some preordained harmony by which saving and investment were necessarily equal . . . But unfortunately this is not so, I venture to say with certainty that it is not so" (Ibid.: 355). There was "a fatal flaw . . . in orthodox reasoning . . . largely due to the failure of the classical doctrine to develop a satisfactory theory of the rate of interest" (Keynes 1934: 489). (Not a word, of course, about rigid wages being even part of the problem.)

The liquidity preference theory of interest with which Keynes sought to replace that classical doctrine was however worse than unsatisfactory. He had somehow convinced himself that if saving and investment determined income (and not interest), then liquidity preference and money supply must determine interest (and not money income or the price level). This meant, among other things, that it was *not possible* that the interest rate would adjust so as to coordinate consumption and production over time. A far stronger assertion than the proposition that it *cannot be depended on* always to do the job! This doctrine became a rich source of confusion.[5] From it stems the Paradox of Thrift, for example, and also the proposition that "investment causes saving but saving does not cause investment" to which Keynes's Cambridge followers have shown a fierce attachment but which fails to throw much light on postwar problems of capital formation around the world.

In the early years of the depression, Keynes had expressed himself forcefully on the dangers of debt-deflation. An "immense burden of bonded debt, both national and international . . . fixed in terms of money" had been built up (Keynes 1930b: 128). This entire structure would be "deranged" with disastrous consequences (1931). It is surprising, therefore, that this theme is hardly mentioned in the *General Theory* – despite the evident fact that this is exactly what had happened in the United States

in the meantime. But Keynes had apparently convinced himself that the contemporary capitalist economy had a chronic tendency to save more than its capital markets would channel into investment. Now, the saving that does not end up financing investment will be directed towards improving liquidity positions. But there is a crucial qualitative difference between the temporary tidying up of balance sheets in an ordinary recession and the desperate scramble for liquidity in the wake of a financial crash. The *General Theory* does not make the distinction. In the midst of the great depression with the capitalist economies in dire straits, while the totalitarian threats were mounting on both left and right, it was surely exceedingly difficult to maintain a balanced perspective on the "fatal flaw" in the economic system that Keynes thought he had identified. But the result was that he generalized from the depression and consequently came to exaggerate the economy's instability, specifically with regard to the frequency with which to expect strong deviation-amplifying effects through the consumption multiplier.[6]

To Keynes, it was always the saving of households that was the problem. He rather slighted the liquidity management of firms and its role in propagating negative (or positive) impulses.[7] In a simple model with only outside money, the excess of saving over investment constitutes a straightforward excess demand for money. More generally, in recessions, firms which temporarily are not finding profitable investment opportunities will seek to improve their balance sheets by repaying bank loans while the banks, in turn, respond to this "reflux" by returning their borrowed reserves to the Central Bank. The inside money stock thus covaries with the output. This mechanism was explained very clearly in the *Treatise* but is missing in the *General Theory*. It might have been helpful to the Keynesian side in the monetarist controversy if it had been an integral part of their canon.

The focus on households directs attention to the inability of the unemployed to exercise effective demand for consumption goods. Most unemployed workers in interwar Britain were surely income-constrained to an extent that, in large part due to the influence of Keynes, has become uncommon in more recent decades. But it also directs attention away from another effective demand failure which, again, pertains to the balance sheets of firms in a depression. In an intertemporal general equilibrium model, it is assumed that firms can exchange the promise of revenues from future output for the present command over the resources required to realize that output. If this exchange cannot be effectuated, very little growth can be realized. Yet, this is what happens in the wake of a financial crash that leaves large volumes of bad loans outstanding between banks and businesses. Policies that do not tackle this specific effective demand failure directly are unlikely to help, as the experience of Japan over the last decade demonstrates. Conventional Keynesian analysis will suggest public works and low interest rates as the measures to be tried but the

Japanese problem is neither income-constraints nor a "liquidity trap" in the usual sense.

Keynesianism in the United States

The Liquidity Preference hypothesis of interest determination is, I have contended, an error which put Cambridge Keynesians onto the wrong track. But it had virtually no influence on the development of Keynesianism in the United States and played no role in its eventual failure. Liquidity Preference vs. Loanable Funds was a contest between two hypotheses about what drives the economy's adaptation to shocks and it proved more or less impossible to make sense of the issue between them with the help of the simultaneous equation approaches that became standard after the war. Indeed, the American Keynesians were constantly surprised at the evident indignation vented at them by their Cambridge counterparts, not comprehending that growth models in which "saving *did* cause investment" were an outrage to those who counted themselves true believers.

The theoretical troubles of American Keynesianism stem instead from the so-called "Neoclassical synthesis", so named and long insisted upon by Paul Samuelson whose unequalled prestige did much to gain it general acceptance. The "synthesis" maintained that Keynesian theory was a special case of the more general neoclassical theory, obtained by adding the restriction of "rigid wages" to the latter. As an interpretation of Keynes, this was calumny but, given the scholarly standards of the profession as they are today, the sticky wages calumny sticks and cannot be revoked no matter how often it is shown to be false. In any case, for a couple of decades most everyone was pleased to accept the implied contention that Keynesianism was wrong in pure theory but right in policy practice.

However, in switching back to the pre-Keynesian preoccupation with sticky wages, American Keynesianism gradually lost sight of the intertemporal coordination problem. Keynes's distrust of the financial markets was only kept alive by a small band of post-Keynesians while the major American Keynesians made great contributions to the development of modern finance theory (which has been anything but distrustful of them). It is instructive to recall how Keynes's idea that interest rates, rather than wages, might be too high for full employment came to disappear from the theory that was made to bear his name. In Modigliani (1944), two different unemployment solutions to the simultaneous equation model were given. One, the basic one, where the money wage was fixed at too high a level relative to the money stock to allow full employment; the other, called "Keynes's special case", with the rate of interest too high for investment to equal saving at the full employment level of income. The latter, however, Modigliani thought of as a case of "rigid interest" which could apply only if the economy was in a true liquidity trap. A few years later, Patinkin

(1948) pointed out that, in principle, the Pigou effect could take the economy out of the trap and back to full employment and in subsequent revised versions of his famous paper, Modigliani (e.g., 1980) conceded the point. This left the basic case – unemployment is due to wages that are too high and its persistence is due to their downward rigidity.

Although the Pigou effect argument is flawed almost to the point of fraudulence, the liquidity trap was generally regarded as a theoretical curiosity and its elimination was mourned by nobody. Keynes would not have mourned it either. His concern had been with interest rates that, while not in the least "inflexible", were kept too high by central bank policy or by bearishness on the exchanges, thus setting in motion a decline in real output which in turn would bring with it a worsening of investment expectations. This could happen at *any* level of the interest rate, not just at some small fraction of 1 percent. But this his central hypothesis now disappeared as it were through a logical "trap" door.

This development allowed a great simplification of "Keynesian" economics which greatly helped its marketing throughout the world. For short-run purposes, all you now needed to know was (a) the volume of nominal expenditure and (b) the level of the inherited rigid money wage. A truncated version of IS-LM will serve as a theory of nominal income and, if just a historically fixed wage rate is too crude, it may be replaced by a Phillips curve.

It is worth taking note of what this simplification of textbook Keynesianism entails. A complete simultaneous equation representation of the "analytical nucleus" of the *General Theory*, as provided by Barens (1999), will have separate supply-price as well as demand-price functions for capital and for consumer goods. Shocks to investment expectations impinge on this system through the demand price for capital goods. This model has a relative price structure and reasonable "microfoundations". It could be modified so as to include the response of the supply price of labor to the volume of employment. But it would not allow the *ad hoc* addition of a Phillips curve (and even less, of course, of an aggregate supply function).[8] The Phillips curve came in handy, however, as an appendage to an IS-LM model from which relative prices had been removed.

IS-LM understood simplistically as a theory of nominal income and augmented by the Phillips curve, but deprived of the intertemporal coordination issue, constituted the poor, enfeebled, benighted "Keynesian economics" that succumbed to monetarism.

The natural rate

Milton Friedman's (1968) Presidential Address to the American Economic Association, in which he introduced the concept of the Natural Rate of Unemployment, is a milestone in the great shift in economists' worldview. Snowdon (1999: 32) has collected a few testimonies: "very likely the most

influential article ever published in an economics journal" (James Tobin); "the most influential article written in macroeconomics in the past two decades" (Robert Gordon); "one of the decisive intellectual achievements of postwar economics" (Paul Krugman); "easily the most influential paper on macroeconomics published in the postwar era" (both Mark Blaug and Robert Skidelsky).

The influence of that paper is undeniable. It misled just about everybody and diverted economic theory down a track that was to prove destructive also of Friedman's monetarism. It acquired immense authority by the fact that the inflationary decade of the 1970s appeared to verify all its predictions. But even so it could not have swept everything before it had not all the intellectual sloppiness surrounding the neoclassical synthesis enfeebled its potential opposition.

There were two things wrong with it. First, it assumed that all money could be analyzed as if it were outside money – and therefore neutral. Second, it presumed that as long as nominal wages adjusted to monetary changes, the economy would settle down to the natural rate of unemployment *whether or not* it were also at the natural rate of interest.[9]

Once anticipated, an outside money inflation would surely displace the Phillips locus vertically. If we were to suppose that the historical growth rate of outside money had been x percent higher than it actually was, it is plausible (abstracting from inflation tax effects) that the historically given scatter of points would have been x percent higher in Phillips space. Two things that do *not* follow are worth noting. First, it is not true that all central bank actions entail, or amount to the same thing as, changing the equilibrium nominal scale of the system.[10] Second, while correctly anticipated inflation means that the scatter will be displaced vertically, it does *not* by itself mean that the scatter will collapse to a single vertical "long run Phillips curve" as soon as wage-adjustments have caught up with changes in the outside money stock. If saving were to exceed investment[11] at NAIRU, employment will not converge to NAIRU.

Two stylized facts should be kept in mind. One is that while the U.S. short-run Phillips curve misbehaved in the 1970s and early 1980s as predicted by Friedman, we also have to explain why it behaved quite differently before the late 1960s or after the mid-1980s. The answer, I suggest, is that the 1970s was *the* decade of American outside money inflation. The other is that NAIRU has proved empirically hopeless to nail down. The thing is found to move as required to fit the new worldview. So unemployment under Reagan in the 1980s was blamed on the natural rate of unemployment having shifted up, not on the high real rates of interest. In the same manner, unemployment in France in the 1990s was widely blamed – Professor Fitoussi notably dissenting – on the inflexibilities of the French labor market, not on the high interest rates. In the same vein, the long Clinton boom was supposedly due to the great flexibility of the American labor market, not to low interest rates and loose liquidity.

Friedman's natural rate hypothesis was accepted by all and sundry. "Keynesians" who had opposed him on a range of issues did not themselves have an explanation for unemployment other than lagging wage-adjustment and, in order to make any case at all for stabilization policy, found themselves in the position of having to argue that money wage inflexibility was more of a problem than monetarists were willing to recognize. The so-called "New Keynesians" have subsequently set themselves the task of finding reasons for wage stickiness that are "rational" enough not to earn them the disapprobation of Lucas & Co.

Acceptance of the natural rate hypothesis left increased labor market "flexibility" as the only remaining remedy for unemployment. So this has become the one-note tune of macroeconomic debate in Europe. No doubt, many European countries have allowed institutional arrangements to develop that inhibit competition and mobility in a way that benefits some groups to the expense of others. A good argument can be made for some of these arrangements to be modified. The problem is that it is not clear how far the flexibility argument is going to be pushed. If "inflexibilities" of one sort or another really are the only reasons for unemployment, are we going to demand of labor that, whenever rates of interest are too high or animal spirits too weak, workers take enough of a cut to encourage business to provide employment? The idea that labor should be obliged to subsidize business whenever intertemporal coordination fails seems fraught with moral hazard.

Economic ideas have political consequences. The natural rate hypothesis may be logically weak and empirically useless but it has become an ideological pillar of the new right. And its acceptance on the other side of the political spectrum explains much of the evident intellectual incoherence on the left. Stabilization policy used to be something tangibly useful that government could do for lower income citizens. But if you accept that unemployment is the fault of labor and that income distribution has to be whatever a globalized economy may dictate, it becomes difficult to formulate a message that will carry conviction for the traditional voters on the left.

In Friedman's famous address, the natural rate doctrine is combined with the hypothesis that monetary policy does not affect any real magnitudes once prices have been appropriately adjusted. In Friedman's own theory, the equilibration of prices would occur with long and variable lags. In the interim, monetary policy could have significant real effects. In Lucas's rational expectations version of monetarism, the long and variable lags evaporate leaving only a soon-to-be-resolved momentary confusion resulting from policy "surprises". The idea that only "unanticipated money" could have real effects completely eviscerated the monetarism of Friedman and of Brunner and Meltzer.

Since monetary policy surprises are to no sensible purpose, the rational expectations version of monetarism left central banks with nothing of

"real" usefulness to do. There developed, however, a curious literature in which it was assumed that politicians thought voters would reward them for using monetary policy to fool citizens into working harder than consistent with the natural rate of unemployment. This line of analysis produced two policy recommendations which, although not well founded, have had considerable influence, namely, (a) to make central banks independent so as to constrain the politicians and (b) to steer the banks away from any tendency to play the Phillips curve on their own by giving them only the single objective of low inflation.

This policy doctrine is the perfect expression of the new worldview: the private sector will take care of itself, stabilization policy is a vain ambition, and the government had better be constrained from pursuing it.

But central banks don't just "print money". Within reasonably stable monetary regimes, they have considerable short-run powers to regulate real rates of interest and the real volume of bank credit. Unless the private sector is *perfectly* stable, these powers will be potentially useful. The financial crises which have multiplied around the world in recent years make the argument for totally foreswearing their use anything but persuasive.[12]

Examples

The worst case scenario of the old worldview was that of a great depression brought about by widespread insolvency in the private sector. The cure would be deficit spending by a solvent government. The worst case envisaged in the new view is high inflation engaged in by an insolvent government. It can be cured by trimming government spending – or by higher taxes, although that is less often recommended.[13]

The Washington consensus reflected the new worldview in the emphasis put on restraining government spending and reducing the size of the public sector through privatizations. In Europe, the debt- and deficit-limits of the Maastricht Treaty express the same philosophy. In the United States, California law requires the state government to balance the budget each year.

The Washington consensus was to a large extent formed by the economic history of several Latin American countries in the 1980s – as seen by economists and policy makers in the IMF, the World Bank and the US government. The monetization of large deficits, the high inflation and economic disruption of the problem countries was contrasted to the economic success of Chile. Through the spectacles of the new worldview, one prescription will do for all cases: constrain the government so as to unleash the initiative of the private sector.

The story behind some of these Latin American cases is however not so simple and straightforward.[14] In more than one instance, it starts with large defaults in the private sector threatening the solvency also of the banking system. The government then stepped in to stop the financial crises from

developing into a general depression. This required in effect the "nationalization" of the bad loans. But the poor ordinary tax payer is not easily persuaded to pay for the mistakes or swindles of the once rich and powerful who by then may be living in fair comfort abroad. It was not in the social contract that he should pay. But someone has to pay, and if the ordinary tax payer will not vote the taxes, he has nonetheless to foot the bill through the inflation tax.

Thus distributive conflict often lurks at the bottom of these inflations. When this is the case, it is not to be solved, and may be exacerbated, by IMF conditionality although outside pressure may help a particular government to enact responsible policies. In the case of Chile, which managed to resume growth with price level stability, the distributive conflict that had boiled up under Allende was suppressed by Pinochet.

The 1990s also showed the new worldview in action in the wave of privatizations that swept around the world, enthusiastically supported by the international financial organizations and of course by the private institutions that earned the fees in these transactions. Without a doubt governments everywhere have gotten into areas of economic activity that could be handled as well or better by the private sector. But the pell-mell privatizations in Russia, for example, seem to have been carried through on the simple conviction that getting the people's assets out of government control and into private hands in whatever way and on whatever terms would surely be welfare improving. *Whose* welfare was improved did not seem important. Typically, questions of distributive justice were not given high priority.

The East Asian crises of 1998 also showed the new worldview to be too simple. The Indonesian crisis was triggered by large corporations defaulting on dollar-denominated debt. The ensuing scramble for dollars sent the exchange rate plummeting, thereby demonstrating the new risk that countries run in a world of globalized finance, namely, debt-deflation in a foreign currency. Although the finances of the central Indonesian government had been in reasonable shape for a decade, the IMF, acting on "consensus" instinct, at first demanded fiscal retrenchment and then had to reverse itself. The Indonesia crisis gave rise to the concept of "crony capitalism", a concept that allowed one to maintain the view that the government is still always the problem.

It is not likely that any crises in the emerging economies will have much of an impact on today's American-dominated economics. Perhaps the asset-price bubbles that have grown and burst in countries such as Japan and Sweden will cause some rethinking. Closer to home, the United States stock market bubble (or its exchange rate bubble?) may do more to shift attention back to problems with the private sector and the role of government regulation and macropolicy in stabilization.

Growth theory instead?

The development of equilibrium business cycle theory has had the logical consequence that the leading people in macroeconomics today are losing interest in business fluctuations and are instead turning their attention to economic growth. Assuming that people adjust to business cycles as smoothly and efficiently as to seasonal fluctuations, they calculate that the welfare loss from cycles is relatively trivial. By the force of compound interest the welfare gain from an increase in the growth rate even if by only a fraction of 1 percent will on the other hand be of great magnitude within a generation or two. One recalls that the compound interest argument was made quite eloquently some 70 years ago by Keynes (1930a). But with Keynes there was a crucial difference. When he contemplated the "Economic Prospects for our Grandchildren", he also knew that the train to that future might very well derail. He had an ever-present sense of "the fragility of the existing social order".[15] That sense must have been shared by most educated Europeans between the two wars. Fitoussi is prominent among the fewer Europeans that share it today. But the fragility of the social contract – what an un-American idea!

Conclusion

The two worldviews with which I started are obviously simplistic extremes and as such they will at times prove dangerous guides in a more complex reality. The accumulation of external events in recent years which demonstrate that the untrammeled market system does not always work for the best and that not all troubles can be laid at the door of governments may start the pendulum swinging back once more. Perhaps it is already in motion.

But these long swings in economic understanding do us no good. Although the impulses of external events play a causal role in bringing them about, their propagation internal to the economics profession seems subject to herd behavior. If we are ever to stabilize the pendulum of economic opinion somewhere in the judicious – but ideologically always unsatisfactory – middle and recognize the capacity for good *and* its limits of both markets and governments, it will have to be through people who do not run with the herd but think for themselves and tirelessly follow their convictions. Like our friend, Jean-Paul Fitoussi.

Notes

1 The growing inequality of incomes in recent years is, of course, attracting the attention of prominent economists but hardly within the real business cycle school. In the United States, James K. Galbraith (1998) has played a role corresponding to that of Fitoussi in France.

2 I vividly recall that at the 1997 IEA conference in Trento where Jean-Paul's

paper (Fitoussi 2001) gave both the theoretical argument for and much empirical evidence consistent with this view, Robert Lucas expressed himself shocked to have to listen to such "inflationary" views. Surely, he argued, we knew better by this time. Of course, the ECB was eventually to give Fitoussi the monetary policy that the Bank of France had denied him. The result was that French unemployment finally began to shrink – without inflation.

3 See Jean-Paul Fitoussi (1995).

4 Cf. Leijonhufvud (1998).

5 To be fair, Keynes had set himself a difficult analytical task. On the one hand he had to explain what was going wrong with relative prices and therefore with allocation; on the other, he had to keep track of the flow-of-funds through the economy because households were cash-constrained while firms adjusted output in response to cash flow. Modern theory is easier in not having any reason to deal with these "hydraulic" problems. Even when a cash-in-advance constraint is introduced it only serves to determine aggregate money demand and does not give rise to effective demand failures.

6 In the 1950s, the problem for the "new theories of the consumption function" was to explain why the multiplier was not as large as Keynes had envisaged. In more recent years, the problem has become to explain why it is not as small as rational expectations theory would have us believe.

7 Peter Howitt (2002) stresses multiplier-propagation within the business sector rather than through household consumption behavior. His model, however, hinges on a thick market externality rather than on the liquidity argument stressed here.

8 A number of people have pointed out the overdeterminacy of adding an aggregate supply function to models that include an IS schedule which subsumes reduced forms of market demand *and* supply functions. Cf., e.g., Barens (1997).

9 Keynesian economics in its early days had overcome this kind of partial equilibrium reasoning about the aggregate labor market. But there has been backsliding among Keynesians later.

10 Letting borrowed reserves expand or making them contract are examples of central bank actions that do not change the volume of outside money.

11 Or, if $(S + T + IM) > (I + G + X)$.

12 If low inflation is not to be the only goal of central banks, however, the independence doctrine is in trouble. Making decisions on trade-offs that have distributive consequences is surely an intrinsically *political* responsibility, not to be handed to unelected technocrats.

13 It is reasonably clear what should and can be done in each worst case scenario. It is not easy to prescribe a medicine when, as in Argentina, the worst of both worldviews are combined.

14 Cf. Daniel Vaz (1999).

15 The phrase is Moggridge's (cf. his 1993: 154).

References

Barens, Ingo (1997) "What went wrong with IS-LM/AS-AD analysis – and why?", *Eastern Economic Journal* 23 (1): 89–99.

Barens, Ingo (1999) "From Keynes to Hicks – an aberration? IS-LM and the analytical nucleus of the *General Theory*", in Peter Howitt, Elisabetta de Antoni and Axel Leijonhufvud (eds), *Money, Markets and Method: Essays in Honour of Robert W. Clower*, Cheltenham, Edward Elgar.

Fitoussi, Jean-Paul (1995) *Le débat interdit: monnaie, Europe, pauvreté*, Paris, Arléa.

Fitoussi, Jean-Paul (with Pierre Rosanvallon) (1996) *Le nouvel âge des inégalités*, Paris, Le Seuil.

Fitoussi, Jean-Paul (2001) "Monetary Policy and the Macroeconomics of 'Soft' Growth", in A. Leijonhufvud (ed.), *Monetary Theory as a Basis for Monetary Policy*, London, Palgrave and the International Economic Association.

Friedman, Milton (1968) "The Role of Monetary Policy", *American Economic Review*, 58, March: 1–17.

Galbraith, James K. (1998) *Created Unequal: The Crisis in American Pay*, New York, The Free Press.

Howitt, Peter (2002) "The Microfoundations of the Multiplier Process", paper presented at the January 2002 ASSA meetings.

Keynes, John Maynard (1930a, 1972) "The Economic Possibilities for our Grandchildren", *The Collected Writings of John Maynard Keynes*, Vol. IX: *Essays in Persuasion*, London: Macmillan, pp. 321–32.

Keynes, John Maynard (1930b) "The Great Slump of 1930", *The Collected Writings of John Maynard Keynes*, Vol. IX: *Essays in Persuasion*, London: Macmillan, pp. 126–34.

Keynes, John Maynard (1931, 1973) "An Economic Analysis of World Unemployment", *The Collected Writings of John Maynard Keynes*, Vol. XIII: *The General Theory and After. Part I: Preparation*, London: Macmillan, pp. 343–67.

Keynes, John Maynard (1934, 1973) "Poverty in Plenty: is the Economic System Self-Adjusting?", *The Collected Writings of John Maynard Keynes*, Vol. XIII: *The General Theory and After. Part I: Preparation*, London: Macmillan, pp. 485–92.

Leijonhufvud, Axel (1973, 1981) "Effective Demand Failures", *Swedish Economic Journal*, March, reprinted in *idem*, *Information and Coordination*, Oxford, Oxford University Press.

Leijonhufvud, Axel (1998) "Mr Keynes and the Moderns", *European Journal of the History of Economic Thought*, pp. 169–88. Also in Luigi Pasinetti and Bertram Schefold (eds), *The Impact of Keynes on Economics in the 20th Century*, Cheltenham, Edward Elgar.

Modigliani, Franco (1944, 1980) "Liquidity Preference and the Theory of Interest and Money", *Econometrica*, 12: 45–88, rev. repr. in A. Abel (ed.), *The Collected Papers of Franco Modigliani*, Vol. I, Cambridge, MA, MIT Press.

Moggridge, D.E. (1976, 1993) *Keynes*, 3rd edn, Toronto, University of Toronto Press.

Patinkin, Don (1948) "Price Flexibility and Full Employment", *American Economic Review* 38: 543–64.

Snowdon, Brian and Howard R. Vane (1999) *Conversations with Leading Economists: Interpreting Modern Macroeconomics*, Cheltenham, Edward Elgar.

Vaz, Daniel (1999) "Four Banking Crises: Their Causes and Consequences", *Revista de Economia* IV (1) May: 29–344.

8 The difficult dialogue between the development of macroeconomic theory and macroeconomic policy concerns

Edmond Malinvaud

1 Introduction

Explaining macroeconomic phenomena, and the resulting macro performance of various economies, is a valuable objective for macroeconomic theory. Indeed, a very large part of macroeconomics has no other purpose. The intention of this contribution is not to downgrade in any way the explanatory objective which in all sciences naturally comes prior to concern about applications. But, in order to pay tribute to Jean-Paul Fitoussi, this book rightly stresses economic policy, which was and remains the dominant driving force in his life.

Moreover it makes sense to argue that, since the Great Depression, the main motivation inspiring development of macroeconomic theory was to guide policies. This was obviously true within the Keynesian movement, but also outside it whenever for instance alternative policy rules and their effectiveness were at stake.

Most economists of my generation, particularly among those working for policy advising, have painfully lived the two decades from 1975 to 1995. This was in part because the macroeconomic performance strongly deteriorated in our countries and the policy instruments at our disposal turned out to be much less powerful than we had thought. This was also because we disagreed with much of what was argued during the concomitant reconsideration of macroeconomic theory, as it was led by members of a younger generation, a reconsideration which made a definite impact in the academic world. We felt that these younger theorists were overly self-confident in their criticism, in special features of their scientific approach, and in the adequacy of our market institutions to spontaneously meet perceived macroeconomic challenges.

Now, as the multifarious progress realized by macroeconomic research during the two decades in question is assessed in retrospect, a better sense of realities again prevails. It may be the proper time now to coldly analyse the drama we have lived through. The analysis is a valuable objective not only for this article but also more generally. Indeed, the story is revealing of a permanent difficulty in the dialogue between

the development of theory and the satisfaction of policy concerns in macroeconomics.

We shall start with a brief look at the two decades before 1975, when exaggerations in the then dominant Keynesian camp laid it open to valid criticism. The ground will so be prepared for us to understand first the force of the so-called rational-expectation revolution, second the attraction of the modelling exercises promoted by the real-business-cycle movement. We shall then be in a position to wonder about the possible maintenance in present macroeconomic research on theoretical practices which would still be exploring the same vein that appeared so rewarding during two decades. While recognizing the many valuable contributions brought by macroeconomic research since 1975 to policy analysis, we shall venture a diagnosis about the origin of the deviations we could also witness. The article will end with the, hopefully valid, enunciation of better prospects for the dialogue between macroeconomic policy making and macroeconomic theory.

2 First signs of a mounting storm

Troubles began earlier than 1975, in particular when on two opposite sides of the macroeconomic profession some extremists forcefully argued in favour of policies which were too precise to be warranted by the then existing state of objective macroeconomic knowledge. On his part, Milton Friedman, was certainly right when pleading for consideration of medium and long-run effects of policies, strained what was known about the worth of the quantity theory of money when presenting in 1960 his "Program for monetary stability" focused on a fixed target for the growth rate of a monetary aggregate.

Without insisting on this particular policy recommendation it makes sense here to comment on what could be perceived as being the initial push given by Friedman to the ideas which would flourish in the two later decades attracting our interest. On the one hand, he stood against demand management policies, arguing that in the long run they would do more harm than good. He also pointed to the crucial role of expectations in macroeconomics. He contributed to maintain a vision that imputed to government interventions the macroeconomic evils. This vision had been prevalent for long in a large part of the economic profession and would again appeal to a new generation. But on the other hand, his approach to macroeconomic research was so different from the one which was going to serve as a battlehorse! He always wanted to remain close to observation of phenomena. He never gave much credit to the idea of short-run market clearing. He indeed acknowledged the role of market disequilibria in the short run, for instance in his discussion of the Phillips curve and in his 1968 attempt at defining the natural rate of unemployment. He was quite open-minded about the formation of expectations, in which he recognized the

importance of adaptations. He was never an adept of the modelling strategy that was applied in the so-called new-classical macroeconomics school.

Probably more important than Friedman's disputes with the Keynesians were, for the success of the emerging new-classical school, the obviously excessive policy-analysis claims made in the 1970s by some enthusiasts of structural macroeconometric models. From its inception in the 1940s aggregate demand analysis had been meant to apply as much to policy analysis as to forecasting. Structural macroeconometric models were built in order to serve both purposes, more realistically than simple theoretical models could do. A methodology for policy analysis then appeared, based on a quite natural logic. It was progressively enriched as experience in the use of structural models requiring more precise guidelines and as the ideas embodied in the models became more complex. Confidence in macroeconometric models reached a climax in the early 1970s. It was such as to support in the minds of some econometricians the idea that the mathematical technique of optimal control, developed in engineering for physical systems, could also be used in the determination of dynamic macroeconomic policies, the structural model being then taken as a good representation of the economic system (Chow 1975). The claim was excessive because optimal control techniques so applied had not the appropriate flexibility and could not reach a sufficient degree of reliability.[1] More generally, predictions based on structural models notoriously had a low accuracy, even though they were better than simple extrapolations of the data series. Also, the models were assuming that constant patterns ruled the feedback from observations to behaviours, which was a disputable assumption when policies to be determined could, when announced, induce atypical reactions in expectations and behaviours.

Overall, we may however say in retrospect that such manifestations of over-confidence either by Friedman and a few other "monetarists" or by adepts of structural macroeconometric models, were less outrageous than those which would later come in two major waves.

3 "The rational expectation revolution"

This is not the place for a full and fair history of the so-called revolution. In retrospect we may however concentrate attention on a few points: introducing the rational expectation hypothesis in macroeconomic research was a valuable contribution; promoters of the revolution deliberately ignored market disequilibria, a stand for which they had no justification, given their claims about the effectiveness of monetary policy; the theoretical models presented as sustaining these claims had other reasons to be obviously incredible; the promoters showed no serious interest for empirical validation; subsequent empirical research invalidated their claims; but the revolution had set up a dramatically inadequate reference for imitations, which polluted for too long macroeconomic policy research.

(i) Clearly, relevant studies of macroeconomic policies often have to recognize both the role of expectations and the fact that policies act on a dynamic context in such a way that exogeneity of expectations would be an inadequate assumption. It was only natural that, in a first response to the challenge so posed, research assumed instead that expectations adapted to what was observed. Various formalizations of this simple idea were considered and empirically compared, all of them being "backward-looking" (for simplicity we shall speak here of adaptive expectations). The remark that, by their very definition, expectations are forward-looking, is not a determinant objection against adaptive expectations, but rather points to the fact that the formalizations in question amount to assuming forms of stationarity in the contexts studied.

An alternative approach built on the idea that the proper assumption to make about expectations in a model serving for the study of a policy had to be consistent with the model. In this sense they had to be "rational". This idea indeed makes sense, but it puts a heavy requirement on the definition of the model, which has to be realistic enough in all respects, including the formation of expectations by economic agents. We shall not go further here in discussing all the implications of this remark, we shall rather take a more pragmatic stand.

There is a fair degree of uncertainty about the choice of the good model for research on a policy problem. Experience shows that in many cases it is instructive to have available more than one model, for instance a fully consistent model with rational expectations and another model with adaptive expectations. Working out the conclusions in each case, comparing them, finally testing which one of the two sets of conclusions best fits the data is not objectionable in principle for progress in the analysis and understanding of the problem.

A valuable contribution of theoretical macroeconomic research in the 1970s was to explore how to deal with formal models which include rational expectations. A methodology was then worked out. In this respect it was an irreversible achievement. Similarly, a valuable contribution of research in the 1980s and 1990s was to adopt the pragmatic viewpoint I just indicated and to sharpen our judgement on the respective merits of adaptive and rational expectations in the various fields where they make a difference.[2]

(ii) Here I have to make the point that the most active promoters of the rational-expectation revolution deliberately ignored market disequilibria and that this stand makes their contributions of the time unsuitable for the study of monetary policies aimed at short-run regulation of inflation and activity. Indeed, the fact that the degree of market slacks or tensions often changes in the short run is obvious to any observer, and also obvious is the fact that monetary policy operates through its impact on this degree. But my point is so clear that I need not elaborate on it, before I turn my

attention to other features of the models then used for discussing the effectiveness of anticipated monetary policy.[3]

(iii) In the late 1960s, M. Friedman (1968) argued that the long-term Phillips curve had to be vertical because in the long term adaptive expectations about inflation had to coincide with the trend in actual inflation. The impact of monetary policy on demand pressure could then only be temporary. By the same token it looked natural to think that anticipated monetary policy would more generally be ineffective. As long as this was presented as an intuitive proposition, it only called for empirical validation. But the promoters of rational expectations thought they could prove, by a purely deductive argument, that the proposition was indeed true. Looking in retrospect at the proofs so given, any economist gifted with a fair amount of common sense can only wonder how they could be found persuasive. Indeed, they were based on models whose realism was questionable, to say the least.

Let us have a brief look at a model which is quite close to those most often quoted for the purpose in the 1970s. The economy is made of a number of "islands", namely of sectors which do not trade in the main good but buy inputs from the rest of the economy at prices reflecting an average price level. The output of the main good in island z is determined simultaneously with its price by perfect equilibrium between supply and demand in the local market. The supply is all the higher as the price exceeds more than expected cost and the demand is all the higher as the anticipated quantity of money is more abundant. On both sides of the market agents have rational expectations. The model moreover specifies the properties of the stochastic processes which determine for instance random variations in the demand for goods and in the money supply. Agents have imperfect information about the realizations of these processes, but they make inference from for instance perfect knowledge of the laws driving their dynamic stochastic economy, information about the growth of the money supply in the recent past, and direct observation of the equilibrium of their local market. Solving his model the economist finds that anticipated money supply has no effect on the equilibrium path, which however exhibits a positive correlation between aggregate output and the price level.[4]

Why was such a result so often quoted as supporting the proposition that anticipated monetary policy was ineffective? Probably not because the model would have been found realistic in its representation of the economic structures or in its representation of economic behaviour. On both accounts it was obviously unrealistic. But the proposition was intuitively appealing and the authors who solved the models had shown a degree of mathematical expertise which impressed readers, colleagues and students.

(iv) Such being the case, it is remarkable that promoters of the revolution

paid so little attention to empirical research about the effectiveness of monetary policy. Their disciples indeed went on teaching the same proposition. There was, however, a lively research field on the subject, which may be traced back to the work of M. Friedman in association with A. Schwartz (1963). We may recall also a long list of contributions from C. Sims starting in 1972, or a number of articles more particularly addressed to the same proposition we just discussed or still many developments up to now (see Malinvaud (2000: 1116–43)). By the early 1980s it was already clear that this empirical research did not support the idea that anticipated monetary policy had no effect on output.

We may note in passing that simultaneously it became more and more clear that the formation of expectations by economic agents was seldom well characterized by a pure form of the rational-expectation hypothesis.

(v) The rational-expectation hypothesis and the *de facto* neglect of empirical evidence against attractive mathematical models were only two aspects of the methodology that was forcefully imposed in dominant academic circles by promoters of the revolution. Without listing here all the mottos that so appeared, I believe it is appropriate to mention "the microeconomic foundations" that any macroeconomic model should have in order to be worth examining, not any kind of such foundations but at the very least a utility-maximizing representative agent, the parameters of the utility function being then declared "deep".

The campaign for such a particular methodology was successful, in the sense that the methodology and most of its mottos were adopted, even by many young economists who were called neo-Keynesians because they did not accept full market-clearing. A style of macroeconomic research had been imposed. Macroeconomists had to imitate it. Since the style was very tolerant about the realism of the mathematical model, it was fairly easy to imagine new models which would lead to any pre-selected conclusion that pleased the research worker.

What could informed outside scientists think about a profession in which leading figures so behaved over two decades? Indeed, the behaviour did not fundamentally change after the middle 1980s, as we are going to see.

4 The real-business-cycle movement

Born in the early 1980s and actively developed during again more than a decade, the so-called real-business-cycle theory embarked on the way that had been opened by the rational-expectation revolution, with an equivalent degree of self-confidence and with the same deliberate bias in favour of market clearing. But there was, from the viewpoint taken in this article, an important shift of focus: the main concern was no longer to predict effects of monetary policy, but rather to explain business cycle facts. Thus research became more detached from the demands of policy makers.

Nevertheless, considering the RBC movement is relevant in a discussion about the dialogue between theoretical research and policy making.

In two ways the movement took stands about macroeconomic policies. There was first the proposition that macroeconomic regulation did not really matter, something which can be placed as a parallel to the proposition that, anticipated monetary policies were not effective. Independently, there was also a tendency to play down the role of those policies the study of which did not easily fit into the kind of modelling that was promoted by the movement, no matter how eclectic it became.

Here we shall first consider the real-business-cycle theory strictly defined and the view that regulation was uninteresting. Next, we shall pay attention to the substantial evolution of the movement, insisting on its practice for embodying empirical verification in its methodology, a feature that reveals a different attitude from the particularly objectionable one that was exhibited in the rational-expectation revolution. We shall end commenting on the place given to policy analysis in the work of the movement.

(i) Emergence of the real-business-cycle theory in 1982/3 was associated with that of the provocative proposition according to which business cycles were optimal reactions of the economic system to unavoidable shocks. The proposition contained one assertion, namely that there were no other kind of shocks. It took advantage of a rather simple isomorphism between, on the one hand, an aggregate model earlier developed for the discussion of optimal growth and directly generalized to a stochastic environment, and, on the other hand, a competitive growth model for an economy with in particular a single infinitely-living representative consumer. In view of the assumptions made, the isomorphism indeed implied the optimality proposition.

I shall not discuss here the assertion that all macroeconomic shocks would be technological. By now it is commonly agreed that a business-cycle theory based on such an hypothesis cannot really apply. But R. Lucas (1987) argued that the hypothesis was not necessary for showing that business cycles, such as those that occurred in the US since the Second World War, had really little importance. His basis was the utility function of the representative consumer, from which he speculated about the cost experienced by "the average American family", in terms of consumption level, as a consequence of the variability of consumption around its trend. His results led him to a really small cost, something like one tenth of a percentage point. He then recognized that the cost was a function of the square of variability (as measured by a standard deviation for instance), so that more variability than experienced in the postwar could be definitely more costly. He went on, considering income variability at the level of individual consumers and the role of private or social insurance against at least part of it. Finally, he concluded suggesting

> that the main social gains from a deeper understanding of business cycles . . . will be in helping us to see how to avoid large mistakes with policies that have . . . inefficient side-effects, not in devising ever more subtle policies to remove the residual amount of business-cycle risk.

I understand the point, but I must say that I feel uncomfortable about the argument leading to it (and this applies as well to many other estimates of welfare gains or losses based on more or less special social utility functions, the arguments of which are consumption volumes measured at aggregate levels). The welfare consequences of economic booms and depressions look to me as involving much more than consumption levels (on this point macroeconomists should hear testimonies from labour economists, sociologists and others). This gives me reasons to keep being concerned by the study of macroeconomic stabilization.

(ii) From its appearance, the methodology preached by the RBC school explicitly allocated a place for empirical validation. This recognition certainly made possible, first, a serious scrutiny of the inductive side of the methodology and, second, subsequent improvements brought to it. All along, however, the power of the recommended tests remained low.

In the initial phase validation was not thought to require much more than had been found sufficient during the first decade of the rational-expectation revolution. As we saw, the authors who worked out the "islands model" had pointed out that it led to a positive correlation between changes in aggregate output and in the price level, which was indeed observed in the macroeconomic time-series data (this result of the model was a reflection of the hypothesis that agents were imperfectly informed about the origin of fluctuations in the market equilibrium). Similarly, the model exhibited by the real-business-cycle theory leads to changes in consumption, investment and the labour input which are correlated with the computed changes in output (simultaneously and at a number of lags and leads). The set of the numerical values found for the correlation coefficients after "calibration" of the parameters of the model looks close to the corresponding set computed on observed time series.[5] An agreement is also found on the computed and observed relative amplitudes of fluctuations in the same variables.

Clearly, this is not a stringent test. In particular it is easy to obtain a similarly good fit to the data with traditional Keynesian models in which fluctuations are driven by changes in aggregate demand. There are, of course, changes in the predictions derived from the two kinds of model, but not concerning the variables put on stage, at least for the chosen calibration.

In order to look deeper into the difficulty we must recall how roughly the same fluctuations of the labour input can be derived from two models one of which assumes perfect clearing of the labour market whereas the

other rejects the assumption. In the RBC model the changes in labour input from one period (one quarter, say) to the next, are changes in the labour supply of the representative individual as well as changes in the labour demand of the productive sector. A favourable technological shock boosts labour productivity, hence the return on labour, which leads the representative individual to increase its labour supply. We may speak of intertemporal substitution because the labour supply will on the contrary decrease later when the technological shock will be unfavourable.[6]

Before looking more precisely at the lack of robustness of the empirical validations reported above, we must acknowledge that research of the RBC school developed many models which went far beyond the strict real-business-cycle theory. This appears for instance in the book edited in 1995 by T. Cooley, a faithful member of the school. Some of these developments were immediately embodied in the teaching of the school, in particular models in which aggregate demand shocks played a major part. Others, which took liberties with the competitive equilibrium hypothesis, were more coldly endured, particularly when they allowed for price-stickiness, for non-market clearing and for involuntary unemployment. So, within a school that has become eclectic and now mainly pleads for dynamic-stochastic general-equilibrium models, there are heretics, who accept the methodology but not some of the substantial messages.[7]

Turning back to the methodology of empirical validation, we are also finding a range of contributions addressed to a spectrum of difficulties raised by the main practice of the school, some of these contributions coming from school members. The difficulties all have the same origin, namely an inclination for looking at favourable verdicts whereas unfavourable ones are not so difficult to discover, this in three directions.

Initially a rather small set of variables was selected for second-order multivariate analysis of deviations from trends. Looking at broader sets was advisable. For instance, in T. Cooley's book, at the beginning of his chapter on "Business cycles and aggregate labor market fluctuations", F. Kydland uses results from a broader database in order to list ten notable regularities in observed labour market fluctuations. He then observed that some of these cyclical patterns had been regarded as deviations from the first formulation of the real-business-cycle theory. He then wrote: "Through the interaction of theory and measurement, the deviations or anomalies relative to theory have led to stronger theory as well as to better measurement." It is not the place to go deeper in this chapter. But the starting point is revealing of a difficulty, which could also be exemplified by other references.

Still more critical is the fact that the results of the models do not mimic some of the most relevant dynamic features which are revealed by data analysis of the time series, according to methods developed for the purpose by mathematical statisticians. It so appears that in comparison the RBC models have weak endogenous propagation mechanisms, do not gen-

erate interesting business-cycle dynamics via their internal structure and have to rely instead on exogenous sources of dynamics in the stochastic process of shocks (for references see Malinvaud (2000: 1369–72)).

The empirical methodology of the RBC movement is also quite disputable on its practice for the choice of values given to parameters of the models. These values do not result from estimations of the models themselves, but are rather "calibrated", which means that they are claimed to result from other econometric sources. But when carefully looked at, these sources do not warrant the firm statements with which they are called upon in the RBC literature: they are typically microeconometric studies the results of which are imprecise and not directly transposable to the macroeconomic models under discussion. The point, forcefully made by L. Hansen and J. Heckman (1996), needs no elaboration here. Particularly critical, for the significance of the match with the data claimed in the RBC literature, is the calibration of the elasticity of the intertemporal substitution in labour supplies, a parameter the value of which remains highly disputed.

(iii) I announced at the beginning of this section that, when speaking of policy analysis, promoters of the RBC movement had their own views about which policies should be evaluated. This clearly appears in the chapter written by V. Chari, L. Christiano and P. Kehoe in Cooley's book.

The chapter is "Policy analysis in business cycle models". The aim is to "illustrate how [the RBC] framework can be applied in practice by computing ... optimal [fiscal and monetary] policies". Two examples are worked out, with state-contingent policy instruments, which means that the values of each instrument is fixed in each period (each quarter, say) depending on what has been observed for the shocks up to the beginning of this period. In the article the instruments are tax rates and, for the second example, issues of nominal public debt; the real value of government consumption is exogenously specified. Welfare to be optimized is evaluated according to the utility function of the representative individual. Can any policy maker seriously consider a period-by-period modulation of tax rates? Is the utility of the representative individual the proper measure of welfare for evaluating fiscal policies, in the determination of which redistributive concerns are so important?

The first example is a real economy with three tax rates, on the values of respectively the labour input, the capital input and the interest on the public debt. In the optimal policy modulation is almost fully born by the third tax. The capital input is not really taxed (a result which will not surprise those familiar with the fiscal literature). The second example is a monetary economy without capital: a stochastic constant-returns-to-scale technology transforms labour into output, which is allocated to three parts, the third of which being government consumption. In the utility function the two first parts are not perfect substitutes. Money is devoted to

buying the first part only, which is subject to a cash-in-advance constraint. In the optimal policy modulation is provided by the issue of public debt, i.e. by money growth. It operates through adaptations in a fully flexible price level, hence period-by-period changes in the speed of inflation, and intertemporal substitution in the labour supply. It is hard to view the problems solved in these two examples as good representatives of the demands addressed by policy makers to their economists.

(iv) Let us make, in passing, a brief comment about the positive objective of explaining business cycles, which is not the topic examined in this article. It would be hazardous to pretend today that the aggregate modelling scheme, imposed by the RBC school on the representation of the economy, paid off in terms of a definitely better knowledge of the actual dynamic process generating macroeconomic time series, and this notwithstanding the wealth of specifications to which the scheme was applied. Indeed, the initial enthusiasm about the methodology progressively weakened, even though the modelling scheme was still applied to many extensions of the basic models, thus providing subjects for hundreds of dissertations throughout the world.

5 Subsequent currents and waves

Notwithstanding the many limitations stressed above, we should not underestimate the resistance of some preconceived norms about what might qualify as being an interesting contribution to macroeconomic theory, the resistance also of some entrenched habits in research practice. A distorted concept of what might constitute valid microeconomic foundations was instilled in the minds of too many students and young professors who keep repeating the mottos popularized in the 1980s. Contempt of serious attempts at exacting empirical validation is still now witnessed. Whereas applied policy making did not stop using forms of macroeconometric models that were promoted in the 1960s, little interest is shown still today in the work of the few academic macroeconomists who seriously study the performance of such models and the conditions under which they usefully incorporate such alternative hypotheses as rational expectations.

No discredit seems to mark the utilization of the same basic approach which flourished so much in the 1980s and consists in overselling mathematical models built on specifications which were chosen either in view of the preselected conclusions to be established or for the purpose of launching extravagant theories that may look likely to attract attention in academic circles. The strict theory of the real business cycle was of the latter type. And still recently we saw something similar with the emergence of the "fiscal theory of the price level". Since this new case is revealing and bears on economic policies, it deserves some attention in this chapter.[8]

The denomination of the new theory is a good measure of its content. Indeed, we read sentences like the following: The fiscal theory says that the price level is determined by the ratio of nominal public debt to the discounted value of the stream of future real primary surpluses (see for instance the first line of the abstract in J. Cochrane (2001)). We may see that, with the above positive tone, it was found by a number of economists to provide a plausible explanation of changes in the price level.

This looks like a challenge to common wisdom which holds that, as a first approximation, the price level is determined by monetary policy. Certainly the quantitative theory is not exact in the short term. But indirectly, through effects on demand pressures, on the progressive revision of price expectations, on price making and on wage bargaining, it tends to apply in the long term. Such a common wisdom is now based on a long series of theoretical and empirical works, at both the microeconomic and macroeconomic levels.

How can it be that such a challenge gained force on the basis of so little evidence during the last decade?[9] Was it simply because it was provocative and could feed new mathematical arguments? I shall not go up to the point of maintaining it.

Probably, concerns about the sustainability of the public debt in many parts of the world drew attention to the intertemporal government budget constraint and to the ways to cope with it. When the ratio of the debt to national income keeps increasing, the natural dictum says that taxes will have to be raised or public expenditures to be cut, so that the real primary surplus will be raised. Experience, however, showed that in many cases a burst of inflation wiped out a large part of the real value of government debt, so substantially easing the intertemporal budget constraint. But such events had not been expected, they could seldom be said to have been deliberate, and in any case they were viewed as outcomes of bad policies because they were unfair to holders of government bonds, even more generally unfair to creditors, not to speak of other costs of inflation. The most manifest bursts of inflation indeed occurred during wars, revolutions or other periods of domestic unrest, at times in which it would have been indecent for monetary authorities to stand in the way.

What is new with the fiscal theory of the price level is to consider as normal what was earlier viewed as accidental. It is a new twist given to fiscal theory, a twist favouring the vision in which monetary policy would be permanently passive and public debts would not explode beyond the sustainable.

What can be the worth of this new positive theory of the determination of the price level? A few microeconomic foundations have been proposed, and they looked, at least to me, too far-fetched to be convincing. A macroeconomic test was thus far found negative: it supports the traditional dictum, according to which, when the weight of the public debt increases, fiscal authorities actually tend to react by raising the primary surplus. This

was shown by H. Bonn (1998) on the US annual data from 1920 to 1995, after deletion of the period covering the war (1940–1945). This was also the result obtained by J. Creel and H. Sterdyniak (2001) on a panel of four countries (France, Germany, the UK and the USA) for the years 1970–1999.

In order to discard the interpretation according to which such regressions did not necessarily invalidate the new fiscal theory which holds that the real public debt is a function of future primary surpluses, M. Canzoneri, R. Cumby and B. Diba (2001) proceed to a VAR analysis and look at the impulse-response functions implied by the US data for the years 1951–1995. They conclude that the traditional dictum gives a definitely more plausible explanation than the new fiscal theory, which in order to fit the facts, needs assistance from a rather convoluted explanation. I may add that assuming price makers to be able to forecast future primary surpluses looks difficult to swallow, unless they are said to rely on the traditional dictum for so doing, in which case they will perhaps also forecast that the restrictive fiscal policy will have depressing rather than inflationary effects.

It does not seem that anybody thus far drew normative implications from the new theory to policy making. We may wonder what they might be, whether for monetary or fiscal policy. But as long as the theory is not found to be positively valid, such implications would be premature.

It appears very unlikely that the fiscal theory of the price level will ever reach as high a fame as did the rational-expectation revolution or the real-business-cycle methodology.[10] However, the interest surrounding it shows that echoes of former storms are still being felt.

6 Mischiefs, grieves and harms of the star system ruled by the "publish or perish" motto

The positions expressed in this chapter should not be misinterpreted. I am not an enemy of those who were working on the fields of macroeconomic theory I discussed. Studying the macroeconomic effects of alternative hypotheses about expectations, introducing random shocks into the traditional growth models, even looking precisely at the meaning and implications of the government intertemporal budget equation, all were interesting research objectives. The outcome of such research contains material that is a definite acquisition for macroeconomic theory and applications. Perplexity, uneasiness, and the conviction that I should express them, considering the surrounding enthusiasm, better characterize my feelings during two long decades. The real needs of macroeconomic policy making were too much overlooked by the stars of the time.

Indeed, the star system which tends to prevail in our academic world is the culprit when it becomes dominant. We well understand the virtue of the norm "publish or perish", given to students, and of the norm "publish

what is new and imaginative" given to editors of scientific journals. Their results well stand the comparison with the alternative norm often used in older times when it was rather said: "devotedly serve renowned professors". But in science the main motto always ought to be: "search for what is really true and relevant".

When theorists indulge in scholastics that are less and less relevant, or when they take liberties in selecting favourable empirical support while ignoring more exacting tests, something goes wrong. This is particularly so when all the tricks, commonly used in advertising new products, are also used by authors of scientific papers. Then the scientific ethics is really lost. An old macroeconomist may be allowed to voice a warning: "resist the temptation of fuelling primacy of the star system in our profession".

Reflection leads us to detect two reasons why the mischiefs of this system are particularly damaging in macroeconomics. In the first place, a very large part of relevant research cannot be glamorous. Think for instance about the importance of correct assessments of the medium-term effects of policies. We know that, in many cases, such effects matter more than those materializing during the first year after a decision. But these effects are difficult to gauge: they go through various channels, such as consumption, investment, price making and wage negotiations; a proper modelling has a good chance to lead to an unwieldy specification; data at the disposal of econometricians are not so adequate because most of them rather exhibit short-term effects. Indeed, we cannot expect to really learn about medium-term effects before a number and variety of investigations have delivered their results. None of these investigations taken in isolation will provide a definitive proof, hence attract much attention in the media through which scientific excellence is asserted. In the second place, precisely because there is often a substantial diversity in the published results about a relevant effect, you may pick among them whichever best fits the point you want to make. Clearly such a behaviour, if intentional, does not conform with scientific ethics. But your reader will not know whether indeed that was intentional. In other words, in macroeconomics lapses from scientific ethics may remain long undetected. This gives freedom of manoeuvre to anyone who is unscrupulous and may see an opportunity for taking advantage of the star system.

Beyond the difficult research conditions which explain why the honest results of relevant work in macroeconomic policy making hardly ever generates glamour, beyond the fact that we witness, among research workers on macroeconomic theory, lapses from strict obedience to ethics, also stands the fact that ideological neutrality is difficult to maintain when deciding whether results, even when agreed to follow from a given macroeconomic policy move, would mean an improvement. Anyone who thinks that the ultimate aim of macroeconomic research is objective improvement in policies naturally stands in a bad fix and ought to be excused for his or her sensitiveness.

7 Better prospects for the dialogue between policy making and theoretical research in macroeconomics

Fortunately I can end this chapter with a more optimistic tone. The macroeconomic literature of the last 15 years contains much valuable material for those in charge of advising monetary authorities. Nothing prevents us from expecting that the spirit of this new trend in research will overflow on the whole field of macroeconomic policy making.

About monetary policy the roots of the new trend may be found in various developments: advocacy of VAR analyses by C. Sims since the beginning of the 1980s; as a by-product of the interest brought to expectations, reflections about the credibility of policies and the reasons why authorities may benefit from choosing strategies of commitment; decision taken by Christina and David Romer to closely study the minutes of the US Federal Open Market Committee. These developments led to a lot of serious works on the choice of the macroeconomic targets aimed at by the monetary authorities and on the policy rules followed by these authorities (see for instance L. Svensson (1997) and J. Taylor (1993)). Contributions came from those in charge of preparing decisions with the assistance of macroeconomic tools (for instance F. Brayton *et al.* (1997)), as well as from theorists (for instance R. Clarida *et al.* (1999), discussed in Malinvaud (2000: 1576–80)). Overall, this is a very lively field of relevant research, which monetary policy makers do not ignore.

No fundamental reason ought to bar the generalization of this model to the whole domain of macroeconomic policy making. Everywhere decision makers are using the service of economic advisers who really are intermediaries between the academic economists, who are supposed to know, and those who are supposed to act. Quite a few among renowned scientists spent time as economic advisers. Unfortunately when they wrote for their academic colleagues about their experience, they most often exposed how policy makers received their advice rather than why their advisers were sometimes embarrassed for want of appropriate results coming from macroeconomic theories. There is, however, a too seldom read literature that it would be useful to search for finding at least some answers to this second interrogation (see for instance H. Stein *et al.* (1996), and Malinvaud (2000: 1631–47)).

Notes

1 I tried to formulate elsewhere a balanced judgement on the role of structural macroeconometric models in policy analysis (see Malinvaud 2000: 1545–8).
2 In Malinvaud (2000: 1560–93) I tried to assess the outcome of this second line of research.
3 My aim here is not to discuss a broader question, namely how rational expectations can be embodied in the study of monetary policy. For this different purpose it would be appropriate to examine how the question was later

approached by some of those who had been associated to the initial phase of the rational-expectation revolution. But these subsequent writings made much less of an impact on our profession. See for instance Sargent (1987).

4 For details see Malinvaud (2000: 1112–16).

5 See below for a definition of calibration.

6 The intertemporal substitution in labour supply was first stressed in R. Lucas and L. Rapping (1969) and was for years claimed to be high notwithstanding contrary evidence coming from the vast majority of empirical work on micro data.

7 For a personal survey see Malinvaud (2000: 1359–90).

8 A subsidiary reason here is to signal the survey of the theory by J. Creel and H. Sterdyniak (2001), two members of the staff of OFCE, the institute chaired by J.-P. Fitoussi.

9 Notice this was precisely when, programme after programme, the Japanese fiscal policy engineered downward shocks on the primary surplus which, if anything, only seemed to depress the Japanese price level.

10 Apparently impressed by empirical results and anxious to rescue the theoretical innovation, some now seem to suggest that the positive theory of the price level applies only to periods of rapid inflation. This is a drastic retreat from the claim announced in the definition of the theory, as recalled here. Moreover, we still have to see in what sense the theory, so limited to particular historical episodes, explains the price level during, or at the end of, these untidy episodes.

References

Bonn, H. (1998) "The behavior of US public debt and deficits", *Quarterly Journal of Economics* 113: 949–63.

Brayton, F., A. Levin, R. Tyron and J. Williams (1997) "The evolution of macro-models at the Federal Reserve Board", *Carnegie-Rochester Conference on Public Policy.*

Canzoneri, M., R. Cumby and B. Diba (2001) "Is the price level determined by the needs of fiscal solvency?", *American Economic Review* 91: 1221–38.

Clarida, R., J. Gali and M. Gertler (1999) "The science of monetary policy: a new Keynesian perspective", *Journal of Economic Literature* December.

Chow, G. (1975) *Analysis and Control of Dynamic Economic Systems*, Wiley, New York.

Cooley, T. (ed.) (1995) *Frontiers of Business Cycle Research*, Princeton University Press.

Cochrane, J. (2001) "Long-term debt and optimal policy in the fiscal theory of the price level", *Econometrica* 69: 69–116.

Creel, J. and H. Sterdyniak (2001) "La théorie budgétaire du niveau des prix, un bilan critique", *Revue d'Economie Politique* novembre–décembre, 111: 909–39.

Friedman, M. (1960) *A Program for Monetary Stability*, Fordham University Press, New York.

Friedman, M. (1968) "The role of monetary policy", *American Economic Review* March.

Friedman, M. and A. Schwartz (1963) *A Monetary History of the United States, 1867–1960*, Princeton University Press.

Hansen, L. and J. Heckman (1996) "The empirical foundations of calibration", *Journal of Economic Perspectives* Winter.

Lucas, R. (1987) *Models of Business Cycles*, Blackwell, Oxford.
Lucas, R. and L. Rapping (1969) "Real wages, employment and inflation", *Journal of Political Economy* 77: 721–54.
Malinvaud, E. (2000) *Macroeconomic Theory*, vol. C, North-Holland, Amsterdam.
Romer, C. and D. Romer (1989) "Does monetary policy matter? A new test in the spirit of Friedman and Schwartz", *NBER Macroeconomic Annual* no. 4.
Sargent, T. (1987) "Rational expectations", in *The New Palgrave Dictionary*, Macmillan, London.
Sims, C. (1972) "Money and causality", *American Economic Review* September.
Stein, H., C. Schultze and B. De Long (1996) "Fifty years of the Council of Economic Advisors", *Journal of Economic Perspectives* Summer.
Svensson, L. (1997) "Inflation forecast targeting: implementing and monitoring inflation targets", *European Economic Review* June.
Taylor, J. (1993) "Discretion versus policy rules in practice", *Carnegie-Rochester Conference on Public Policy*, pp. 195–214.

9 What structuralism is – and what errors and omissions it avoids in supply-side and RBC models

Edmund S. Phelps[1]

When Jean-Paul Fitoussi and I wrote our book *The Slump in Europe: Reconstructing Open-Economy Theory* (1988) we drew upon three conceptual departures that, although some two decades old, had not been much exploited by the economics profession. One of those departures was firm-specific investment in employees, which had been introduced in my paper on wage dynamics (Phelps 1968) and a second was investing in customers, which had been introduced in a paper with Sidney Winter for the conference volume *Microeconomic Foundations of Employment and Inflation Theory* (Phelps 1970). Jean-Paul and I combined those novel elements with some monetary building blocks from Keynes and Mundell in order to construct our narratives about Europe's deep slump. Subsequently I wondered whether the new elements would work to tell broadly similar stories – and perchance some new stories – *without* those monetary blocks and with nonmonetary ones in their place. From 1988 to 1992 I developed three nonmonetary models resting on the three elements, bringing them together in my book *Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest and Assets* (1994).

I dubbed the perspective presented by this collection of models *structuralism*. It seemed sufficiently different from the other "isms" to warrant a name. That name was in part a small joke – "to indicate," as Michael Woodford surmised, "that some of the crucial insights were developed on Paris's Left Bank."[2] (Woodford 1992). The idea for that name may have been put in my mind by Hayek's remark somewhere that Mr. Keynes had "not enough structure" in his model. Although the name has not become a household word it has got some recognition. Paul Volcker referred to "Keynesianism" and "structuralism" when introducing a luncheon speaker last year and no one looked puzzled. But what is this structuralism? I have sometimes asked myself the same question and thought for a moment of my emphasis on "incentive wages," which I rested on considerations of employee turnover and others later rested on employee malfeasance. Of course it is a great presentational virtue that a model being used to discuss unemployment actually *has* unemployment in it. However, I would say that the presence of incentive wages and its

byproduct, involuntary unemployment, is not the essential element of the structuralist models.

In my view, the *essence* of structuralism is the richness of its view of business life and thus of the business assets in which firms invest and the expectations they have to form. To exhibit this richness it will be convenient to put aside the view of the labor market in *Structural Slumps* and adopt instead the neoclassical view of labor–leisure choice, in which the labor market clears and is always in equilibrium, i.e., expectations are correct; in the usual formulation, the fixed number of persons in the labor force are always employed and only their hours worked is to be determined. In what follows I first recall the way the parameter shifts that are the focus of supply-side thinking and "real-business-cycle" modeling impact on labor input (i.e., hours) and net pay in such a neoclassical labor market and then proceed to contrast the impacts of the parameter shifts that are the focus of the structuralist models.

The supply-side economics, embodied in an assortment of what I will label SS models, grew out of the dissatisfaction of Samuelson (1951, 1956) and Friedman with the Keynesian effective-demand view of the role of fiscal policy and was effectively launched with the formal argument by Robert Mundell (1971) that fiscal policy is best set with focus on its supply-side effects upon employment and monetary policy best focused on the price level. The bold thesis that evolved is that the level of employment is importantly, if not predominantly, driven by disturbances to the "tax wedge" caused by tax rates, subsidies and tariffs operating through familiar neoclassical mechanisms. The supply-side perspective is succinctly presented in a paper by Casey Mulligan (2002) that seeks to establish the part played by public finance distortion in the movements of labor supply of American workers over the period 1889–1996. For his SS model he starts with the neoclassical model of labor–leisure choice, with its condition $\mathrm{MRS}(C, \overline{L} - L) = v^h$, where the MRS function gives the marginal rate of substitution (MRS) between consumption and work, or "marginal value of time" in terms of the final good, and is increasing in current consumption C and in hours worked L, hence decreasing in leisure; the right-hand side variable, v^h, is the after-tax wage. The latter is related to the firms' demand wage v^f and to the proportional tax rate τ on after-tax wage income by $v^h \equiv (1+\tau)^{-1} v^f$. Invoking pure competition, he equates v^f to the marginal product of labor (MPL). The result is $\mathrm{MRS}(C, \overline{L} - L) = (1+\tau)^{-1} MPL$. The implication is that an increase of τ, in decreasing the right-hand side, operates to decrease hours, given consumption and the value of MPL. Mulligan argues from his empirical exercise that marginal tax rates are well correlated with labor–leisure distortions at low frequencies, but they cannot explain the distortions during the Great Depression, the Second World War and the 1980s. He concludes that the decade-to-decade aggregate fluctuations in consumption, wages, and labor supply do not jibe very well with this competitive equilibrium model.

Another difficulty, about which more later on, is that increases (decreases) of τ may be part of a fiscal-policy shift toward permanently lower (higher) public debt expressed as a ratio to, say, the GDP, in which case the capital market comes into the picture and may possibly make a critical difference for the net effects on initial asset prices and employment.

Real-business-cycle economics, embodied in the typically stochastic RBC models, grew out of the neoclassical tradition (Ramsey 1928; Solow 1956) and was sparked to life principally by Edward Prescott (Kydland and Prescott 1982). Its main thesis is that employment activity is driven by shifts in the *MPL* function – less formally, by exogenous productivity fluctuations stemming from technical changes. If technological advance produces pure labor augmentation, A, the function gives $MPL = A\ \mathrm{F}_L\ (K/A, L)$, where F_L is the derivative of the net production function F with respect to *augmented* labor, AL. Substituting for consumption the expression for net product *minus* net investment gives $\mathrm{MRS}([\mathrm{F}(K,\ AL) - \mathrm{d}K/\mathrm{d}t], \overline{L} - L) = A\ \mathrm{F}_L\ (K,\ AL)$. The RBC theorists are interested primarily in high-frequency fluctuations, hoping to beat the Keynesians at their own game, so, as a consequence, the "vibrations" of A, in the imagery of Robert Hall, have negligible effect on the paths of K and of the rate of net investment, $\mathrm{d}K/\mathrm{d}t$. Hence a temporary rise of A above its trend path, in increasing the right-hand side, operates to increase hours in the MRS function, given consumption. (In fact, consumption also increases, which operates to increase leisure demanded at a given wage rate, so an unambiguous increase in hours requires a further specification of the utility function; the usual one has the property that, supposing household wealth is simply K, hours would remain unchanged if K rose as much proportionately as A and increase if K rose less than that. Under that specification, therefore, the desired result is unambiguous if K changes negligibly with vibrations in A.) The empirical difficulty of the RBC models that has provoked the most comment is the failure of employment to demonstrate the systematic increase in response to positive shocks to A that is the central prediction of the theory. There is also the difficulty that the neoclassical condition that lies at the core of the RBC perspective has the unfortunate implication that, given the wage, consumption and leisure should move in the same direction, which is hard to square with the data. This latter difficulty, however, is the fault of the RBC theory's link to the neoclassical labor market rather than to the shocks that are its focus.

The structuralism in *Structural Slumps*

How is structuralism fundamentally different from supply-side and RBC theories – at any rate the structuralism existing up to and including *Structural Slumps* of 1994? This early structuralism is just as nonmonetary as supply-side and RBC models, so it does not differ there. It does not exclude tax rates or unanticipated technological advances, so it does not

differ in either of those ways. Its fundamental difference is its attempt to build (nonmonetary) intertemporal models on elements of business life that are central to a market economy, especially to the more enterprising economies, which might be called capitalist. I will emphasize in the next section that there are other important differences between my structuralist view and the views of supply-side and RBC theorists – differences that perhaps derive from the aforementioned difference.

Let me now enumerate these aspects of business activity that are central to the structuralist perspective and to go on to show how some of the consequent behavioral mechanisms would impact (if we really wanted to model employment that way) on the neoclassical sort of labor–leisure choice that the other perspectives, supply-side and RBC, have relied upon. To this end it is necessary to surmount or circumvent a small hurdle. To be faithful to the Yaari–Blanchard demographics used in the structuralist models (so as to admit non-Ricardian effects of public debt) we must recognize that workers will plan trajectories of rising consumption and leisure, which, since it means that older people will be consuming less and working less, could complicate the analysis of aggregate, or average, hours worked. In order that the analysis be kept simple it is supposed that the function giving the current rate of utility belongs to a class of functions, which includes $\log c + \zeta \log(\ell - l)$, that make the marginal value of time depend only on the *ratio* of leisure to consumption, not on their absolute values; i.e., the MRS function is homogeneous of degree zero in consumption and leisure. In this case, with all workers facing the same net wage, it follows that, provided the current-utility maximum with respect to hours supplied is interior for workers of every age and thus wealth, we may speak of *the* leisure–consumption ratio, since all workers will exhibit the same ratio, denoting it by $(\ell - L)/C$, which now denotes *average* leisure per *average* consumption.

A dimension of business life introduced by the structuralism is the role of customers in a product market where "friction" in communication impedes (but does not block) the flow of information about prices. This is the Phelps–Winter customer market, in which, in general, a firm profits from the sluggishness of information: a firm can "mark up" its price above marginal cost without at once losing all its customers, and this transient monopoly power gives value to its current stock of customers. Let m denote the markup $(P - MC)/P$, where P is price and MC is marginal cost. Then it is straightforward to deduce that $1 + m \equiv \psi$, where the function ψ makes m inversely related to $\tilde{q}$, the shadow price that firms attach to a customer when taken as a *ratio* to how much output a customer has to be supplied. (That ratio is fully analogous to Tobin's Q ratio.) In this imperfectly competitive framework resulting from information costs, the analogue to Mulligan's labor–equilibrium relationship is $\text{MRS}(C, \overline{L} - L) = (1 + \tau)^{-1}[\psi(\tilde{q})]^{-1} MPL$, where I have chosen to leave in the tax rate, there being no reason to proceed without it. If we substitute for MPL

the parameter Λ and, in the closed economy case, substitute ΛL for C in MRS, then L is fully determined. An increase of $\tilde{q}$ pulls up the right-hand side (i.e., increases the v^h that firms are willing to offer); and, since MRS(ΛL, $\overline{L} - L$) is doubly increasing in L, that induces an increase in hours supplied. Thus the *markup* wedge between net pay and labor's marginal value productivity joins the *tax* wedge as a potential factor in the determination of the equilibrium (i.e., correct-expectations) path of employment, here average hours. Sometimes both are needed in an analysis because they move in opposite directions, so the one helps to escape from the other.

In a subsequent paper (Hoon and Phelps 2002) it is shown that, theoretically, through the "Wall Street" channel from tax increase to the demand for labor that was championed in recent years by Robert Rubin and Lawrence Summers, a current and prospective pay-down of the public debt would operate to elevate asset prices – an effect that exists whether or not the long rate of interest is simultaneously lowered; and that effect could – again, theoretically – elevate the normalized shadow price of customers, $\tilde{q}$, by *enough* to pull up v^h and L by *more* than the contractionary effect from the *supply-sider* channel pushes them down. In this framework, one cannot expect to understand well the medium-term responses of employment (here hours) to fiscal shocks without considering the asset price responses to such shocks, in particular, to current as well as prospective tax changes. The increase in the tax rates introduced in the mid-1990s under the Clinton administration may have served on balance to *boost* employment, contrary to what would be predicted by the competitive equilibrium framework, because the expectation of a decline in the debt-GDP ratio boosted asset prices and thus reduced firms' markups.

Another dimension of business activity that structuralism brings in is the production of capital goods, whose speculative nature, as the Austrians were fond of saying, makes it unlike the production of pure consumer goods (i.e., consumer goods other than consumer durables). Nothing much will come of that distinction if the former always requires renting capital and using labor in the same proportion as that found optimal in the latter, unless one posits, as done by Kenneth Arrow and Mordecai Kurz, some ex post immobility of machines or structures that makes it impossible to – too costly – move capital from one use to the other. My structuralist models take the easier route, supposing instead that production of the capital good is typically less labor-intensive than is production of the consumer good. It will simplify the exposition here without any loss of generality to go to the extreme of supposing that producing the capital good requires no capital, only labor, to which output is proportional, given technology. (This was the assumption of the Austrians and also Wicksell and Kaldor.) Then the marginal value productivity of labor is q_I/Λ_I, where q_I is the real per-unit valuation currently placed on the capital good (in terms of the consumer good) and $1/\Lambda_I$ is the marginal physical product of a manhour in

the capital good industry. Accordingly, the neoclassical condition in the labor market becomes $\text{MRS}(\text{F}(K_I), \bar{L} - L) = q_I/\Lambda_I$. It is now clear that an increase in the valuation per unit placed on the capital good, whether by firms producing on speculation or by firms deciding on the size of their capital-good orders, pulls up the demand wage, thus inducing an increase in the amount of hours supplied.

Readers having some familiarity with my work will know the third element added by my structuralist models is the firm's stock of employees in which it has had to make costly firm-specific investment in the form of "training" for the sake of turning the employee into an asset that is useful to the firm. The value placed on such an asset will depend on the expected stream of quasi-rents generated, which is a matter of the quit and mortality rates, the wage and long-term interest rates. An increase in this shadow price would in the normal structuralist model cause an increase in hiring and thus, with time, an increased stock of employees, but in the neoclassical labor market adopted here there is already full employment and the impact of the increased shadow price in that market can fall only on the wage and thus on hours supplied. Intuitively, one wants to argue that an increase of the shadow price, q_N, would prompt firms to try to hire employees away from one another but, in any equilibrium scenario, the wage must jump up immediately to forestall such an expectational disequilibrium; and the higher wage would induce an increase in hours supplied. However, a full discussion would bring us to the matter of firms' wage policy and its uses to combat quitting, which would take us away from the neoclassical setting in which this chapter has been set. In the incentive wage story to which the employee turnover–training model naturally gives rise, however, we can say that an increase in the shadow price to a higher level would pull up both the number of persons employed and the wage rate.

A unifying theme in *Structural Slumps* is that, in fact, swings in the unit value placed on each of these business assets play a large part in the big, medium-term swings exhibited by the unemployment rate – a theme not invalidated by the further truth that equilibrium models make such asset value swings derivable from parameter swings or shifts. It has continued to be a credo of mine that every capitalist economy is prone to big swings in business activity and that huge swings in the (real) shadow prices of business assets account – in some proximate way – for the large swings in business activity seen over the past century and a half. The idea of a natural unemployment rate, properly set forth, never promised a rose garden of (seasonally adjusted) tranquility. Gylfi Zoega and I developed some econometric evidence for this perspective in Chapter 17 of the 1994 book. And Chapter 18 offered an interpretative "nonmonetary history" of the postwar period from this asset-price perspective – a chapter that could now be improved upon. In recent years Hoon and I have begun to suggest that the role played by asset prices in the determination of employment

serves to illuminate key American episodes over a considerably longer period. These include swings in U.S. labor input that Mulligan found were not adequately accounted for by the competitive-equilibrium framework.

Take the period of the Great Depression, which is essentially the 1930s, and the depressionary epoch running from the late 1970s through the early 1980s. (I have sometimes referred to a "second world depression," which started in the mid-1970s and from which most OECD countries emerged by the mid-1990s – with important exceptions in continental western Europe.) Mulligan found that the wedge between MRS and *MPL* was considerably swollen in these two depressionary periods and that tax distortions alone cannot quantitatively explain the swelling in these two periods. "What drove a 40 percent wedge between marginal product and value of time?" he asks. In a paper on tax cuts using the customer market formulation of the product market (Hoon and Phelps 2002), we obtain $MPL/\mathrm{MRS}(C, \overline{L} - L) = (1 + \tau)[\psi(\tilde{q})]$ and we suggest in view of this relationship is that the part of the wedge enlargement that cannot be explained by an increase in the tax rate is accounted for by a depression in the real (shadow) prices put on business assets; our example is a sag in the real value put on the customer, which tends to enlarge firms' markups, thus driving down the "demand wage" facing workers.

The first half of the 1940s, dominated by the Second World War, is another challenging epoch. Here Mulligan found that despite an increase in federal tax rates from practically zero to more than 20 percent during the war, leisure during the war was lower than implied by the labor–equilibrium condition given by his competitive-equilibrium model. To explain why the wedge between marginal product and value of time was smaller than that implied by the enlarged tax distortion of that period Hoon and I draw upon the empirical finding, highlighted by Mankiw (1985, 1987), of a reduction of real interest rates during the war. Theoretically, this dip can be explained either by Mankiw's own introduction of consumer durables into the standard neoclassical growth model or by the introduction of the differences in relative labor intensiveness in the consumer and capital-good producing sectors (Phelps 1994). In the former case, an increase in government spending on the aggregative good, and in the latter case, an increase in government spending on the relatively labor-intensive capital good, reduces the real interest rate, and raises asset prices, including the shadow prices firms place on their operating business assets, such as their customers. This effect operates in the direction of counteracting the distortionary effect of the increased federal income tax rates in the period.

The last epoch pointed to by Mulligan may be called the Reagan era, which effectively runs in terms of annual data from 1982 to 1989 or so. Mulligan found a reduction in the wedge that could not be fully explained by the decreases in federal labor income tax rates in that era. Although the Rubin–Summers channel would imply that the stock market should

decline if agents formed expectations of a build-up of public debt, it has been argued by Blanchard and Summers (1984) that a favorable shift occurred in expected future profitability as evidenced by the strong stock market performance in the Reagan era and the rise of corporate investment demand in the face of extraordinarily elevated real interest rates over that period. If that is the right story, such improved expectations operated, according to the structuralist models, to lift the shadow price of business assets in general, thus to have induced firms to lower their markups and brought about a decline in the wedge beyond what was achieved by income tax cuts.

I have been going along with a highly filtered, dumbed-down reading of *Structural Slumps* purely for purposes of the above exposition. The book *does* postulate that the underlying forces, such as productivity and tax rates and other influences, are not constants but instead moving parameters, or "forcing functions." And, true, the book *could* have argued, without changing the structural equations, that these moving parameters *swing up and down*, thus causing swings up and down in the values (or shadow prices) put on the various kinds of business assets. That is the kind of thing that RBC theory supposes, in which productivity "jumps" to a higher path for a while, then drops down to a lower path, always averaging around an *unchanged mean*, so that the detrended time series fluctuates randomly (in an AR1 process) around some mean level of detrended productivity. However, such a reading of the 1994 book is too narrow. I was thinking of slumps as very often *permanent*! (Apologies to baseball fans, for whom a slump is implicitly temporary.) The signature themes of the book were mostly, if not all, about (permanent) shocks or parameter shifts having a nonvanishing effect on the natural rate of unemployment: a strong increase in the world real interest rate from one plateau to a higher plateau for the indefinite future, which appears to have happened in the early 1980s; a marked decrease in a country's or a region's rate of labor augmentation, which appears to have happened to continental western Europe and the U.S. in the early 1980s; a huge enlargement of the welfare state, such as occurred in the OECD nations in the 1960s and 1970s.

I am perfectly aware that Robert Lucas, beacon of the rational expectations movement so prevalent in the past three decades, has strenuously opposed that method of analyzing fluctuations. "Beware of theorists," Lucas memorably warned, "bearing free parameters." The contention seems to be that if investigators are to have the license to shift parameters in the model at will, there is no discipline, no rules, regarding permissible parameter shifts to which they will have to submit; consequently there will be no basis for deciding whether the macroeconomic disturbance (temporary or permanent) under study has or has not been fairly explained and understood by any or all those economists supplying their rival explanations. But if all parameters are to be treated as constants, no particular fluctuation observed over some interval can ever be explained; it can only

be seen as the result of disturbance terms. (The situation is no brighter, so far as I can see, if parameters are to be treated as given, never shifting forcing functions of time.) Historical fluctuations, such as the slump initiated in the 1970s, are beyond explanation. True, that truth does not make Lucas's stricture unacceptable. But is it compelling? I do not think so. Lucas's argument overlooks the possibility that we can choose among rival explanations of an historical fluctuation by various criteria: an explanation that invoked only one parameter shift would be more attractive, other things equal, than rival explanations that invoked many such shifts. A parameter shift that could be seen as explaining other historical fluctuations as well would be considered more attractive, perhaps more plausible, than a shift that was unique to the fluctuation under study because it was ungeneralizable.

The RBC theorists did not come down on my head for violating the diktat against exogenous and unanticipated parameter shifts in part, I think, because they suppose that such analyses can be brought under the rational-expectations rules by postulating that there are long- or medium-term *regimes*, perhaps even very *long* regimes, with a supposedly known and unchanging probability attached to each regime; and there are occasional *shifts* from one regime to another. This is the modeling strategy famously proposed – apparently in the spirit of rational expectations – by Hamilton (1988, 1990). An important example has been worked out for the case of two regimes in a recent Columbia dissertation by Steffen Reichold (2001) and a case with any finite number of regimes has been recently studied by Thomas Sargent (2003). I appreciate the skill and ingenuity with which these papers proceed to implement the basic idea. (Reichold's paper makes an interesting empirical estimation from U.S. historical data of how low the probability of the boom phase subject to the constraint that the model offer stock returns in the two states that match historical data.) But I have to say that the basic construct is as much of a stretch of the imagination as the usual rational-expectations premise. Now we need centuries of data to assess "the probability" of each of the regimes, of which there may exist a great number. The regimes are supposed to retain their meaning, whatever it is, over the centuries. For me, such regimes lack any correlates in real life. Why not just identify each regime with the height of the stock market, as one would do in the crudest modeling? Why not just think of a "regime" as the "state of long-term expectations" about the unit value entrepreneurs put on the business asset(s) they use to produce?

Let me recap my theme to this point. I have been arguing that structuralism has a role to play because both real-business-cycle and supply-side economics, in attempting to dispense with the rich structure of business life, generally explain only a smallish portion of the swings in the wedge that they suggest they do and thus only a fraction of the swings observed in the unemployment rate; and this austere economics is

inhospitable to the reality that parameters not infrequently shift, thus shifting the steady-state natural unemployment rate – a phenomenon they "acknowledge" only by invoking "regime shifts." But that is not the only role for which structuralist models are required.

Outright errors in supply-side and real-business-cycle economics

Let me briefly go beyond the indictment that RBC and supply-side economics leave some room for the richness of structuralism. I want to argue now the more radical claim that both RBC and supply-side economics get some very important things wrong. If I am right in that, the best defense of structuralism is offense!

Take first the RBC models. Two errors of theirs are on my mind these days, the first of which is notorious. They get wrong the *algebraic sign* of the effect of an unanticipated rise expected to prove temporary on current employment. Those models predict that workers will increase hours per day or per week to take advantage of the blip in productivity – "making hay while the sun shines." In contrast, the structuralist model resting on the turnover–training problem of firms predicts that, since the shadow price put on having another employee will not rise in proportion to productivity in view of the latter's expected temporariness while the productivity increase itself represents an increase in the opportunity cost of keeping an employee off the production line in order to train new hires, Tobin's Q ratio is reduced, with the result that hiring drops and so the employment level begins falling. The structuralist model resting on the problem of keeping/gaining customers predicts a share drop of employment, since the firms cannot find (and would not want to invest in) the incremental customers to buy the temporarily enlarged capacity output. In fact, it is the latter, structuralist prediction that we find in the data describing temporary deviations in productivity from trend path, not the former, RBC prediction.

The RBC models also get wrong the algebraic sign of the employment effect of the sudden expectation of a *future* step-increase in productivity of a supposed permanent nature. First of all, in the structuralist models, which are centered around the trinity of business assets, such a sudden expectation has immediate and positive effect on the market value of wealth holdings, as asset prices jump. RBC economics must on this account predict a consequent drop in the supply of hours at any given wage – an effect that the structuralist models of the 1994 book blithely disregard, preferring to regard hours as adjusting with glacial slowness to changes in preferences for a shorter work week. Thus RBC economics predicts a drop in labor input and hence in output as a result of this sudden expectation of bountiful shareholder returns starting at some future date. In contrast, the structuralist models uniformly imply increased hiring and a steep rise of employment to result from the lift in the shadow

value that firms will place on the business assets that the sudden productivity expectations will inspire: the firms will want to get started immediately with extra hiring, extra commercial structures and extra customers in anticipation of their increased returns at the future date. The data strongly associate surges in stock markets to increased levels with an (temporary) investment boom and (temporary) boom in employment, not with the predictions of RBC economics.

A minor qualification: the actual RBC models abstract from these assets, thus averting the implication of RBC economics contrasted here with the structuralist implication. If we wanted to work with the literal RBC models there would be no rise in visible market wealth in reaction to the sudden expectation of the future productivity step-increase. But there is no doubt that in a Ramsey-type formulation, the newly expected step-increase would cause an immediate jump in both consumer demand *and* leisure demand – roughly to their new Friedman–Hall permanent levels in the case of an open economy. (In a closed economy case, of course, some RBC theorists would suggest that the interest rate will increase to a point where people's consumption is curtailed to the hours people are willing to work; but whether increased consumption wins the tug-of-war or instead increased leisure looks to be theoretically ambiguous. RBC economics would do better with an OLG model here, but its roots lie in Ramsey–Barro theory.)

The two above criticisms are not unfair, since the RBC modelers believe that the *forte* of these models is their ability to explain fluctuations as the result of productivity fluctuations – and little else, except possibly tax rates. It is not as if they were fundamentally about something else and took some justifiable shortcuts in regard to productivity. (Needless to say, if I have got wrong the implications of the RBC models in these respects I would like to know about it and to correct my mistakes, with apologies. But I do not believe I have got them wrong.)

Do the structuralist models pass these tests. Certainly they do. They predict that employment will decline in response to an unanticipated, actual increase of productivity above the trend-path that is expected to be temporary. They also predict that employment will be driven up by an unanticipated expectation of a permanent step-increase of productivity in the future.

Now take the SS models, the analytical flavor of which is so well caught by Mulligan. His study focused on decade-to-decade changes in the wedge. Yet that is somewhat curious, since the primary message of supply-side economics is about the *permanent* consequences of economic policies that push tax rates to a higher plateau and thus drive up the wedge for all time. Paradoxically, however, it is precisely with regard to permanent consequences that the SS models run into trouble. Their fundamental equation, it will be recalled, is the wedge relation, $\mathrm{MRS}(C, \bar{L} - L) = (1 + \tau)^{-1} MPL$. For Mulligan's medium-term analysis it is plausible to take consumption, C, as given or, equivalently, to take $\mathrm{d}K/\mathrm{d}t$ as given in consumption supply, namely $\mathrm{F}(K, AL) - \mathrm{d}K/\mathrm{d}t$. But for a *long-term*

analysis, which is the normal mode of supply-side economics, C and dK/dt cannot be taken as given. Consider therefore the steady state of an open economy in which the interest rate, r, is given externally and the domestic MPL is given by the factor-price-frontier function. One equation equates steady-state consumption per worker to $v^hL + rW$, where W denotes wealth per worker. Another equation gives steady-state consumption per worker as $\mu(\mu + \rho)W/(r - \rho)$, where μ is the exponential force of mortality and ρ is the rate of pure time preference. Solving for steady-state W as a function of L; we obtain positive W for large enough ρ and C in the latter equation as a more steeply increasing function of W than in the former equation. It follows that an increase in the tax rate shifts down steady-state W while also decreasing v^h. Then, unambiguously, an increase of the tax rate – suppose the tax rate is needed for national defense – lowers consumption; and since a decrease of consumption decreases MRS (and a decrease of leisure increases it) it appears to be possible that the steady-state L satisfying the wedge relation is not decreased.

I would add that in structuralist models, which I have analyzed in some detail, it is quite possible for the steady-state unemployment rate to be invariant to an increase in the burden of the public sector, just as the steady unemployment rate could be theoretically invariant to a decrease of productivity. It appears that supply-side economics must be more careful in its assertions that a higher tax rate inevitably spells higher unemployment or a lower work week. Employment and labor input are not necessarily contracted in the long term, in which wealth and income have experienced a downward adjustment to the decline in the after-tax wage.

This long-term lacuna in standard SS economics – its failure to deliver in any guaranteed way a contraction of labor effort in response to an increase of government purchases and hence the tax rate on labor income, despite the ease with which it delivers a contraction of labor and output in the short term – can be averted in the more general models of the structuralist type. In the 1994 book and some subsequent work I show that high tax rates have an unambiguous bite to them if there are various "inhomogeneities" in the economy that cause the public debt and the welfare state *to shrink less* in proportionate terms than after-tax wages are reduced; then the declines of wealth and nonwage income, including the welfare-type benefits from "social wealth," that emerge alongside the sharp cut of after-tax wage rates are insufficient to bring labor input back to its previous level. A related theme of mine in discussions of the properties of the structuralist models is that in the 1960s, 1970s and 1980s, a very large share of the increased public spending for which tax rates were raised was spending for the welfare state. It is rather like imposing a tax rate on labor in order to issue a lump-sum "demogrant" to the entire working-age population. The possible disincentive effects provoking workers to increase their propensities to quit, shirk, be absent, etc. are obvious; and when the welfare state created is as massive as it was in Europe these effects are apt

to be important – perhaps raising the unemployment rate by a whole percentage point or two.

Did structuralist models err in the 1980s?

Fairness requires that I address the possibility that the structuralist models also err. Notably, it has been suggested by some macroeconomists that the 1994 book does not manage to fit its structuralist models to the record of the U.S. unemployment rate between 1975 and 1989. It is argued that the rate of interest rose as much or more in the U.S. as in Europe in the early 1980s and the rate of growth of productivity slowed down in the U.S. too, even if not as sharply as in France, in the mid-1970s and later years.

The Fitoussi–Phelps monograph of 1988 comes in here. The essential argument was that there was an asymmetry between the U.S. and Europe: the former was largely the cause of the more-or-less shared rise in the real rate of interest, through a sharp increase in public expenditures (the defense buildup), an equal surge in welfare spending and tax cuts opening up a further increase of the budgetary deficit. (Each item amounted to about 1 percent of the GDP.) Some of these tax cuts were corporate investment tax credits, which directly pulled up interest rates but did *not thereby contract* employment by as much as the increased investment induced by the tax credits *expanded* employment. Also, in some models the increase in defense purchases of domestic output also operate to pull up the interest rate while at the same time expanding output in the short term, at any rate. So that is one part of the solution to the puzzle: in Europe the rise of the interest rate was "imported" and so there were no forces behind it that were directly expansionary.

The other part of the solution to the puzzle is that, because there were forces *pulling* up interest rates in the U.S. while in Europe the rates were only being *pushed* up by the rise of rates overseas, there was a consequent real exchange rate appreciation in the U.S., mirrored by a real exchange rate depreciation in Europe. This induced a trimming of markups in the U.S., relative to whatever trend was going on at the time, and a widening of markups in Europe. The evidence in favor of this hypothesis is somewhat problematic, as there were sharp trends appearing in the data before late 1981, when interest rates rose so sharply. I would say that the data are not so strongly against the hypothesis as to require us to reject it. Other, more statistical evidence in long-term time series give some significant support to the general hypothesis that the markup is pushed up by a drop in the shadow value put on another unit of customer stock.

Structuralism beyond *Structural Slumps*

Structural Slumps made compromises in order that it could be read and digested in a few sittings. Each of the three models is characterized by a

phase diagram for the closed-economy case and the open-economy one. The modeling cuts corners at a few places, the acceptability of which will differ from reader to reader. Yet the analysis is pretty straightforward and the results are, for the most part, readily understood.

These simple models are remote from the high-powered models that could be constructed if one were willing to sacrifice tractability for the sake of greater explanatory power and the ability to embrace some new causal forces excluded from the prototype models of the 1994 book. A leading example is the exclusion of a "modern" capital market from the models. There are no Greenwald–Stiglitz managers unwilling to increase debt finance only beyond some point because, though more debt would be good for shareowners, it would increase the risk they would sometimes lose their jobs in a bankruptcy. I felt that informational imperfections were already so rampant in the labor market and product market of my models that I had better not let them spread to the capital market. The cost has been that the 1994 book did not encompass the employment effects of high corporate debt burdens and bad loans at the commercial banks. But I do not believe that this is the main limitation of the 1994 framework.

What was left out of the 1994 framework was the essential character of capitalism, namely, ideas – new commercial products or methods, to be specific, or new technologies making possible new commercial products or methods. In one respect, changes in technologies *were* embraced by the 1994 models, but such technological changes sprang instantly out of nowhere; also, no new investments were required to realize, or implement, them.

The first, small step beyond the mold of the 1994 book was the idea of "structural booms" in which the *prospect* or, more tellingly, the *expectation* of the possibility of a new technology for producing one or more new products excites anticipatory investments in new employees, new customers (in an open economy) and new plant or equipment (Phelps and Zoega 2001). There is nothing objective, measurable here: the expectations are a creative interpretation on the part of entrepreneurs (in the broad sense that includes CEOs of established firms) of future business prospects. It is a big step toward a more subjectivist view of one of the drivers – perhaps the greatest driver – of medium-term swings in business activity. It is the sudden *expectation* of the possibility of engineering and marketing at a profit some new thing, something not heretofore tried at any rate, that fuels the anticipatory investment in new employees and new customers. (The rise of Amazon.com is the paradigm case, in which vast losses were made in preparation for the time when the production and marketing capabilities would be at a level to justify the anticipatory investments.)

While the capitalism genie is out of the bottle, there are already attempts being made to get the genie back in the bottle. A paper by Paul Beaudry and Franck Portier (2003) speaks neither of expectations nor of

ambiguous indications of an uncertain future. They see business fluctuations as driven primarily by a "shock" that does not affect productivity in the short term – on the rise of employment is in evidence – but "affects productivity in the long term." In the interpretation of this shock that they propose, "it represents news about future technological opportunities." This view of the economy rejects the 1911 theory of Schumpeter, in which entrepreneurs create the future, using whatever engineering knowledge has been left by past scientific research, in favor of a return to the views of Arthur Spiethoff and his German school, which attributed investment booms to the sudden expectation of commercial applications of just-announced breakthroughs in science, engineering, navigation, and geographical exploration. My impression is that some of the more spectacular booms of history were indeed Spiethoffian, such as the boom triggered by visions of electrification. And an element of that type, the protocols that made possible commercial use of the Internet, almost surely played a role in the 1996–2000 boom, which was virtually proclaimed with the creation of Netscape and its IPO in 1995.

Another step has not yet begun to be taken. The 1994 book relied upon the investment theory associated with Tobin's *Q* ratio (Tobin 1968). But in an economic system that is predominantly capitalism it cannot be said that any shareowners know the correct valuation of the complex of business assets at the firm that "backs" the firm's shares – or even that a firm's CEOs in any sense "knows" the shadow value to put on an extra unit of this or that business asset. It cannot even be said that the shareowners know the value that the CEO currently puts on having another unit of each of the various business assets. The avant garde argument is that, even if the CEO knows that the shareowners do not know the worth of the underlying business assets, the CEO following Tobin's rule will respond to the arbitrage opportunity presented by any excess in the value of the shares per unit of the asset stock over the cost of acquiring another unit of assets; by investing more in response to such a discrepancy the CEO will increase the number of assets represented by each share. Though the CEO and the shareowners alike are totally ignorant, Tobin's rule of investing when *Q* exceeds one brings about an increase in the underlying stock of assets behind each share; so the initial shareowners end up enriched: each such owner has an "ownership claim" to a greater stock of assets than before, however much or little that asset is worth. (From a macro standpoint, however, the firms, in issuing shares as long as their *Q*s are greater than one, will presumably drive *down* the share price; even if saving is unmoved, so investment in the aggregate cannot increase, the sale of indefinitely more shares must drive down the price (real or nominal) of shares. Hence the shareowners are made worse off, since their shares are worth less on the market and, in the aggregate, they are not backed with more assets in physical terms than before.) To justify selling more shares the CEO can say that whether or not he takes advantage of the high *Q*

others will do so and thus send all share prices down; so it is better for his shareowners if he tries to act first or early, thus enriching existing shareowners. But the shareowners can say that they are not made better off by any increase in assets per share, that only the share price matters; so there is no mandate, nor obligation, of the CEO to issue more shares in the hope of investing more at the expense of other firms.

Another difficulty for *Q* theory is that it is not obvious that, if investing more, the CEO ought to invest in more assets of the kind already in stock; perhaps a different kind of business would be more valuable. Or perhaps it would make sense to the CEO to issue more shares with which to acquire more bonds, say, corporate bonds, including junk bonds. Then the CEO will be able to say that he has increased the stock of such bonds to which each share is an ownership claim. If these brief remarks are right, a necessary step in the development of an investment-based theory of employment will be to formulate the theory of corporate investment under capitalism, with its vast uncertainty and ignorance on the part of both CEOs and shareowners as a consequence of the great novelty, variety, and change that capitalism presents. After all, much of the structuralist message is that employment must be depressed where – and when – CEOs are unwilling to invest in new employees, new customers, and new facilities. And yet at the very core of that theory we realize that we do not really understand very well what it is that CEOs are trying to do or how they want to portray themselves and how, as a result, they respond to changes in the environment they face. Existing models suppose that everyone knows the value of all business assets and of all shares. But that is at odds with capitalist mechanisms of entrepreneurship and innovation.

The third step in my structuralist agenda is the introduction of institutions, including capitalist institutions, into a theory of corporate investment and innovation. It is plausible to believe that much of continental Europe's low employment, both its low participation rates and its high unemployment rates, can be laid to a deficiency in the dynamism fostered by its operating system – the infrastructure of institutions that support its market economy. It is a reasonable hypothesis that Europe's employment levels would pick up to permanently higher levels if several economic institutions were retired or reformed and several new institutions installed. The modeling that will make this view operational and testable largely remains to be done.

Notes

1 I would like to acknowledge and emphasize that the development of "structuralism" in my books and papers over the past dozen years owes a huge debt to extensive collaborations with Hian Teck Hoon, Singapore Management University, and Gylfi Zoega, Birkbeck College and the University of Iceland.

2 In the "structuralism" of the Sorbonne's Claude Levi-Strauss, if I have it right, a society's world-view – what it thinks about things – is encoded and transmitted through a rich structure of narrative myths.

References

Beaudry, Paul and Franck Portier (2003) "Stock Prices, News and Economic Fluctuations," CEPR Discussion Paper 3844.

Blanchard, Olivier and Lawrence H. Summers (1984) "Perspectives on High World Real Interest Rates," *Brookings Papers on Economic Activity*, issue 2.

Fitoussi, Jean-Paul and Edmund S. Phelps (1988) *The Slump in Europe: Reconstructing Open-Economy Theory*, Oxford, Blackwell.

Hamilton, James D. (1988) "Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control* 12, June/September: 385–423.

Hamilton, James D. (1990) "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics* 45, July/August: 39–70.

Hoon, Hian Teck and Edmund S. Phelps (2002) "Tax Cuts, Employment and Asset Prices: A Real Intertemporal Model," Columbia University, Department of Economics, Discussion Paper Series #0102-70.

Kydland, Finn and Edward C. Prescott (1982) "Time to Build and Aggregate Fluctuations," *Econometrica* (6), November: 1345–69.

Mankiw, N. Gregory (1985) "Consumer Durables and the Real Interest Rate," *Journal of Political Economy* 93, May: 353–62.

Mankiw, N. Gregory (1987) "Government Purchases and Real Interest Rates," *Journal of Political Economy* 95, June: 407–19.

Mulligan, Casey (2002) "A Century of Labor-Leisure Distortions," National Bureau of Economic Research, Working Paper 8774.

Mundell, Robert A. (1971) "The Dollar and the Policy Mix: 1971," Essays in International Finance Section, Princeton University, no. 85, May: 3–28.

Phelps, Edmund S. (1968) "Money-Wage Dynamics and Labor-Market Equilibrium," *Journal of Political Economy* 76 (4), Part 2, August: 678–711.

Phelps, Edmund S. (ed.) (1970) *Microeconomic Foundations of Employment and Inflation Theory*, New York, W.W. Norton and Co.

Phelps, Edmund S. (1994) *Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest and Assets*, Cambridge, MA, Harvard University Press.

Phelps, Edmund S. and Sidney G. Winter Jnr (1970) "Optimal Price Policy under Atomistic Competition," in Phelps (ed.), *Microeconomic Foundations of Employment and Inflation Theory*, New York, W.W. Norton and Co.

Phelps, Edmund S. and Gylfi Zoega (2001) "Stocks and Jobs: The Valuation of Firms and Business Assets as a Factor in Structural Slumps and Booms," *Economic Policy* April: 85–116.

Ramsey, Frank P. (1928) "A Mathematical Theory of Saving," *Economic Journal* 38, December: 543–59.

Reichold, Steffen (2001) *"New Economy" and "Productivity Slowdown": Can Learning about Rare Regime Shifts Explain Aggregate Stock Market Behavior*, Doctoral dissertation, Columbia University. November.

Samuelson, Paul A. (1951) "Principles and Rules in Modern Fiscal Policy," in *Money, Trade and Economic Growth: Essays in Honor of John Henry Williams*, New York, Macmillan.

Samuelson, Paul A. (1956) "The New Look in Tax and Fiscal Policy," Joint Committee on the Economic Report, 84th Congress, 1st session, *Federal Tax Policy*

for Economic Growth and Stability, November 9, 1955. Washington, DC, U.S. Government Printing Office, 1956.

Sargent, Thomas J. (2003) "Acknowledging the Possibility of Regime Shifts." NYU Discussion Paper, February.

Solow, Robert M. (1956) "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70: 65–94.

Tobin, James (1968) "A General Equilibrium Approach to Monetary Policy," *Journal of Money, Credit and Banking* 1.

Woodford, Michael (1992) "Seven Schools of Macroeconomic Thought by E.S. Phelps: A Book Review," *Journal of Economic Dynamics and Control* 16: 391–8.

10 What should we mean by "growth policy?"

Robert M. Solow

The underlying theme of this tribute to Jean-Paul Fitoussi is the perennial one of the relation between economic theory and policy. But the surface argument is much more contemporary, concrete and specific.

The story I want to tell, in a nutshell, is this. For various reasons – empirical, technical and casual – modern ("neoclassical") growth theory has centered its attention on steady-state exponential growth. The main function of a model appears to be to determine or "explain" the long-term rate of growth. To this end, special assumptions are made whose only real purpose is to guarantee the existence of one or more exponential steady states. This purpose is soon forgotten, and the assumption of convenience become standard.

When the models are intended to serve as a guide to policy, the tacit presumption is that the goal of growth policy is to increase the long-term growth rate. Further assumptions are needed to allow conventional policy instruments to have an effect on the steady-state rate of growth; these assumptions also become standard, as if they have some independent validation. It is easy, dramatic and satisfying for policy makers to speak of their intention to raise the growth rate. And so the very vocabulary of growth policy becomes identified with moving the growth rate. (In everyday political discourse, even business-cycle upswings are described as "growth." This elementary confusion of supply and demand is not the issue here.)

I think that this pattern is unnecessary and, what is worse, dysfunctional both for theory and for policy. It seems more natural to define as "growth policy" anything that permanently lifts the cyclically-corrected trend path of the model economy, even if it only adds a constant percentage to a preexisting path. What else would one call a policy that increases output by 2 percent forever? The deeper point is that once growth theory is freed of the need to generate exponential paths and ways to tilt them, the way is open to a more discriminating choice of assumptions, governed more by empirical validity and less by artificial convenience.

My task now is to put some flesh on this skeleton of a narrative. There is always a complicated mix of internal and external factors underlying the

emergence of a new theory, and shaping the form it takes. In the case of neoclassical growth theory, these certainly included Nicholas Kaldor's influential statement of a handful of "stylized facts" that, as he claimed, characterized the long-run macroeconomic behavior of advanced industrial capitalist economies. That neat characterization amounted to a compact description of an exponential steady state. This picture was supported and extended by Lawrence Klein on "The Great Ratios of Economics." It is not surprising, in that context, that the simplest complete growth models should feature not only a steady state, but a steady state with a substantial basin of attraction. The model was thus a reading of the background facts as they were understood at the time. (After 50 more years of data, the stylized facts now look a bit more complicated.)

For so simple a construction, there is not much to be said about matters of technique. The assumption of universal constant returns to scale was known to be a convenient simplification, especially useful in allowing and even simplifying the competitive-market interpretation of the model. Decreasing returns to scale would be much less promising for the existence of sustainable steady state.

I should also mention a wholly casual factor of a sort that can sometimes matter in a small research community. The early stages of modern growth theory were strongly influenced by the work of Roy Harrod, beginning with his 1939 "An Essay in Dynamic Theory," and even more so his 1954 book, *Toward a Dynamic Economics*. Harrod thought in terms of somehow *replacing* levels of output and employment with rates of growth of output and employment as the central theoretical objects of a macroeconomic model. This helped to establish the habit of talking about rates of growth, and thus the habit of focussing on the steady state as the central characteristic of a growth model. For all his creativity, Harrod had little command of analytical technique. An ordinary differential equation would not have been a natural mode of thought for him. If we know the rate of growth of x as a function of time, then we know x as a function of time, up to a multiplicative constant; and if we know x, then of course we know its rate of growth. In Harrod's mind, however, to move from statics to dynamics was somehow to move to "the" rate of growth as a more or less independent concept. This idiosyncrasy may have left its trace on the origins of modern growth theory and its emphasis on the exponential steady state.

One might think that talking about rates of growth is a merely grammatical shift. Anything sayable in one vocabulary is sayable in the other. But that is not the whole story. There is an effect on the models themselves. If one is after an exponential steady state, one will be tempted to choose assumptions that favor the existence of an exponential steady state. The most pervasive example of this tendency is no doubt the almost universal assumption that technological change (whether exogenous or endogenous) is purely labor-augmenting. There is no serious empirical

basis for this choice; but it is necessary to guarantee that steady-state growth should be at least possible under constant returns to scale.

A more extreme example is the common use, usually without comment, of the Cobb–Douglas production function, one of whose great conveniences is the unique property that any factor-augmenting technological change can be expressed as purely labor-augmenting. We even have very interesting models of endogenous factor bias that lead to the conclusion that labor-augmentation is the natural outcome of the choices made by firms. But it is far from clear that the assumption of pure labor-augmentation is empirically defensible. There are some econometric tests that sharply reject it. At a minimum, one should be wary of conclusions from growth models that depend on such assumptions; more broadly, it would seem to be a good idea to explore what happens under alternative assumptions about the incidence of technological change.

An exactly analogous situation holds with respect to increasing returns to scale. The presence of increasing returns to scale, without further restrictions, rules out the possibility of steady exponential growth. But if the underlying aggregate production function has the particular form $F(K, L^a)$, with $a > 1$ and $F(x, y)$ homogeneous of degree one in x and y, then increasing returns to scale prevail (in the sense that doubling K and L will more than double output), and yet steady exponential growth is possible. Once again, the Cobb–Douglas production function automatically has this property.

More subtle examples of a similar tendency are ubiquitous in the endogenous growth models that have dominated the literature since the pioneer articles of Paul Romer and Robert Lucas. Instead of citing particular instances, I will just describe a sort of generic example. Many endogenous growth models work through a variable that I will call $A(t)$ (where A can stand for "Anonymous"). In some models A represents the level of labor-augmenting technology. In others it may represent the range of intermediate goods provided to the final-demand-producing industry, with the assumption that use of a wider variety increases productivity. In another group of models it may represent the stock of human capital or some analogous productivity-enhancing input. Generally $A(t)$ governs the level of output (*ceteris paribus*), so that the rate of growth of output is related to $A^{-1}\mathrm{d}A/\mathrm{d}t$.

In an endogenous growth model, increments to A have to be produced; that is to say $\mathrm{d}A/\mathrm{d}t$ is itself an output. So the model needs a production process for $\mathrm{d}A/\mathrm{d}t$. The natural formulation is that the inputs into that production process are the pre-existing stock of A and the current allocation of some generalized resource R. One could write the production function as $\mathrm{d}A/\mathrm{d}t = G(A, R)$. But that will not generally serve the purpose of endogenous growth, if by that phrase we mean the endogenous determination of a steady-state rate of growth of A, and therefore of aggregate output. No such steady-state rate of growth may exist.

The key step, sometimes inconspicuous, is to specialize $G(A,R)$ to the form $Ag(R)$. Suppose one started with something a little more general, say $G(A, R) = A^b g(R)$, it is easy to see that the rate of growth of A would then tend ultimately to zero as A becomes very large, for any bounded R, if $b < 1$. It is slightly less easy to see that $b > 1$ implies that A would explode to infinity in finite time for any fixed level of R. To do the required job, the parameter b must be exactly unity, i.e. $G(A,R) = Ag(R)$. Then, of course, $A^{-1}\mathrm{d}A/\mathrm{d}t = g(R)$. In other words, the rate of growth of A, and therefore the rate of growth of aggregate output is controlled by the *level* of resource R allocated to R&D or education or whatever activity is being emphasized.

This device accomplishes two purposes. It assures the existence of an endogenously determined steady-state growth rate. And it offers a quick route to policy: the steady-state growth rate depends on the *level* of R, the resources devoted to increasing the value of A, and that is exactly the sort of magnitude that can be influenced by differential tax and subsidy policies or by other traditional kinds of incentive programs. All of a sudden it has become easy in principle to bring about major changes in the path of output. It should be clear that $A(t)$ can be endogenous *without* recourse to the special assumptions needed to insure the existence of a steady-state growth rate. The general formulation $\mathrm{d}A/\mathrm{d}t = G(A,R)$ is enough to make $A(t)$ endogenous, but there is no guarantee that any steady-state path exists. It is the unargued belief that growth theory is about steady-state growth rates that leads to the domestication of these highly special assumptions within the standard models.

Growth models, like most macroeconomic models, are intended to say something about policy, in this case especially when the stated objective is faster economic growth. (I leave aside the different issue of economic development, with emphasis on institutional change.) The notion of "faster economic growth" is loosely stated, and properly so in my view. But there is a tendency for extra-scientific discussion to focus on "the growth rate," more out of intellectual laziness than anything else. Nevertheless I think the tendency is abetted by the literature of growth theory, along the lines I have been discussing. Both in the political arena and in the *Journal of Economic Growth*, it seems to be taken for granted that a policy aimed at faster economic growth is by definition a policy aimed at increasing the sustainable long-term growth rate. The nature of political discussion encourages theory and theory encourages political discussion in these terms, to the disadvantage of both.

The classical case is policy designed to generate a permanent increase in the fraction of output devoted to investment in plant and equipment, in other words to straightforward capital-deepening. In old-fashioned growth theory, this has no effect on the steady-state growth rate. But aggregate output per worker will be permanently higher, even in the absence of technological change, and aggregate output itself will be higher than it

would otherwise have been, and by an increasing absolute amount. Why should this not be characterized as a contribution to economic growth? (Assuming all this happens on the "right" side of the Golden Rule, the path of consumption per person will first dip, then rise, surpass its earlier level, and stay permanently higher.)

We are talking about "mere" definitions, of course; but definitions can affect priorities. In this case the effect is to downgrade important *long-term* policy options. It goes without saying that a successful attempt to raise the steady-state growth rate can have fantastically favorable effects, eventually dwarfing lesser achievements. My suspicion, however, is that theory plants these possibilities mainly in order to find them, and thus conveys the impression that using public policy to increase the growth rate is much easier than the reality suggests it is. No one can afford to sneeze at a policy that merely succeeds in making aggregate output 2 percent higher, year after year, than it otherwise would have been.

I used the familiar example of capital-deepening to make the point easily. There are many other sorts of policies that work in the same way. A permanent increase in the stock of human capital per worker is the closest analogy; further afield are such possibilities as a maintained improvement in economic efficiency (through the elimination of monopoly power, the reduction of labor-market discrimination, or through the exploitation of comparative advantage, for instance). It is hard to imagine why any action that leads to a sustained proportional increase in output should not be regarded as a policy favoring economic growth.

The same case is also valid within the same framework of endogenous growth theory, and perhaps even more sharply. In the notation adopted earlier, suppose that $\mathrm{d}A/\mathrm{d}t = A^b f(R)$, where $b < 1$. A permanent increase in R will bring about an increment in A that is unbounded with time, but increases "only" like a power of t ($1/(1-b) > 1$ to be exact). There is no permanent increase in the rate of growth of A; indeed the rate of growth falls through time like $1/t$. Something similar will be true of aggregate output. In our current vocabulary, this is not a contribution to economic growth, and trying to increase R does not constitute "growth policy." This seems to me to be misleading; the vocabulary suggests a distinction that goes against common sense, and for no good reason. It would seem more natural to count as growth policy anything that generates an increment to output that increases through time. (I will not quibble over the proper label for a constant increment.)

Thinking this way has implications for growth theory. The proper object of study is the growth path, not just a possible exponential path that may not exist. No matter how convenient it is to do comparative dynamics on exponential steady states, the cost of that convenience seems to be too high. Perhaps that simplification was justified when numerical solution of dynamical equations was recondite and expensive. Modern computers and computer graphics have certainly shifted the balance. Comparative

dynamics through computer simulations can accommodate a more valid collection of assumptions, with no apparent loss of definiteness.

A final note: are there other examples in macroeconomic theory where a fairly casual choice of specification has led to a deformation of theory and a corresponding awkwardness in policy? One possible parallel may be modern monetarism. It was not necessary to focus so much attention on *the* "stock of money." A more general formulation was available via the Tobin–Brainard many-asset general-equilibrium approach. Perhaps as a legacy of the quantity theory and the quantity equation, perhaps because of Milton Friedman's aversion to Walrasian ways of doing things, perhaps for other reasons, monetarism and its offshoots put the money-stock in the center of its picture of theory and of policy. Advocacy of a policy of steady growth of the money supply was only the sharpest example of this tendency.

The advent of modern financial engineering pulled the rug from under this emphasis on a well-defined money-stock. When substitute assets can be created ad hoc, no single asset is likely to be a reliable index of the general state of liquidity, and monetary policy can not rely on control of a single asset to accomplish its purpose. It is possible that this failure, in addition to any other structural weaknesses, has contributed to the decline of monetarism as a theory and as a guide to policy.

If there is a general lesson for macroeconomic theory in these considerations, it is not that convenient simplifications should be avoided. That would be the end of systematic reasoning about the economy. It is rather that convenient simplifications have to be chosen with care. Of course they have to be convenient and simplifying; it is also important that they do not have the unintended effect of deforming the theory by excluding important possibilities or directing attention to arbitrary special cases. The mobilization of economic theory in the interests of economic policy is a difficult balancing act between realism and abstraction. That is why we honor people like Jean-Paul Fitoussi with the will and the intellect to practice it.

11 Rational expectations equilibria

A recursion theoretic tutorial*

K. Vela Velupillai

1 Preamble

> If a macro-system as a whole has coherence, perhaps it would be useful to study directly the reasons that determine its coherence. This probably is the course underlined by Keynes when he stressed his intention of studying 'the system as a whole'. If a macroeconomic logic partially independent of that which determines individual behaviour exists – and underemployment equilibrium is surely an equilibrium relative to the system and not to the individuals composing it – perhaps that logic deserves to be analysed in itself. . . . My conviction is that macroeconomics has its own dimension which must be considered and not just alluded to.
>
> [8], pp. 27–8

The two fundamental principles that underpin the study of a macroeconomic 'system as a whole' are, first, the 'fallacy of composition' and, second, the idea known variously as the 'paradox of thrift', 'paradox of saving' or, more dramatically, as the 'Banana parable' (cf. [12], pp. 176–8). The ubiquitous 'representative agent' has dispensed with these homely wisdoms of a macroeconomic logic. As a result the momentous macroeconomic issues of growth, fluctuations, unemployment and policy are disciplined by the logic of microeconomic behavioural determinants. It could so easily have been otherwise had we, for example, paid more serious attention to one of the great masters of our subject who straddled the micro-macro divide, John Hicks, when, in his summarising statements of the 'Final Discussion' after the IEA Conference on 'The Microeconomic Foundations of Macroeconomics', pointed out:

> We had been supposed to be discussing the microeconomic foundations of macroeconomics, but we had come to realise that there were several kinds of macroeconomics, each probably requiring its own foundations, and though they overlapped they were not wholly the same. One had to distinguish at the least between macroeconometrics

> and 'macro-political-economy'.... [W]e had been much more concerned with 'macro-political-economy'.
>
> There was a close relation between macro-political-economy and social accounting, so ... it might be useful to arrange our problems in relation to the social accounting framework in order to see how they fitted together.
>
> [11], p. 373

One of the great merits of Jean-Paul Fitoussi's work as a macroeconomic theoretician, and as a passionate advocate for an active role for policy, has been his ever vigilant attention to the above 'two fundamental principles of macroeconomics' underpinned by their relation 'to the social accounting framework'. Macro-theoretical propositions derived solely on the basis of microeconomic theories, particularly if they are not constrained by the 'two fundamental propositions of macroeconomics', have always left him with a sense of unease. Thus, policy ineffectiveness propositions, based as they are on strong rational expectations hypotheses, time inconsistency results and equilibrium interpretations of fluctuations and unemployment are examples where Fitoussi's critical antennae have been seriously disturbed over the past two decades.

For years I have, myself, been struck by a strange anomaly. Many of the fundamental concepts that lie at the basis of Newclassical Macroeconomics – policy ineffectiveness, credibility, time inconsistency, rational expectations, the advantages of (transparent) rules over (enlightened) discretion, etc. – were also those that informed the work of the 'old' Stockholm School economics – particularly the work of Erik Lindahl and Gunnar Myrdal from the early 1920s through the late 1930s. They, in particular Lindahl, also worked these themes and concepts into dynamic equilibrium schemes. I cannot find a better, clearer, statement of the dynamic economic environment, in which what eventually came to be known as the rational expectations hypothesis, than Lindahl's discussion of the idea in a presentation of his vision of the Keynesian system (but it was only a rewording of a basic idea that had been almost a touchstone of his work on monetary policy and capital theory during the 1920s and early 1930s):

> It also seems reasonable to postulate an interdependence between the variables entering an economic system in the case concerning the determination of the conditions for *correctly anticipated processes*. **These conditions are that the individuals have such expectations of the future that they act in ways which are necessary for their expectations to be fulfilled.** It follows that the interdependence between present and future magnitudes is conditioned in this case by the fact that the latter, via correct anticipations, influence the former. If we also choose to describe such developments as **equilibrium processes**, this implies

> that we widen the concept of equilibrium to include also economic systems describing changes over time where the changes that take place from period to period do not cause any interruption in, but, on the contrary, are an expression of the continual adjustment of the variables to each other.
>
> [13], p. 27; bold emphasis added

However, their fundamental political sympathies were very similar to those espoused by Fitoussi and they made their framework – accounting systems *par excellence* – substantiate an active role for policy. This made me wonder whether there was something special about the language[1] within which the newclassicals developed their concepts and made them work had a role to play in the scope of the conclusions they reached.

Thus, in recent years, I have tried to resolve the anomaly mentioned above by framing aspects of Newclassical Macroeconomics with the formalism of an alternative mathematics, of recursion theory, and asking pertinent algorithmic and dynamic questions. Essentially, I have replaced the use of the standard topological fixed-point theorems that have been used to encapsulate and formalise *self-reference* (rational expectations and policy ineffectiveness), *infinite regress* (rational expectations) and *self-reproduction* and *self-reconstruction* (growth), in economic contexts, with two fundamental theorems of classical recursion theory.[2] The idea of self-referential behaviour is, for example, formalised by considering the action of a program or an algorithm on its own description.

A theoretical framework must mesh smoothly with – be consistent with – the empirical data generating process that could underpin it from methodological and epistemological points of view. I do not use these loaded words with grand aims in mind; I refer to the simple fact that a process that generates the macroeconomic data that is the basis on which the processes of scientific validations of any sort can be performed must do so in a way that is consistent with the way the theoretical model postulates the use of the data. I refer to this as a 'simple fact' in the elementary and intuitive sense that data that must be used by rational agents will have to respect their cognitive structures and the structures of the processing and measuring instruments with which they – and the macroeconomic system as a whole – will analyse and theorise with them. There is no point in postulating data generating mechanisms that are incompatible with the cognitive and processing and measuring structures of the analysing agents of the economy – at the individual and collective levels. In one of my own collaborative writings with Fitoussi, we have touched upon themes of this sort ([9], esp. pp. 225–32).

In this essay I try to formalise the idea of *Rational Expectations Equilibria* (*REE*) recursion theoretically, eschewing all topological assumptions. The title has the qualifying word 'tutorial' to emphasise the fact that I want to try to suggest a modelling strategy that can be mimicked for other

concepts and areas of macroeconomics: *policy ineffectiveness*, *time inconsistency*, *growth*, *fluctuations* and other dynamic issues in macroeconomics. All recursion theoretic formalisations and results come, almost invariably, 'open ended' – meaning, even when uniqueness results are demonstrated there will be, embedded in the recesses of the procedures generating equilibria and other types of solutions, an indeterminacy. This is due to a generic result in computability theory called the *Halting Problem for Turing Machines*. It is a kind of generic *undecidability* result, a counterpart to the more formal, and more famous, Gödelian *undecidability* results. It is this fact, lurking as a backdrop to all the theorems in this essay, that makes it possible to claim that *Computable Macroeconomics* is not as determinate as Newclassical Macroeconomics. This is also the reason why the Swedes, again Lindahl and Myrdal in particular, were able to work with concepts that were, ostensibly, similar to those being used by the Newclassicals, but were actively engaged in proposing and devising enlightened, discretionary, policies at the macroeconomic level. To be categorical about policy – positively or negatively – on the basis of mathematical models is a dangerous sport.

This chapter is organised as follows. In the next section I outline the origins of the rational expectations problem as a (topological) fixed-point problem. Next, in the third section, I suggest its reformulation in recursion theoretic terms. This reformulation makes it possible to re-interpret a rational expectations equilibrium as a recursion theoretic fixed-point problem in such a way that it is intrinsically computable *ab initio*. Thus, there is no separation between a first step in which the existence of a rational expectations equilibrium is 'proved' and, then, an *ad hoc* mechanism devised to determine it – via uncomputable, equally *ad hoc* learning processes. Moreover, every recursion theoretic assumption, and their consequent formalisms I have employed or invoked, in this chapter, is consistent with the known results and constraints on human cognitive structures and all known computing devices, artificial or natural, ideal or less-than-ideal.

In the fourth section, respecting existing tradition, I accept any given *REE* solution from some, prior, economic model or analysis – in the particular case considered it is a standard OLG generated *REE* solution – and devise a recursion theoretic learning mechanism to determine it.

In the concluding section I try to fashion a fabric, or at least its design, from the sketches of the threads outlined earlier, that depicts a possible research program on Computable Macroeconomics as an alternative to the Newclassical Recursive Macroeconomics.

2 Topological rational expectations

> We can now clearly see the unavoidable dilemma we are facing if we want to apply the Brouwer theorem in the present situation: *if we restrict ourselves to a discrete variable, i.e. consider the reaction func-*

> *tion f merely as a mapping of P into P* [the discrete (finite) set of percentages] *we are not entitled to use the Brouwer theorem because of the non-convexity of P. Besides, continuity represents a vacuous condition in this case. On the other hand, if we use a continuous variable we can use the Brouwer theorem, but the fixed point is then generally located outside of P and hence meaningless in the empirical situation at hand.*
>
> [2], p. 330; italics in the original

In a critical discussion of the use of the Brouwer fixed-point theorem by Herbert Simon [19], that presaged its decisive use in what became the definition of a rational expectations equilibrium, Karl Egil Aubert, a respected mathematician, suggested that economists – and political scientists – were rather cavalier about the domain of definition of economic variables and, hence, less than careful about the mathematics they invoked to derive economic propositions. I was left with the impression, after a careful reading of the discussion between Aubert and Simon ([2], [20], [3], [21]), that the issue was not the use of a fixed point framework but its *nature*, *scope* and *underpinnings*. However, particularly in a rational expectations context, it is not only a question of the nature of the domain of definition but also the fact that there are self-referential and infinite-regress elements intrinsic to the problem. This makes the choice of the fixed point theorem within which to embed the question of a rational expectations equilibrium particularly sensitive to the kind of mathematics and logic that underpins it. In this section I trace the origins of the 'topologisation' of the mathematical problem of rational expectations equilibrium and discuss the possible infelicities inherent in such a formalisation.

There are two crucial aspects to the notion of rational expectations equilibrium – henceforth, *REE* – ([18], pp. 6–10): an individual optimisation problem, subject to *perceived constraints*, and a system wide, autonomous, set of constraints imposing a consistency across the collection of the perceived constraints of the individuals. The latter would be, in a most general sense, the accounting constraint, generated autonomously, by the logic of the macroeconomic system. In a representative agent framework the determination of *REE*s entails the solution of a general fixed-point problem. Suppose the representative agent's perceived law of motion of the macroeconomic system (as a function of state variables and exogenous 'disturbances') as a whole is given by H.[3] The system wide autonomous set of constraints, implied, partially at least, by the optimal decisions based on perceived constraints by the agents, on the other hand, imply an actual law of motion given by, say, H^*. The search for fixed-points of a mapping, T, linking the individually perceived macroeconomic law of motion and the actual law of motion:

$$H^* = T(H) \tag{1}$$

as the fixed points of H of T:[4]

$$H = T(H) \tag{2}$$

determines *REE*s.

What is the justification for T? What kind of 'animal' is it? It is variously referred to as a 'reaction function', a 'best response function', a 'best response mapping', etc. But whatever it is called, eventually the necessary mathematical assumptions are imputed to it such that it is amenable to a topological interpretation whereby appeal can be made to the existence of a fixed point for it as a mapping from a structured domain into itself. So far as I know, there is no optimising economic theoretical justification for it.

There is also a methodological asymmetry in the determination of H and H^*, respectively. The former has a self-referential aspect to it; the latter an infinite regress element in it. Transforming, mechanically, (1) into (2) hides this fact and reducing it to a topological fixed-point problem does little methodological justice to the contents of the constituent elements of the problem. These elements are brought to the surface at the second, separate, step in which ostensible learning mechanisms are devised, in *ad hoc* ways, to determine, explicitly the uncomputable and non-constructive fixed points. But is it really impossible to consider the twin problems in one fell swoop, so to speak?

This kind of tradition to the formalisation and determination of *REE*s has almost by default forced the problem into a particular mathematical straitjacket. The mapping is given topological underpinnings, automatically endowing the underlying assumptions with real analytic content.[5] As a consequence of these default ideas the problem of determining any *REE* is dichotomised into two sub-problems: a first part where non-constructive and non-computable proofs of the existence of *REE*s are provided; and a subsequent, quite separate, second part where mechanisms – often given the sobriquet 'learning mechanisms' – are devised to show that such *REE*s can be determined by individual optimising agents.[6] It is in this second part where orthodox theory endows agents with an *ad hoc* varieties of 'bounded rationality' postulates, without modifying the full rationality postulates of the underlying, original, individual optimisation problem.

Now, how did this *topological fixed-point REE tradition* come into being? Not, as might conceivably be believed, as a result of Muth's justly celebrated original contribution [16], but from the prior work of Herbert Simon on a problem of predicting the behaviour of rational agents in a political setting [19] and an almost simultaneous economic application by Franco Modigliani and Emile Grunberg [10]. Let me explain, albeit briefly, and to the extent necessary in the context of this essay.[7]

Simon, in considering the general issue of the feasibility of public prediction in a social science context, formalised the problem for the particular case of investigating how 'the publication of an election prediction

(particularly one based on poll data) might influence [individual] voting behaviour, and, hence – . . . – falsify the prediction'. Simon, as he has done so often in so many problem situations, came up with the innovative suggestion that the self-referential and infinite-regress content of such a context may well be solved by framing it as a mathematical fixed-point problem:

> Is there not involved here a vicious circle, whereby any attempt to anticipate the reactions of the voters alters those reactions and hence invalidates the prediction?
>
> *In principle*, the last question can be answered in the negative: there is no vicious circle.
>
> . . .
>
> We [can prove using a classical theorem of topology due to Brouwer (the fixed-point theorem)] that it is always possible in principle to take account of reactions to a published prediction in such a way that the prediction will be confirmed by the event.
>
> Simon, op. cit. [19], pp. 82–4; italics added

The 'vicious circle' refers to the self-referential and infinite-regress nature of any such problem where a (rational) agent is placed in a social situation and the individual's behaviour determines, and is determined by, the mutual interdependencies inherent in them. Almost simultaneously with Simon broaching the above problem, Grunberg and Modigliani took up a similar issue within the more specified context of individually rational behaviour in a market economy:[8]

> The fact that human beings react to the expectations of future events seems to create difficulties for the social sciences unknown to the physical sciences: it has been claimed that, *in reacting to the published prediction of a future event, individuals influence the course of events and therefore falsify the prediction.* The purpose of this paper is to verify the validity of this claim.
>
> [10], p. 465; italics added

Grunberg and Modigliani recognised, clearly and explicitly, both the self-referential nature of the problem of consistent individually rational predictions in the face of being placed in an economic environment where their predictions are reactions to, and react upon, the aggregate outcome, but also were acutely aware of the technical difficulties of infinite regression that was also inherent in such situations (cf., in particular [10], pp. 467, 471). In their setting an individual producer faced the classic problem of expected price and quantity formation in a single market, subject to public prediction of the market clearing price. It was not dissimilar to the crude cobweb model, as was indeed recognised by them ([10],

p. 468, footnote 13). Interestingly, what eventually came to be called *rational expectations*, by Muth was called a *warranted expectation*[9] by Grunberg and Modigliani (ibid., pp. 469–70). In any event, their claim that it was '*normally* possible' to prove the *existence* of 'at least one correct public prediction in the face of effective reaction by the agents' was substantiated by invoking Brouwer's fixed-point theorem (ibid., p. 472). To facilitate the application of the theorem, the constituent functions[10] and variables – in particular, the *reaction function* and the conditions on the domain of definition of prices – were assumed to satisfy the necessary real number and topological conditions (continuity, boundedness, etc.).

Thus it was that the tradition, in the rational expectations literature of 'solving' the conundrums of self-reference and infinite-regress via *topological fixed-point theorems* was etched in the collective memory of the profession. And so, four decades after the Simon and the Grunberg–Modigliani contributions, Sargent, in his influential *Arne Ryde Lectures* [18] was able to refer to the fixed-point approach to rational expectations, referring to equation (2), above, without blinking the proverbial eyelid:

> A rational expectations equilibrium is a fixed point of the mapping T.
> [18], p. 10

Now, 50 years after that initial introduction of the topological fixed-point tradition by Simon and Grunberg–Modigliani, economists automatically and uncritically accept that this is the only way to solve the *REE* existence problem – and they are not to be blamed. After all, the same somnambulent complacency dominates the fundamentals of general equilibrium theory, as if the equilibrium existence problem can only be framed as a fixed-point solution. Because of this somnambulent complacency, the existence problem has forever been severed of all connections with the problem of determining – or finding or constructing or locating – the processes that may lead to the non-constructive and uncomputable equilibrium. The recursion theoretic fixed-point tradition not only preserves the unity of equilibrium existence demonstration with the processes that determine it; but it also retains, in the forefront, the self-referential and infinite-regress aspects of the problem of the interaction between individual and social prediction and individual and general equilibrium.

3 Recursion theoretic rational expectations

> Suppose that we want to give an English sentence that commands the reader to print a copy of the same sentence. One way to do so is to say:
>
> Print out this sentence.
>
> This sentence has the desired meaning because it directs the reader

> to print a copy of the sentence itself. However, it doesn't have an obvious translation into a programming language because the **self-referential** word 'this' in the sentence has no counterpart.
>
> . . .
>
> The recursion theorem provides the ability to implement the **self-referential** *this* into any programming language.
>
> [22], p. 200; italics in original;[11] bold added

There is nothing sacrosanct about a topological interpretation of the operator T, the reaction or response function. It could equally well be interpreted recursion theoretically, which is what I shall do in the sequel.[12] I need some unfamiliar, but elementary, formal machinery, not normally available to the mathematical economist or the macroeconomist.

Definition 1 *An **operator** is a function:*

$$\Phi: F_m \to F_n \tag{3}$$

where F_k $(k \geqq 1)$ *is the class of all partial (recursive) functions from* $\mathbb{N}^k$ to $\mathbb{N}$.

Definition 2 Φ *is a **recursive operator** if there is a computable function* ϕ *such that* $\forall f \in F_m$ *and* $\mathbf{x} \in \mathbb{N}^k$, $y \in \mathbb{N}$:

$$\Phi(f)(\mathbf{x}) \simeq y \text{ if } f \exists \text{ a finite } \theta \subseteq f \text{ such that } \phi(\tilde{\theta}, \mathbf{x}) \simeq y$$

where[13] $\tilde{\theta}$ *is a standard coding of a finite function* θ*, which is extended by f.*

Definition 3 *An operator* $\Phi: F_m \to F_n$ *is **continuous** if, for any* $f \in F_m$*, and* $\forall \mathbf{x}$*, y:*

$$\Phi(f)(\mathbf{x}) \simeq y \text{ iff } \exists \text{ a finite } \theta \subseteq f \text{ such that } \Phi(\theta)\ (\mathbf{x}) \simeq y$$

Definition 4 *An operator* $\Phi: F_m \to F_n$ *is **monotone** if, whenever* $f, g \in F_m$ *and* $f \subseteq g$*, then* $\Phi(f) \subseteq \Phi(g)$.

Theorem 5 *A recursive operator is continuous and monotone.*

Example 6 *Consider the following **recursive program**,* P*, (also a recursive operator) over the integers:*

$$P\text{: } F(x, y) \Leftarrow \text{if } x = y \text{ then } y + 1, \text{ else } F(x, F(x - 1, y + 1))$$

Now replace each occurrence of F *in* P *by each of the following functions:*

$$f_1(x, y): \textit{if } x = y \textit{ then } y + 1, \textit{ else } x + 1 \tag{4}$$

$$f_2(x, y): \textit{if } x \geqq y \textit{ then } x + 1, \textit{ else } y - 1 \tag{5}$$

$$f_3(x, y): \textit{if } (x \geqq y) \wedge (x - y \textit{ even}) \textit{ then } x + 1, \textit{ else undefined.} \tag{6}$$

Then, on either side of $\Leftarrow$ *in* $\dot{P}$*, we get the* ***identical*** *partial functions:*

$$\forall i(1 \leqq i \leqq 3), f_i(x, y) \equiv \textit{if } x = y \textit{ then } y = 1, \textit{ else } f_i(x - 1, y + 1) \tag{7}$$

Such functions f_i $(\forall i(1 \leqq i \leqq 3))$ *are referred to as* ***fixed points*** *of the recursive program* $\dot{P}$ *(recursive operator).*

Note that these are fixed points of functionals.

Remark 7 *Note that* f_3*, in contrast to* f_1 *and* f_2*, has the following special property.* $\forall(x, y)$ *of pairs of integers such that* $f_3(x, y)$ *is defined, both* f_1 *and* f_2 *are also defined and have the same value as does* f_3*.*

- f_3 *is, then, said to be* ***less defined than or equal to*** f_1 *and* f_2 *and this property is denoted by* $f_3 \subseteq f_1$ *and* $f_3 \subseteq f_2$*.*
- *In fact, in this particular example, it so happens that* f_3 ***is less defined than or equal*** *to all fixed points of* $\dot{P}$*.*
- *In addition,* f_3 *is the* ***only*** *partial function with this property for* $\dot{P}$ *and is, therefore called the* ***least fixed point of*** $\dot{P}$*.*

We now have all the formal machinery needed to state one of the classic theorems of recursive function theory, known variously as the *first recursion theorem, Kleene's theorem* or, sometimes, as the *fixed point theorem for complete partial orders*.

Theorem 8 *Suppose that* Φ: $F_m \rightarrow F_n$ *is a recursive operator (or a recursive program* $\dot{P}$*). Then there is a partial function* f_ϕ *that is the least fixed point of* Φ:

$$\Phi(f_\phi) = f_\phi;$$
$$\textit{If } \Phi(g) = g, \textit{ then } f_\phi \subseteq g.$$

Remark 9 *If, in addition to being partial,* f_ϕ *is also total, then it is the* ***unique least fixed point****. Note also that a recursive operator is characterised by being continuous and monotone. There would have been some advantages in stating this famous theorem highlighting the domain of definition, i.e., complete partial orders, but the formal machinery becomes slightly unwieldy.*

It is easy to verify that the domain over which the recursive operator and the partial functions are defined are weaker than the conventional

domains over which the economist works. Similarly, the continuity and monotonicity of the recursive operator is naturally satisfied by the standard assumptions in economic theory for the reaction or response function, T. Hence, we can apply the *first recursion theorem* to equation (2), interpreting T as a recursive operator and not as a topological mapping. Then, from the theorem, we know that there is a partial function – i.e., a computable function – f_t that is the least fixed point of T. Stating all this pseudo-formally as a summarising theorem, we get:

Theorem 10 *Suppose that the reaction or response function,* T: $H_m \rightarrow H_n$ *is a recursive operator (or a recursive program* Γ*). Then there is a computable function* f_t *that is a least fixed point of* T:

$T(f_t) = f_t$;
If $T(g) = g$, *then* $f_t \subseteq g$

What are the advantages of recasting the problem of solving for the *REE recursion theoretically* rather than retaining the traditional topological formalisations?

An advantage at the superficial level, but nevertheless important, is the simple fact that, as even the name indicates, recursion encapsulates, explicitly, the idea of self-reference because functions are defined, naturally, in terms of themselves. Second, again at the superficial level, the existence of a least fixed point is a solution to the infinite-regress problem. Thus the two 'birds' are encapsulated in one fell swoop – and, that too, with a computable function. There is, therefore no need to dichotomise the solution for *REE* into an *existence* and a separate *process* or *computable* or *learning* part.

Think of the formal discourse of economic analysis as being conducted in a programming language; call it $\Im$. We know that we choose the underlying terminology for economic formalisms with particular meaning in mind for the elemental units: preferences, endowments, technology, information, expectation and so on; call the generic element of the set ς. When we form a compound economic proposition out of the ς units, the meaning is natural and clear. We can, therefore, suppose that evaluating a compound expression in $\Im$ is immediate: given an expression in $\Im$, say $\lambda(\varsigma)$, the variables in λ, when given specific values α, are to be evaluated according to the *semantics* of $\Im$. To actually *evaluate* a compound expression, $\lambda(\varsigma)$, we write a *recursive program* in the language $\Im$, the language of economic theory. But that leaves a key question unanswered: what is the computable function that is implicitly defined by the recursive program? The first recursion theorem answers this question with the answer: the least fixed-point. In this case, therefore, there is a direct application of the first recursion theorem to the semantics of the language $\Im$. The artificial separation between the syntax of economic analysis, when

formalised, and its natural semantics can, therefore, be bridged *effectively*.

If the language of economic theory is best regarded as a very high level programming language, $\Im$, to understand a *theorem* in economics, in recursion theoretic terms, represent the *assumptions* – i.e., *axioms* and the *variables* – as *input data* and the *conclusions* as *output data*. State the theorem as an expression in the language $\Im$. Then try to convert the proof into a program in the language $\Im$, which will take in the inputs and produce the desired output. If one is unable to do this, it is probably because the proof relies essentially on some infusion of non-constructive or uncomputable elements. This step will identify any inadvertent infusion of non-algorithmic reasoning, which will have to be resolved – sooner or later, if computations are to be performed on the variables as input data. The computations are not necessarily numerical; they can also be symbolic.

In other words, if we take algorithms and data structures to be fundamental, then it is natural to define and understand functions in these terms. If a function does not correspond to an algorithm, what can it be? The topological definition of a function is not algorithmic. Therefore, the expressions formed from the language of economic theory, in a topological formalisation, are not necessarily implementable by a program, except by fluke or by illegitimate and vague approximations. Hence the need to dichotomise every topological existence proof. In the case of *REE*, this is the root cause of the artificial importance granted to a separate problem of learning *REE*s. Nevertheless, the separation does exist and I shall approach a resolution of it in recursion theoretic terms in the next section.

4 Recursively learning a rational expectations equilibrium[14]

> The development of computable analysis as an alternative to conventional mathematical analysis was essentially complete by 1975, although today this analysis is largely unknown.
>
> A perfectly natural reaction at this point is to ask 'Why bother? Calculus has been in use now for over three centuries, and what possible reason is there for altering its rules?' The simplest answer to this question must be ... [that] we still do not solve our mathematical problems precisely. [In computable analysis] ... the key mathematical concepts – the real numbers, sequences, functions and so on – are defined in terms of some computation that an ideal computer could perform.
>
> [1], pp. 2–3

In the previous section I took as given by a previous economic analysis the arguments in the operator T. In this section I go behind the scenes, so to speak, and take one of the many possible economic worlds on which T

operates, a simple Overlapping Generation Model (OLG), with standard assumptions, which generates *REE*s as solutions to the following type of *functional* dynamic equation (cf. [4], pp. 414–16):

$$u'(e_1 - m_t) = \mathrm{E}\left\{\frac{m_{t+1}}{m_t}\frac{L_{t+1}}{L_t}v'\left(e_2 + m_{t+1}\frac{L_{t+1}}{L_t}\right) \mid \mathbf{I}_t\right\}, \forall \mathbf{I}_t \tag{8}$$

Where:
u and v are functional notations for the additive utility functions;

The real gross yield on money, $R_t = \dfrac{p_t x_{t+1}}{p_{t+1}} = \dfrac{m_{t+1}}{m_t}\dfrac{L_{t+1}}{L_t}$;

The real per capita currency balances, $m_t = \dfrac{M_t}{p_t L_t}$;

x_t: discrete random variable;
L_t: size of generation t (a discrete random variable with standard assumptions);
M_t: aggregate stock of currency;
p_t: realised price (of the one consumption good);
p_{t+1}: future price (random variable);
e_t: endowment at time t;
$\mathbf{I}_t$: information set defined by

$$\mathbf{I}_t = I\{\mathbf{I}_{t-1}, L_{t-1}, x_{t-1}, p_{t-1}, \theta_t\} \tag{9}$$

θ_t: vector of all other residual variables that the agent believes will influence future prices.

The problem I pose is the devising of an *effective* mechanism to *learn* and *identify* the above *REE* solution, without asking *how* the solution was arrived at – it could have been arrived at by *magic*, by pronouncements by the *Delphic Oracle*, *prayers*, *torture* or whatever. However, it is immediately clear that one must first ensure that the solution is itself a *recursive real*, if an *effective* mechanism is to locate it. *A priori*, and except for flukes, it is most likely that the standard solution will be a *non-recursive real*. To make it possible, therefore, to ensure a recursively real solution to the above functional dynamic equation, this OLG structure must be endowed with an appropriate recursion theoretic basis. I shall, now, indicate a possible set of minimum requirements for the required recursion theoretic basis.

The derivative of the second period component of the additive utility function, v, must be a *computable real function*. Roughly speaking, if the domain of v is chosen judiciously and if $v \in C^2$, and computable, then v' is computable. But, for these to be acceptable assumptions, the arguments of v', i.e., e_2, m_{t+1}, and $\dfrac{L_{t+1}}{L_t}$, must be computable reals. Since this is

straight-forward for e_2 and per capita currency balances,[15] m_{t+1}, a recursion theoretic interpretation for the random variable L_t will ensure that the assumptions underling v' are recursion theoretically sound. Now, the random variables in the OLG model above are characterised by finite means and stationary probability distributions. It is, therefore, easy to construct a *Probabilistic Turing Machine* (*PTM*), endowed with an extra random-bit generator which outputs, whenever necessary, the necessary element that has the pre-assigned probability distribution. Next, there is the question of the recursivity of the information set, $\mathbf{I}_t$. Given that a recursion theoretic learning model requires this information set to be *recursively* presented to the agents, it is only the element θ_t that remains to be recursively defined. However, this is a purely exogenous variable that can be endowed with the required recursive structure almost arbitrarily.

Finally, the expectations operator is interpreted as an integration process and, since integration is a computable process, this completes the necessary endowment of the elements of the above OLG model with a sufficient recursive structure to make the *REE* generated by the solution to the functional equation a *recursive real*. The minor caveat 'sufficient recursive structure' is to guard against any misconception that this is the only way to endow the elements of an OLG model as given above with the required assumptions to guarantee the generation of a recursive real as a solution. There are many ways to do so but I have chosen this particular mode because it seems straightforward and simple. Above all, these assumptions do not contradict any of the standard assumptions and can live with almost all of them, with minor and inconsequential modifications.

With this machinery at hand, I can state and prove the following theorem:

Theorem 11 *A unique, recursively real, solution to (8) can be identified as the REE and learned recursively.*

Proof. See [25], pp. 98–9. ■

Remark 12 *The theorem is about recursive learning; nevertheless it does embody an unpleasant epistemological implication: there is no* ***effective*** *way for the learning agent to know when to stop applying the learning mechanism!*

Remark 13 *Nothing in the assumptions guarantee tractable computability at any stage.*

5 Recursive reflections

> I went out to take a walk and to recollect after dinner. I did not want to determine a route for my stroll; I tried to attain a maximum latitude

> of probabilities in order not to fatigue my expectation with the necessary foresight of any one of them. I managed, to the imperfect degree of possibility, to do what is called walking at random; I accepted, with no other conscious prejudice than that of avoiding the wider avenues or streets, the most obscure invitations of chance.... My progress brought me to a corner. I breathed in the night, in a most serene holiday from thought.
>
> Borges: *A New Refutation of Time*, in [5], pp. 225–6

In recent years Sargent and his collaborators have developed what they call a *Recursive Macroeconomics* and before that there was the encyclopedic treatise by Lucas and Stokey (with Prescott) on *Recursive Methods in Economic Dynamics* ([14], [23]). *Recursive Macroeconomic Theory*, as Sargent *et al.* see it, is *recursive* in view of the three basic theoretical technologies that underpin the economic hypotheses: *sequential analysis*, *dynamic programming* and *optimal filtering*. To put it in terms of the pioneers whose theories underpin *Recursive Macroeconomic Theory*, the core of this approach harnesses the theoretical technologies of Abraham Wald's sequential analysis, Richard Bellman's dynamic programming and Rudolf Kalman's filtering frameworks. This means, the underlying economic hypotheses of *Recursive Macroeconomic Theory* will be framed and formalised in such a way as to be based on the mathematics of sequential analysis, dynamic programming and optimal filtering – whether or not economic reality demands it; whether or not economic behaviour warrants it; whether or not economic institutions justify it; and most basically, whether or not economic data conform to their requirements.

The word *recursive* is heavily loaded with connotations of dynamics, computation and numerical methods. But these connotations are also fraught with dangers. For example the methods of dynamic programming are provably complex in a precise sense; the equations that have to be solved to implement optimal filtering solutions are also provably intractable; ditto for sequential analysis.

The *recursive theoretic* framework for rational expectations equilibria that I have suggested in the main part of this essay is explicitly computational, algorithmically dynamic and meaningfully numerical. Moreover, the theorems that I have derived above, have an open-ended character about them. To put it in blunt words, these theorems tell an implementable story about things that can be done; but they are silent about things that cannot be done.[16] But the stories are always about what can be done with well defined methods to do them – the algorithms. They are never about pseudo-recursive operators that are somnambulatory with regard to computations and numerical methods.

The two exercises presented in the third and fourth sections of this chapter are prototypes of a strategy to be applied to defining areas of macroeconomics: growth, fluctuations, policy, capital, monetary and

unemployment theories. The general idea is to strip the formal models in the respective fields of their topological underpinnings and replace them, systematically, with recursion theoretic elements in such a way that the open-endedness is enhanced and the numerical and computational contents made explicit and implementable. The specific way it was done in section 3 was to concentrate on the use of the topological fixed-point theorem and replace it with a recursion theoretic fixed-point theorem. Similarly, in the case, of growth theory, say of the von Neumann variety, an analogous exercise can be carried out. This will lead to the use of the *second recursion theorem* rather than the one I have harnessed in this chapter and growth will mean self-reconstruction and self-reproduction. In the case of fluctuations, the idea would be to replace all reliance on differential or difference equation modelling of economic dynamics and replace them with naturally recursion theoretic entities such as *cellular automata*.[17] The aim, ultimately, is to produce a corpus of theories of the central macroeconomic issues so that they can be collected under the alternative umbrella phrase: *Computable Macroeconomics*.

The question will be asked, quite legitimately, whether this line of attack aims also to maintain fidelity with microeconomic, rationality, postulates and, if so, in what way it will differ in the foundations from, say, *Recursive Macroeconomic Theory*. The canonical workhorse on which *Recursive Macroeconomic Theory* rides is the *competitive equilibrium model of a dynamic stochastic economy*. A rational agent in such an economic environment is, essentially, a *signal processor*. Hence, *optimal filtering* plays a pivotal role in this approach to macroeconomic theory. The simple answer, as a *Computable Macroeconomist*, would be that the rational agent of microeconomics would be reinterpreted as a *Turing Machine* – a construction I have developed in great detail in, for example [25], chapter 3. The analogous construction for the other side of the market is equally feasible, starting from re-interpreting the production function as a Turing Machine. This endows the production process with the natural dynamics that belonged to it in the hands of the classical economists and the early Austrians but was diluted by the latter-day Newclassicals. What of market structure – i.e., *economic institutions*? Here, too, following in the giant footsteps of Simon and Scarf, there is a path laid out whereby an algorithmic interpretation of institutions is formally natural.

That leaves only, almost, that sacrosanct disciplining rule of economic theory: optimisation. Recursion theoretic problem formulations eschew optimisations and replace them with *decision problems*. Simply stated, one asks whether problems have solutions or not and if they do, how hard they are and if they do not how must one change the problem formulation to make them solvable. *Decidability*, *solvability* and *computability* are the touchstones of a modelling strategy in *Computable Macroeconomics*. I am reminded, once again, as I conclude, of the early Wittgenstein's poignant observations ([27], §6.51):

> For doubt can exist only where a question exists, a question only where an answer exists, and an answer only where something *can be said*.

Notes

* It was Jean-Paul Fitoussi who first introduced me to, and then educated me on, the vast, weird and complex world of *rational expectations* economics, almost a quarter of a century ago, within the salubrious shades of the *Badia Fiesolana* and its magnificent magnolias, laurels and olive groves, in Fiesole. *That* is the length of time over which a friendship has developed and matured.

1 Keeping in mind Samuelson's admonition that *Mathematics is a language* ([17], epigraph on the title page).

2 One of which is also called a *fixed-point theorem*.

3 Readers familiar with the literature will recognise that the notation H reflects the fact that, in the underlying optimisation problem, a *Hamiltonian* function has to be formed.

4 In a space of functions.

5 In the strict technical sense of the mathematics of *real analysis* as distinct from, say, *constructive*, *computable* or *non-standard analysis*.

6 A perceptive (*sic*!) reader may wonder whether there should not also be an optimisation exercise over the set of feasible or perceived learning mechanisms? Carried to its logical conclusion, this would entail the determination of a set of *REE*s over the collection of learning mechanisms, *ad infintium* (or *ad nauseum*, whichever one prefers).

7 My aim is to show that the framing of the *REE* problem as a *topological fixed-point problem* was not necessary. Moreover, by forcing the *REE* problem as a topological fixed-point problem it was necessary to dichotomise it into the proof of existence part and a separate part to demonstrate the feasibility of constructing mechanisms to determine them. This is mainly – but not only – due to the utilisation of non-constructive topological fixed-point theorems in the first, 'proof of *REE* existence', part. In this sense the *REE learning* research program is very similar to the earlier dichotomising of the *general equilibrium* problem. In that earlier phase, a long tradition of using topological fixed-point theorems to prove the existence of economic equilibria was separated from devising *constructive* or *computable* mechanisms to determine them. The later phase resulted in the highly successful *Computable General Equilibrium* (*CGE*) models. It remains a melancholy fact, however, that even after over 40 years of sustained and impressive work on CGE models, they are neither constructive nor computable, contrary to assertions by proponents of the theory.

8 In their first footnote, Modigliani and Grunberg pay handsome acknowledgement to Herbert Simon for, in particular, suggesting 'the use of Brouwer's fixed-point theorem' ([10], p. 465, footnote 1).

Simon himself later, during the 'debate' with Aubert, on the appropriateness of the use of the Brouwer fixed-point theorem in economic contexts, recalled:

> More recently, the question of the self-consistency of predictions has arisen again in connection with the so-called rational expectations theories of economic behavior under uncertainty.... John Muth's important 1961 paper, which introduced the rational expectations theory, acknowledged the Grunberg–Modigliani paper as a direct ancestor.

> It was the purpose of my paper, and that of Grunberg and Modigliani, to demonstrate that it was always *in principle* possible to anticipate the reaction in the forecast, however difficult it may be to make the correct forecast.
>
> [19], p. 608; italics in original

9 I am reminded that Phelps, in one of his early, influential, papers that introduced the concept of the natural rate of unemployment in its modern forms, first referred to it as a *warranted rate*. Eventually, of course, the Wicksellian term *natural rate*, introduced by Friedman, prevailed. Phelps and Grunberg–Modigliani were, presumably, influenced by Harrodian thoughts in choosing the eminently suitable word 'warranted' rather than 'natural' or 'rational', respectively. Personally, for aesthetic as well as reasons of economic content, I wish the Phelps and Grunberg–Modigliani suggestions had prevailed.

10 The relation between a market price and its predicted value was termed the *reaction function*: 'Relations of this form between the variable to be predicted and the prediction will be called *reaction functions*' ([10], p. 471; italics in original).

As became the tradition in the whole rational expectations literature, the functional form for the reaction functions were chosen with a clear eye on the requirements for the application of an appropriate topological fixed-point theorem. The self-reference and infinite-regress underpinnings were thought to have been adequately subsumed in the existence results that were guaranteed by the fixed-point solution. That the twin conundrums were not subsumed but simply camouflaged was not to become evident till all the later activity on trying to devise learning processes for identifying *REE*s.

11 Sipser's reference is to what is called the 'Second Recursion Theorem'. I shall be working with and appealing to the 'First Recursion Theorem'. But, of course, they are related. I want to work, explicitly, with a space of functions as the domain or relevance, i.e., with functionals, because the economic setting is dynamic. In the static economic case, it would have been sufficient to work with the 'Second Recursion Theorem'.

12 I have relied on the following four excellent texts for the formalisms and results of recursion theory that I am using in this part of the chapter: [6], [7], [15], [24].

13 If $f(\mathbf{x})$ and $g(\mathbf{x})$ are expressions involving the variables $\mathbf{x} = (x_1, x_2, \ldots, x_k)$, then:

$$f(\mathbf{x}) \simeq g(\mathbf{x})$$

means: for any $\mathbf{x}$, $f(\mathbf{x})$ and $g(\mathbf{x})$ are either both defined or undefined, and if defined, they are equal.

14 The model and results of this section are an abbreviation, with minor modifications, of what was presented in [25], pp. 94–100.

15 Provided we assume a straightforward recursive structure for prices, which turns out, usually, to be natural.

16 I cannot resist recalling those famous 'last lines' of the early Wittgenstein:

> What we cannot speak about we must pass over in silence.
>
> ([27], §7)

The sense in which this famous aphorism comes to mind is that in the recursion theoretic approach one does not invoke magic, metaphysics or other formal or informal tricks to solve equations. A problem is always posed in a specific context of *effective* methods of solution. The formal mathematical

approach in standard economic theory is replete with magical and metaphysical methods to 'solve', 'prove' or determine solutions, equilibria, etc.

17 There is more to this suggestion than can be discussed here. It has to do with the connections between dynamical systems theory, numerical analysis and recursion theory, if digital computers are the vehicles for experimental and simulation exercises. If, on the other hand, one is prepared to work with special purpose analogue computers, then the connection between dynamical systems and recursion theory can be more direct and it may not be necessary to eschew the use of differential or difference equations in investigating and modelling economic dynamics. I have discussed these issues in [26].

References

[1] Aberth, Oliver (2001) *Computable Calculus*, Academic Press, San Diego, California.

[2] Aubert, Karl Egil (1982) 'Accurate Predictions and Fixed Point Theorems', *Social Science Information* 21 (3): 323–48.

[3] Aubert, Karl Egil (1982a) 'Accurate Predictions and Fixed Point Theorems: A Reply to Simon', *Social Science Information* 21 (4/5): 612–22.

[4] Azariadis, Costas (1993) *Intertemporal Macroeconomics*, Blackwell Publishers, Oxford.

[5] Borges, Jorge Luis (1964) *Labyrinths: Selected Stories and Other Writings*, New Directions Publishing Corporation, New York.

[6] Cutland, Nigel J. (1980) *Computability: An Introduction to Recursive Function Theory*, Cambridge University Press, Cambridge.

[7] Davis, Martin D., Ron Sigal and Elaine J. Weyuker (1994) *Computability, Complexity and Languages: Fundamentals of Theoretical Computer Science*, Second Edition, Academic Press, London.

[8] Fitoussi, Jean-Paul (1983) 'Modern Macroeconomic Theory: An Overview', in: *Modern Macroeconomic Theory* edited by Jean-Paul Fitoussi, Basil Blackwell, Oxford: Chapter 1, pp. 1–46.

[9] Fitoussi, Jean-Paul and Kumaraswamy Velupillai (1993) 'Macroeconomic Perspectives', in: *Monetary Theory and Thought* edited by Haim Barkai, Stanley Fischer and Nissan Liviatan, The Macmillan Press Ltd, London: Chapter 10, pp. 210–39.

[10] Grunberg, Emile and Franco Modigliani (1954) 'The Predictability of Social Events', *The Journal of Political Economy* LXII (6), December: 465–78.

[11] Hicks, John R. (1977) 'Final Discussion', in: *The Microeconomic Foundations of Macroeconomics – Proceedings of a Conference held by the International Economic Association at S'Agaro, Spain*, edited by G.C. Harcourt, The Macmillan Press Ltd, London: Chapter 12, pp. 376–96.

[12] Keynes, John Maynard (1930) *A Treatise on Money – Volume 1: The Pure Theory of Money*, Macmillan and Co. Limited, London.

[13] Lindahl, Erik (1954) 'On Keynes' Economic System', *Economic Record* 30, May: 19–32.

[14] Ljungqvist, Lars and Thomas J. Sargent (2000) *Recursive Macroeconomic Theory*, The MIT Press, Cambridge, Massachusetts.

[15] Manna, Zohar (1974) *Mathematical Theory of Computation*, McGraw-Hill Kogakusha Ltd, Tokyo.

[16] Muth, John F. (1961) 'Rational Expectations and the Theory of Price Movements', *Econometrica* 29 (6), July: 315–35.
[17] Samuelson, Paul Anthony (1947) *Foundations of Economic Analysis*, Harvard University Press, Cambridge, Massachusetts.
[18] Sargent, Thomas J. (1993) *Bounded Rationality in Macroeconomics*, Clarendon Press, Oxford.
[19] Simon, Herbert (1954) 'Bandwagon and Underdog Effects of Election Predictions', in: *Models of Man – Social and Rational*, John Wiley & Sons Inc., Publishers, New York.
[20] Simon, Herbert (1982) '"Accurate Predictions and Fixed Point Theorems": Comments', *Social Science Information* 21 (4/5): 605–26.
[21] Simon, Herbert (1982a) 'Final Comment', *Social Science Information* 21 (4/5): 622–4.
[22] Sipser, Michael (1997) *Introduction to the Theory of Computation*, PWS Publishing Company, Boston, Massachusetts.
[23] Stokey, Nancy L., Robert E. Lucas Jr, with Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, Massachusetts.
[24] Tourlakis, George J. (1984) *Computability*, Reston Publishing Company, Inc., Reston, Virginia.
[25] Velupillai Kumaraswamy (2000) *Computable Economics*, Oxford University Press, Oxford.
[26] Velupillai Kumaraswamy (2003) 'Economic Dynamics and Computation – Resurrecting the Icarus Tradition', forthcoming in: *Metroeconomica – Special Issue on Computability, Constructivity and Complexity in Economic Theory*.
[27] Wittgenstein, Ludwig (1921 [1961]) *Tractatus Logico-Philosophicus*, Routledge & Kegan Paul Ltd, London.

Author index

Subject index