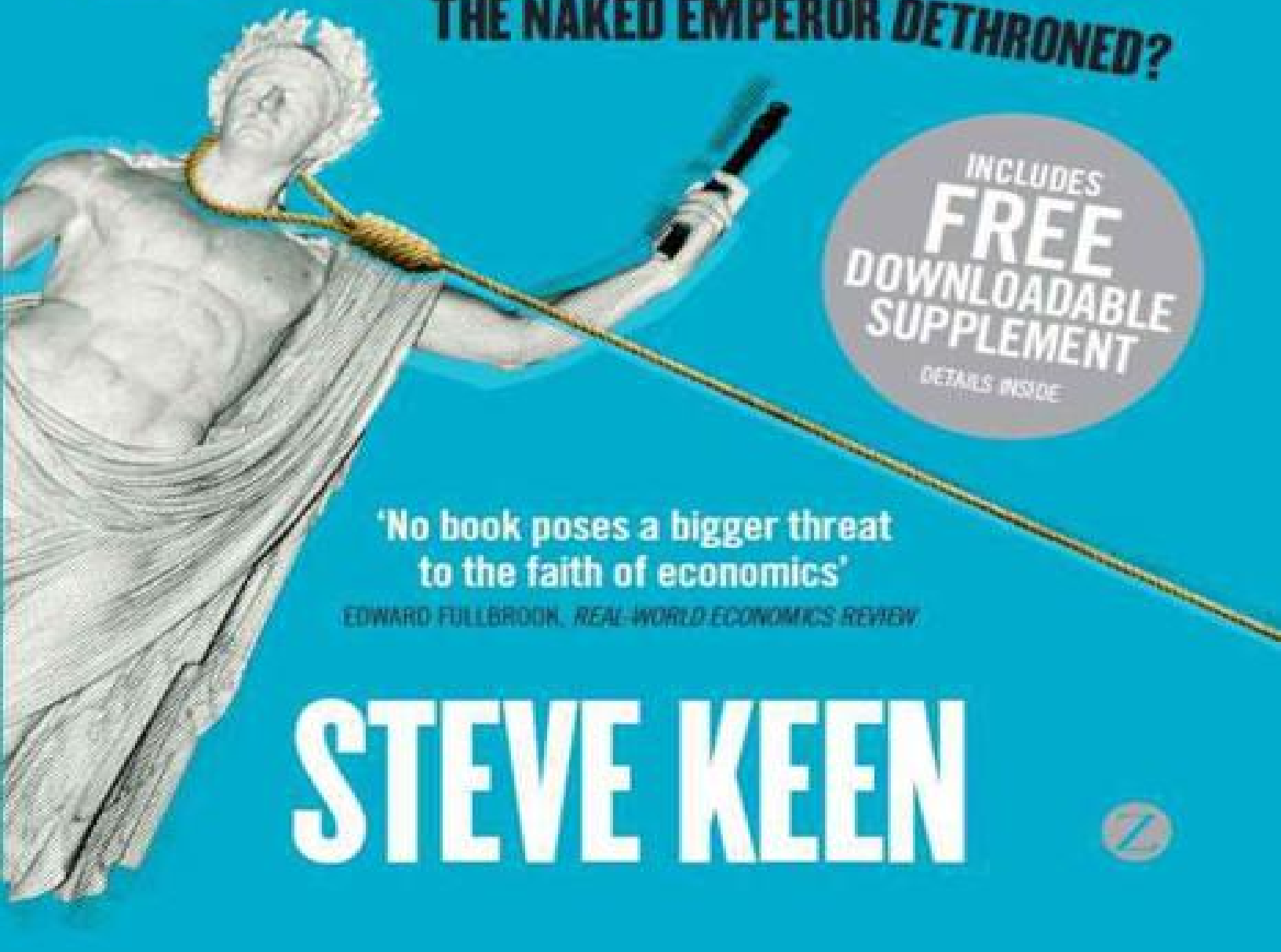
REVISED AND EXPANDED EDITION

DEBUNKING ECONOMICS

THE NAKED EMPEROR DETHRONED?



INCLUDES FREE DOWNLOADABLE SUPPLEMENT

DETAILS INSIDE

'No book poses a bigger threat to the faith of economics'

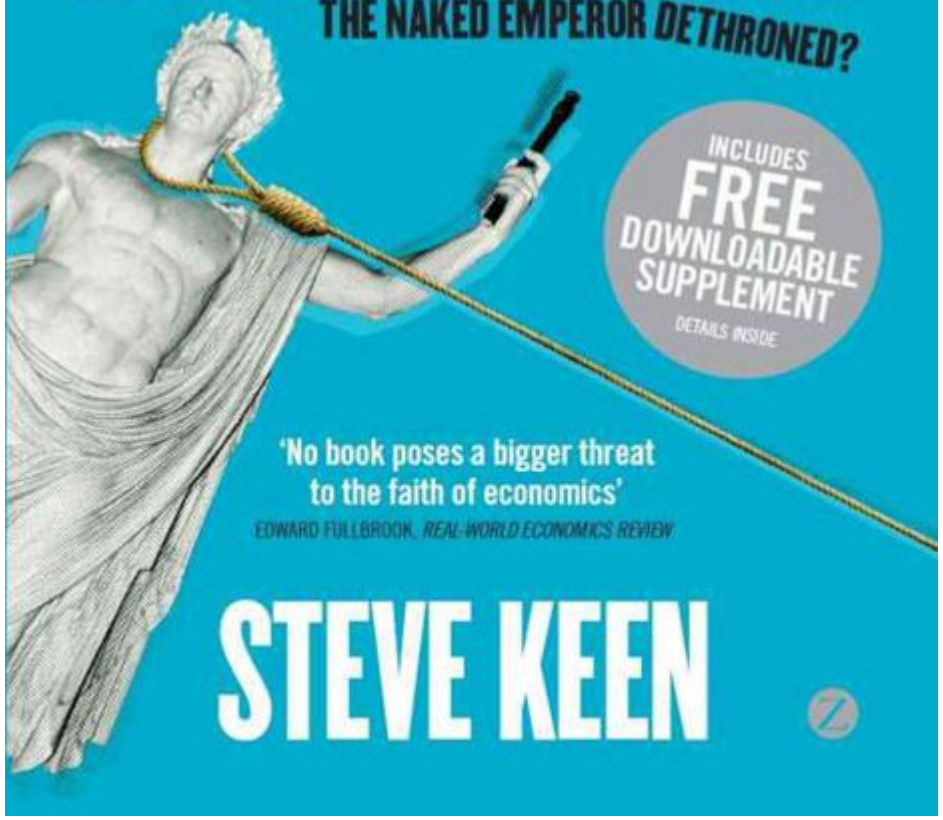
EDWARD FULLBROOK, *REAL-WORLD ECONOMICS REVIEW*

STEVE KEEN

REVISED AND EXPANDED EDITION

DEBUNKING ECONOMICS

THE NAKED EMPEROR *DETHRONED?*



INCLUDES FREE DOWNLOADABLE SUPPLEMENT
DETAILS INSIDE

'No book poses a bigger threat to the faith of economics'
EDWARD FULLBROOK, *REAL-WORLD ECONOMICS REVIEW*

STEVE KEEN

About the author

Steve Keen is professor of economics and finance at the University of Western Sydney. Steve predicted the financial crisis as long ago as December 2005, and warned back in 1995 that a period of apparent stability could merely be 'the calm before the storm'. His leading role as one of the tiny minority of economists to both foresee the crisis and warn of it was recognized by his peers when he received the Revere Award from the *Real-World Economics Review* for being the economist who most cogently warned of the crisis, and whose work is most likely to prevent future crises.

Praise for the first edition

'Professional economists include their own best critics. Steve Keen is one of the very best … translating the algebra into plain language, he deploys a devastating theoretical attack on neoclassical theory.' Hugh Stretton, author of *Economics: A New Introduction*

'Keen's serious but accessible look at the shaky logical and mathematical foundations of neoclassical economics will be of great interest to students and open-minded economists alike.' Don Goldstein, Professor of Economics, Allegheny College

'Particularly useful to those, like myself, who are interested in economics but not formally trained in it. *Debunking Economics* reveals that neoclassical economic doctrines are faulty … because the fundamental assumptions from which such doctrines have been derived are less than self-evident.' Henry C.K. Liu, Chairman, Liu Investment Group

'A wide-ranging yet accessible critique of the staples of neoclassical pedagogy.' Alan G. Isaac, Associate Professor of Economics, American University

'Our hope must be that *Debunking Economics* will be read by enough people to prompt reform of our economic thinking and save our endangered societies.' James Cumes, author of *How to Become a Millionaire – without really working*

'*Debunking Economics* … will transform the way economics is taught and thought.' Jan Otto Andersson, Professor of Economics, Åbo Akademi University

DEBUNKING ECONOMICS – REVISED, EXPANDED AND INTEGRATED EDITION

THE NAKED EMPEROR DETHRONED?

Steve Keen

Zed Books
LONDON | NEW YORK

Debunking Economics – Revised, Expanded and Integrated Edition: The Naked Emperor Dethroned? was first published in 2011 by Zed Books Ltd, 7 Cynthia Street, London N1 9JF, UK and Room 400, 175 Fifth Avenue, New York, NY 10010, USA

The first edition of *Debunking Economics* was first published in the United Kingdom and the United States of America in 2001 by Zed Books Ltd, and in Australia by Pluto Press Australia Ltd.

www.zedbooks.co.uk
www.debunkingeconomics.com
www.debtdeflation.com

Copyright © Steve Keen 2011

The right of Steve Keen to be identified as the author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act, 1988

Set in Monotype Plantin and FontFont Kievit by Ewan Smith, London
Index: ed.emery@thefreeuniversity.net
Cover designed by Rogue Four Design

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior permission of Zed Books Ltd.

A catalogue record for this book is available from the British Library
Library of Congress Cataloging in Publication Data available

eISBN 9781780322209

CONTENTS

Tables, figures and boxes
Preface to the second edition
Preface to the first edition

1 Predicting the ‘unpredictable’
2 No more Mr Nice Guy

Part 1 Foundations: the logical flaws in the key concepts of conventional economics

3 The calculus of hedonism
4 Size does matter
5 The price of everything and the value of nothing
6 To each according to his contribution

Part 2 Complexities: issues omitted from standard courses that should be part of an education in economics

7 The holy war over capital
8 There is madness in their method
9 Let’s do the Time Warp again
10 Why they didn’t see it coming
11 The price is not right
12 Misunderstanding the Great Depression and the Great Recession

Part 3 Alternatives: different ways to think about economics

13 Why I did see ‘It’ coming
14 A monetary model of capitalism
15 Why stock markets crash

16 Don’t shoot me, I’m only the piano
17 Nothing to lose but their minds
18 There are alternatives

Bibliography
Index

TABLES, FIGURES AND BOXES

Tables

2.1 Anticipations of the housing crisis and recession
3.1 'Utils' and change in utils from consuming bananas
3.2 Utils arising from the consumption of two commodities
3.3 The commodities in Sippel's 'Revealed Preference' experiment
4.1 Demand schedule for a hypothetical monopoly
4.2 Costs for a hypothetical monopoly
4.3 Sales and costs determine the level of output that maximizes profit
4.4 Cost and revenue for a 'perfectly competitive' industry identical in scale to hypothetical monopoly
5.1 Input and output data for a hypothetical firm
5.2 Cost drawings for the survey by Eiteman and Guthrie
5.3 Empirical research on the nature of cost curves
7.1 Sraffa's hypothetical subsistence economy
7.2 Production with a surplus
7.3 Relationship between maximum and actual rate of profit and the wage share of surplus
7.4 The impact of the rate of profit on the measurement of capital
10.1 Anderson's ranking of sciences
12.1 The alleged Money Multiplier process
13.1 A hypothetical example of the impact of decelerating debt on aggregate demand
13.2 The actual impact of decelerating debt on aggregate demand
14.1 A pure credit economy with paper money
14.2 The dynamics of a pure credit economy with no growth
14.3 Net incomes

14.4 A growing pure credit economy with electronic money
15.1 Von Neumann's procedure for working out a numerical value for utility
15.2 The Allais 'Paradox'
15.3 The Allais 'Paradox' Part 2
16.1 The solvability of mathematical models
17.1 Marx's unadjusted value creation table, with the rate of profit dependent upon the variable-to-constant ratio in each sector
17.2 Marx's profit distribution table, with the rate of profit now uniform across sectors
17.3 Steedman's hypothetical economy
17.4 Steedman's physical table in Marx's value terms
17.5 Steedman's prices table in Marx's terms
17.6 Profit rate and prices calculated directly from output/wage data
17.7 Marx's example where the use-value of machinery exceeds its depreciation

Figures

2.1 US inflation and unemployment from 1955
2.2 Bernanke doubles base money in five months
2.3 Private debt peaked at 1.7 times the 1930 level in 2009
3.1 Rising total utils and falling marginal utils from consuming one commodity
3.2 Total utils from the consumption of two commodities;
3.3 Total 'utils' represented as a 'utility hill'
3.4 The contours of the 'utility hill'
3.5 Indifference curves: the contours of the 'utility hill' shown in two dimensions
3.6 A rational consumer's indifference map
3.7 Indifference curves, the budget constraint, and consumption
3.8 Deriving the demand curve
3.9 Upward-sloping demand curve
3.10 Separating out the substitution effect from the income effect

3.11 Engel curves show how spending patterns change with increases in income
3.12 A valid market demand curve
3.13 Straight-line Engel 'curves'
3.14 Economic theory cannot rule out the possibility that a market demand curve may have a shape like this, rather than a smooth, downward-sloping curve
4.1 Leijonhufvud's 'Totems' of the Econ tribe
4.2 Stigler's proof that the horizontal firm demand curve is a fallacy
4.3 Profit maximization for a monopolist: marginal cost equals marginal revenue, while price exceeds marginal cost
4.4 Profit maximization for a perfectly competitive firm: marginal cost equals marginal revenue, which also equals price
4.5 A supply curve can be derived for a competitive firm, but not for a monopoly
4.6 A competitive industry produces a higher output at a lower cost than a monopoly
4.7 The standard 'supply and demand' explanation for price determination is valid only in perfect competition
4.8 Double the size, double the costs, but four times the output
4.9 Predictions of the models and results at the market level
4.10 Output behavior of three randomly selected firms
4.11 Profit outcomes for three randomly selected firms
4.12 Output levels for between 1- and 100-firm industries
5.1 Product per additional worker falls as the number of workers hired rises
5.2 Swap the axes to graph labor input against quantity
5.3 Multiply labor input by the wage to convert Y-axis into monetary terms, and add the sales revenue
5.4 Maximum profit occurs where the gap between total cost and total revenue is at a maximum
5.5 Deriving marginal cost from total cost
5.6 The whole caboodle: average and marginal costs, and marginal

revenue
5.7 The upward-sloping supply curve is derived by aggregating the marginal cost curves of numerous competitive firms
5.8 Economic theory doesn't work if Sraffa is right
5.9 Multiple demand curves with a broad definition of an industry
5.10 A farmer who behaved as economists advise would forgo the output shown in the gap between the two curves
5.11 Capacity utilization over time in the USA
5.12 Capacity utilization and employment move together
5.13 Costs determine price and demand determines quantity
5.14 A graphical representation of Sraffa's (1926) preferred model of the normal firm
5.15 The economic theory of income distribution argues that the wage equals the marginal product of labor
5.16 Economics has no explanation of wage determination or anything else with constant returns
5.17 Varian's drawing of cost curves in his 'advanced' microeconomics textbook
6.1 The demand for labor curve is the marginal revenue product of labor
6.2 The individual's income–leisure trade-off determines how many hours of labor he supplies
6.3 An upward-sloping individual labor supply curve
6.4 Supply and demand determine the equilibrium wage in the labor market
6.5 Minimum wage laws cause unemployment
6.6 Demand management policies can't shift the supply of or demand for labor
6.7 Indifference curves that result in less work as the wage rises
6.8 Labor supply falls as the wage rises
6.9 An individual labor supply curve derived from extreme and midrange wage levels
6.10 An unstable labor market stabilized by minimum wage legislation

6.11 Interdependence of labor supply and demand via the income distributional effects of wage changes
7.1 The standard economic ‘circular flow’ diagram
7.2 The rate of profit equals the marginal product of capital
7.3 Supply and demand determine the rate of profit
7.4 The wage/profit frontier measured using the standard commodity
9.1 Standard neoclassical comparative statics
9.2 The time path of one variable in the Lorenz model
9.3 Structure behind the chaos
9.4 Sensitive dependence on initial conditions
9.5 Unstable equilibria
9.6 Cycles in employment and income shares
9.7 A closed loop in employment and wages share of output
9.8 Phillips’s functional flow block diagram model of the economy
9.9 The component of Phillips’s Figure 12 including the role of expectations in price setting
9.10 Phillips’s hand drawing of the output–price-change relationship
9.11 A modern flow-chart simulation program generating cycles, not equilibrium
9.12 Phillips’s empirically derived unemployment–money-wage-change relation
10.1 Hicks’s model of Keynes
10.2 Derivation of the downward-sloping IS curve
10.3 Derivation of the upward-sloping LM curve
10.4 ‘Reconciling’ Keynes with ‘the Classics’
10.5 Unemployment–inflation data in the USA, 1960–70
10.6 Unemployment–inflation data in the USA, 1950–72
10.7 Unemployment–inflation data in the USA, 1960–80
10.8 The hog cycle
11.1 Supply and demand in the market for money
11.2 The capital market line
11.3 Investor preferences and the investment opportunity cloud

11.4 Multiple investors (with identical expectations)
11.5 Flattening the IOC
11.6 How the EMH imagines that investors behave
11.7 How speculators actually behave
12.1 Inflation and base money in the 1920s
12.2 Inflation and base money in the post-war period
12.3 Bernanke's massive injection of base money in QE1
12.4 Change in M0 and unemployment, 1920–40
12.5 Change in M1 and unemployment, 1920–40
12.6 Change in M0 and M1, 1920–40
12.7 M0–M1 correlation during the Roaring Twenties
12.8 M0–M1 correlation during the Great Depression
12.9 Bernanke's 'quantitative easing' in historical perspective
12.10 The volume of base money in Bernanke's 'quantitative easing' in historical perspective
12.11 Change in M1 and inflation before and during the Great Recession
12.12 The money supply goes haywire
12.13 Lindsey, Orphanides, Rasche 2005, p. 213
12.14 The empirical 'Money Multiplier', 1920–40
12.15 The empirical 'Money Multiplier', 1960–2012
12.16 The disconnect between private and fiat money during the Great Recession
13.1 Goodwin's growth cycle model
13.2 My 1995 Minsky model
13.3 The vortex of debt in my 1995 Minsky model
13.4 Cyclical stability with a counter-cyclical government sector
13.5 Australia's private debt-to-GDP ratio, 1975–2005
13.6 US private debt to GDP, 1955–2005
13.7 Aggregate demand in the USA, 1965–2015
13.8 US private debt
13.9 The change in debt collapses as the Great Recession begins
13.10 The Dow Jones nosedives

13.11 The correlation of debt-financed demand and unemployment
13.12 The housing bubble bursts
13.13 The Credit Impulse and change in employment
13.14 Correlation of Credit Impulse and change in employment and GDP
13.15 Relatively constant growth in debt
13.16 The biggest collapse in the Credit Impulse ever recorded
13.17 Growing level of debt-financed demand as debt grew faster than GDP
13.18 The two great debt bubbles
13.19 Change in nominal GDP growth then and now
13.20 Real GDP growth then and now
13.21 Inflation then and now
13.22 Unemployment then and now
13.23 Nominal private debt then and now
13.24 Real debt then and now
13.25 Debt to GDP then and now
13.26 Real debt growth then and now
13.27 The collapse of debt-financed demand then and now
13.28 Debt by sector – business debt then, household debt now
13.29 The Credit Impulse then and now
13.30 Debt-financed demand and unemployment, 1920–40
13.31 Debt-financed demand and unemployment, 1990–2011
13.32 Credit Impulse and change in unemployment, 1920–40
13.33 Credit Impulse and change in unemployment, 1990–2010
13.34 The Credit Impulse leads change in unemployment
14.1 The neoclassical model of exchange as barter
14.2 The nature of exchange in the real world
14.3 A nineteenth-century private banknote
14.4 Bank accounts
14.5 A credit crunch causes a fall in deposits and a rise in reserves in the bank's vault
14.6 A bank bailout's impact on loans

14.7 A bank bailout's impact on incomes
14.8 A bank bailout's impact on bank income
14.9 Bank income grows if debt grows more rapidly
14.10 Unemployment is better with a debtor bailout
14.11 Loans grow more with a debtor bailout
14.12 Profits do better with a debtor bailout
14.13 Bank income does better with a bank bailout
14.14 Modeling the Great Moderation and the Great Recession – inflation, unemployment and debt
14.15 The Great Moderation and the Great Recession – actual inflation, unemployment and debt
14.16 Modeling the Great Moderation and the Great Recession – output
14.17 Income distribution – workers pay for the debt
14.18 Actual income distribution matches the model
14.19 Debt and GDP in the model
14.20 Debt and GDP during the Great Depression
15.1 Lemming population as a constant subject to exogenous shocks
15.2 Lemming population as a variable with unstable dynamics
17.1 A graphical representation of Marx's dialectics

Boxes

10.1 The Taylor Rule
13.1 Definitions of unemployment

PREFACE TO THE SECOND EDITION

Debunking Economics was far from the first book to argue that neoclassical economics was fundamentally unsound. If cogent criticism alone could have brought this pseudo-science down, it would have fallen as long ago as 1898, when Thorstein Veblen penned 'Why is economics not an evolutionary science?' (Veblen 1898). Yet in 1999, when I began writing *Debunking Economics*, neoclassical economics was more dominant than it had ever been.

My reason for adding to this litany of thus far unsuccessful attempts to cause a long-overdue scientific revolution in economics was the belief that a prerequisite for success was just around the corner. As I noted in my concluding chapter, I felt that a serious economic crisis was approaching, and that when this crisis hit, fundamental change in economic theory would be possible:

> I am not wishing an economic crisis upon the modern world – instead, I think one has been well and truly put in train by the cumulative processes described in chapters 10 and 11 [on finance]. If that crisis eventuates – one which neoclassical economic theory argues is not possible – then economics will once again come under close and critical scrutiny. (*Debunking Economics*, 1st edn, p. 312)

When I finished *Debunking Economics*, I hoped to be able to start work on a book with the working title of *Finance and Economic Breakdown*, which would have provided a comprehensive theory of the forces that would cause this crisis. Instead, the reaction from neoclassical economists to Chapter 4 of *Debunking Economics* – 'Size does matter', on the neoclassical model of competition – was so vehement that I spent much of

the next four years developing the arguments in that chapter in response to their attacks.

Finally, in December 2005, I returned to writing *Finance and Economic Breakdown* (for Edward Elgar Publishers). Almost immediately, unforeseen circumstances intervened once more, when I was asked to be an expert witness in a predatory lending case. One look at the exponential growth in the debt-to-GDP ratios for Australia and the USA convinced me that a truly huge crisis was imminent.

I decided that raising the public alarm was more important than writing an academic treatise on the topic, so I reluctantly delayed the book once more and turned to the media and the Internet instead. I published a monthly report on debt, starting in November 2006 (Keen 2006), became sufficiently well known in the media to be described as a 'media tart' by some Australian critics, established the blog Debtwatch (www.debtdeflation.com/blogs), which now has over 10,000 registered users and attracts about 50,000 unique readers each month (with about 25,000 of those being Australian, and most of the rest coming from America and the UK), and in what passed for spare time, worked to complete a model of debt deflation to inform my public comments.

The economic crisis began with a vengeance in September 2007. Unemployment in the USA doubled in the next year, while a 5 percent rate of inflation rapidly gave way to 2 percent deflation.

The complete failure of neoclassical economics to anticipate the crisis also meant, as I expected, that economic theory and economists are under public attack as never before. Their defense has been to argue that 'no one could have seen this coming.' They have taken refuge in the phrase that this crisis was a 'Black Swan,' using Nassim Taleb's phrase completely out of context (Taleb 2007), and ignoring the fact that I and many other non-neoclassical economists did in fact see this coming.

I therefore decided that, for both positive and negative reasons, a new edition of *Debunking Economics* was needed.

The negative reason is that there is no better time to attack a fallacious theory than after it has made a spectacularly wrong prediction. By arguing that the macroeconomy had entered a permanent 'Great Moderation' (the phrase Ben Bernanke popularized to describe the apparent reduction in economic volatility and falls in unemployment and inflation between 1975

and 2007), neoclassical economics couldn't have been more wrong about the immediate economic future. Now is the time to show that, not only was this crisis eminently foreseeable, but also neoclassical economists were about the only ones who were ill equipped to see it coming. The main positive reason is that, with the public and policymakers much more amenable to alternative ways of thinking about economics, now is the time to provide a brief and accessible look at an alternative, realistic model of the economy.

There have also been some important developments in economics since the first edition – notably the growth of econophysics, and the concession by finance academics that the Efficient Markets Hypothesis has been empirically disproven (Fama and French 2004).

Several new chapters have been added on the dynamics of debt-based money, and the continuing economic crisis – currently called the Great Recession in America (and the 'Global Financial Crisis' in my home country, Australia), but which I fully expect to be renamed the Second Great Depression by future economic historians. These new chapters 'break the mold' for the rest of the book, in that they are not critiques of the neoclassical theory of financial instability and economic crises – because there simply is no such theory. Instead they set out, in an introductory way, the non-neoclassical theories of debt deflation and endogenous money that I have played a role in developing (Keen 2008, 2009a, 2009b, 2010), and the model of financial instability that I will cover in detail in *Finance and Economic Breakdown*.

I have also edited a number of chapters where there have been significant theoretical developments since the first edition. By far the most important development here has been a substantial deepening of the critique of the theory of the firm in 'Size does matter.' There is also substantially more information on why the theory of demand is false in 'The calculus of hedonism' and 'The price of everything and the value of nothing,' and a record of the recanting of the Efficient Markets Hypothesis by its major advocates Fama and French in the addendum to 'The price is not right.'

Lastly, a book that was in its first incarnation almost exclusively about microeconomics now covers microeconomics and macroeconomics in roughly equal measure.

The one glaring omission is the absence of any discussion of international trade theory. The reason for this is that, while the flaws in the theory of comparative advantage are, to me, both huge and obvious, a detailed critique of the mathematical logic has not yet been done, and nor is there a viable alternative. That is a task that I may tackle after *Finance and Economic Breakdown* is completed, but not before.

Looking back

The reception of the first edition was both gratifying and predictable. The gratifying side was the public reception: sales far exceeded the norm for this class of book, it continued to sell well a decade after it was first published, and the critical response from the public was almost universally positive.

The predictable side was the reaction from neoclassical economists. They disparaged the book in much the way they have treated all critics – as Keynes once remarked, he expected his work to be treated as being both 'quite wrong and nothing new.' My critique received the same treatment, and as well neoclassicals were incensed by my critique of the theory of the firm.

Their rejoinders to that critique led me to develop it far beyond the version first published in 2001, and in ways that I thought would be very difficult to convey without mathematics, but which in fact I found quite easy to explain in the addendum to 'Size does matter.' However, for a detailed treatment mathematics is still necessary, so for those who can cope with the odd – or rather frequent! – equation, the most accessible papers are in the journals (Keen 2003, 2004; Keen and Standish 2006, 2010) and book chapters (Keen 2005, 2009a). The paper in the free online journal *The Real-World Economic Review* is the most easily accessed of these (www.paecon.net/PAEReview/issue53/KeenStandish53.pdf), while my chapter in the book *A Handbook for Heterodox Economics Education* (edited by Jack Reardon and published by Routledge), 'A pluralist approach to microeconomics,' covers the critique of the Marshallian model of the firm in a manner that should be useful to academics and schoolteachers.

Looking forward

I knew when I wrote the first edition of *Debunking Economics* that its real aim – the elimination of neoclassical economics and its replacement by an empirically based, dynamic approach to economics – could not be achieved until a serious economic crisis called into question the Panglossian view of market economies that neoclassical economics promulgates. That crisis is well and truly with us, and the public has turned on economists as I had hoped it would. Unfortunately, the economics profession is also reacting as I expected – by pretending that nothing is wrong.

As I write these words I have just returned from the 2011 American Economic Association (AEA) annual conference, where close to 10,000 mainly US and overwhelmingly neoclassical economists meet every year to present and hear 'the latest' in the profession. Though there were quite a few sessions devoted to the Great Recession and what its implications are for economic theory (mainly organized by non-neoclassical associations within the AEA, such as the Union for Radical Political Economics), the majority of the profession continues to believe, as Ben Bernanke put it some months beforehand, that 'the recent financial crisis was more a failure of economic engineering and economic management than of what I have called economic science' (Bernanke 2010).

Bernanke's belief could not be farther from the truth: as a means to understand the behavior of a complex market economy, the so-called science of economics is a melange of myths that make the ancient Ptolemaic earth-centric view of the solar system look positively sophisticated in comparison. What his opinion reveals is his inability to think about the economy in any way other than the neoclassical one in which he has been trained – an inability he shares with most of his colleagues.

If we leave the development of economics to economists themselves, then it is highly likely that the intellectual revolution that economics desperately needs will never occur – after all, they resisted change so successfully after the Great Depression that the version of neoclassical economics that reigns today is far more extreme than that which Keynes railed against seven decades ago. I concluded the first edition with the observation that economics is too important to leave to the economists. That remains the case today.

If change is going to come, it will be from the young, who have not yet been indoctrinated into a neoclassical way of thinking, and from those from other professions like physics, engineering and biology, who will be emboldened by the crisis to step onto the turf of economics and take the field over from the economists. It is to those real engines of change in economics that this book is dedicated.

PREFACE TO THE FIRST EDITION

In the preface to the *General Theory*, Keynes commented that its writing had involved a long process of escape from 'habitual modes of thought and expression.' He implored his audience of professional economists to likewise escape the confines of conventional economic thought, and observed that 'The ideas which are here expressed so laboriously are extremely simple and should be obvious. The difficulty lies, not in the new ideas, but in escaping from the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds' (Keynes 1936).

This statement was unfortunately prophetic. Keynes's own escape was incomplete, and the residue of traditional thought the *General Theory* contained obscured many of its most innovative aspects. Faced with a melange of the new and unfamiliar with the old and familiar, the bulk of his audience found it easier to interpret his new ideas as no more than embellishments to the old. The Keynesian Revolution died, slowly but surely, as economists reconstructed the 'habitual modes of thought and expression' around the inconvenient intrusions Keynes had made into economic dogma. Economics failed to make the escape which Keynes had implored it to do, and as time went on, 'modern' economics began to resemble more and more closely the 'old ideas' which Keynes had hoped economics would abandon.

I was initially educated in this resurgent tradition – known as the Keynesian-Neoclassical synthesis – some thirty years ago. The catalyst for my escape from this dogma was extremely simple: my first-year microeconomics lecturer pointed out a simple but glaring flaw in the application of conventional theory.

The economic theory of markets argues that combinations of any sort, whether by workers into unions or manufacturers into monopolies, reduce social welfare. The theory therefore leads to the conclusion that the world

would be better off without monopolies and unions. If we were rid of both, then the economic theory of income distribution argues that, effectively, people's incomes would be determined solely by their contribution to society. The world would be both efficient and fair.

But what if you have both monopolies and unions? Will getting rid of just one make the world a better place?

The answer is categorically no. If you abolish just unions, then according to 'conservative' economic theory, workers will be exploited: they will get substantially less than their contribution to society (equally, if you abolish just monopolies, then workers will exploit companies). If you have one, then you are better off having the other too, and a single step towards the economist's nirvana takes you not closer to heaven but towards hell.[1]

I was struck by how fragile the outwardly impregnable theory of economics was. What seemed self-evident at a superficial level – that social welfare would rise if unions or monopolies were abolished – became problematic, and even contradictory, at a deeper level.

Had I come across that fragility in my Honors or postgraduate education, which is when students of economics normally learn of such things, I would quite possibly have been willing to gloss over it, as most economists do. Instead, because I learnt it 'out of sequence,' I was immediately suspicious of the simplistic statements of economic principle. If the pivotal concepts of competition and income distribution could be so easily overturned, what else was rotten in the House of Economics?

That skepticism initiated a gradual process of discovery, which made me realize that what I had initially thought was an education in economics was in fact little better than an indoctrination. More than a decade before I became an undergraduate, a major theoretical battle had broken out over the validity of economic theory. Yet none of this turned up in the standard undergraduate or honors curriculum – unless it was raised by some dissident instructor. There were also entire schools of thought which were antithetical to conventional economics, which again were ignored unless there was a dissident on the staff.

Thirty years after starting my skeptic's intellectual tour, I am completely free of the 'habitual modes of thought and expression' which so troubled

Keynes. There are many non-orthodox economists like me, who are all trying to contribute to a new, deeper approach to economics.

But still the world's universities churn out economists who believe, for example, that the world would be a better place if we could just get rid of unions, or monopolies.

Worse still, over the last thirty years, politicians and bureaucrats the world over have come to regard economic theory as the sole source of wisdom about the manner in which a modern society should be governed. The world has been remade in the economist's image.

This ascendancy of economic theory has not made the world a better place. Instead, it has made an already troubled society worse: more unequal, more unstable, and less 'efficient.'

Why has economics persisted with a theory which has been comprehensively shown to be unsound? Why, despite the destructive impact of economic policies, does economics continue to be the toolkit which politicians and bureaucrats apply to almost all social and economic issues?

The answer lies in the way economics is taught in the world's universities.

When I became an academic economist, I realized that very few of my colleagues had any knowledge of the turbulent streams in economics. Most were simply dismissive of any attempt to criticize orthodox thinking, and equally dismissive of any of their peers who showed tendencies towards unconventional thought.

This was not because these conventional economists were anti-intellectual – far from it. Even though conventional economics is flawed, it still takes intellectual muscle to master its principles – as you will soon discover. Yet still economists refused to consider any criticisms of economic theory, even when they emanated from other economists, and met rigorous intellectual standards.

Nor were they ill intentioned – most of them sincerely believed that, if only people followed the principles of economic theory, the world would be a better place. For a group of people who espoused a philosophy of individualistic hedonism, they were remarkably altruistic in their commitment to what they saw as the common good. Yet the policies they

promoted often seem to non-economists to damage the fabric of human society, rather than to enhance it.

They also rejected out of hand any suggestion that they were ideologically motivated. They were scientists, not political activists. They recommended market solutions, not because they were personally pro-capitalist, but because economic theory proved that the market was the best mechanism by which to determine economic issues. Yet virtually everything they recommended at least appeared to favor rich over poor, capitalist over worker, privileged over dispossessed.

I came to the conclusion that the reason they displayed such anti-intellectual, apparently socially destructive, and apparently ideological behavior lay deeper than any superficial personal pathologies. Instead, the way in which they had been educated had given them the behavioral traits of zealots rather than of dispassionate intellectuals.

As anyone who has tried to banter with an advocate of some esoteric religion knows, there is no point trying to debate fundamental beliefs with a zealot. After many similar experiences with economists, I abandoned any delusion that I might be able to persuade committed economists to see reason (though there has been the odd exception to this rule). Instead, I prefer to spend my time developing an alternative approach to economics, while persuading others not to fall for the superficially persuasive but fundamentally flawed arguments of conventional theory.

Hence this book, which is aimed at a broader audience than Keynes's target of his fellow economists. Instead, my primary target market is those people who feel that they have been effectively silenced by economists. One of the many reasons why economists have succeeded in taking over social policy is that they have claimed the high intellectual ground against anyone who opposed their recommendations. The object of this book is to show that this claim is spurious.

Though I am the sole author, and thus responsible for all its errors and omissions, I cannot claim sole credit for what is good in it. In particular, I owe an enormous debt to the pioneers of critical thinking in economics.

Pre-eminent amongst these is Piero Sraffa – a name which is known to almost no non-economists, and very few economists. There are many others whose names turn up in subsequent pages – Blatt, Garengani, Goodwin,

Kalecki, Kaldor, Keynes, Minsky, Veblen, to name a few. But none has had quite the impact of Sraffa.

I owe a more personal debt to those few teachers who were, as I am now, dissidents in a sea of believers. Pre-eminent here is Frank Stilwell – the first-year lecturer who, many years ago, introduced me to the first of many flaws in conventional economics. I also gratefully acknowledge the influence which Ted Wheelwright's panoptic knowledge of the many currents in economic thought had upon my intellectual development. My colleagues in HETSA, the History of Economic Thought Society of Australia, have also enriched my appreciation of the many 'roads not taken' by mainstream economics.

Colleagues around the world have provided feedback on the arguments presented here. None can be held liable for what follows, but all influenced it, either directly, in debate, or by providing a forum in which heterodox views could flourish. My thanks go to Trond Andresen, George Argyrous, Tony Aspromorgous, Joanne Averill, Aldo Balardini, Bill Barnett, James Dick, Marchessa Dy, Geoff Fishburn, John Gelles, Ric Holt, Julio Huato, Alan Isaac, James Juniper, Gert Kohler, John Legge, Jerry Levy, Henry Liu, Basil Moore, Marc-Andre Pigeon, Clifford Poirot, Jason Potts, Barkley Rosser, Gunnar Tomasson, Sean Toohey, Robert Vienneau, Graham White, and Karl Widerquist, for reading and commenting upon drafts of this book. I would especially like to thank Karl Widerquist for detailed suggestions on content and the flow of arguments, John Legge for assistance with the proofs of some propositions, Alan Isaac for providing a testing foil to many propositions in the early chapters, and Geoff Fishburn for many years of intelligent and critical discussion of economic theory.

Joyce Hitchings provided valuable feedback on how to make the book's arguments and counter-arguments more accessible to readers with no prior training in economics.

I have also received great encouragement and feedback from my publishers Tony Moore of Pluto Press, and Robert Molteno of Zed Books. My editor, Michael Wall, did a sterling job of making the final product more concise and accessible than the original manuscript.

Sabbatical leave granted by the University of Western Sydney gave me the time away from the everyday demands of an academic life needed to complete a book. The Jerome Levy Institute of Bard College, New York,

and the Norwegian University of Science and Technology in Trondheim, Norway, kindly accommodated me while the finishing touches were applied to the manuscript.

And so to battle.

1 | PREDICTING THE 'UNPREDICTABLE'

A major motivation for writing the first edition of this book was my feeling in 2000 that a serious economic crisis was imminent, and that it was therefore an apt time to explain to the wider, non-academic community how economic theory was not merely inherently flawed, but had helped cause the calamity I expected. At the time, I thought that the bursting of the DotCom Bubble would mark the beginning of the crisis – though I was cautious in saying so, because my work in modeling Minsky's Financial Instability Hypothesis (Keen 1995) had confirmed one aspect of his theory, the capacity of government spending to prevent a debt crisis that would have occurred in a pure credit economy.

Statements that a crisis *may* occur were edited out of this edition, because the crisis has occurred – after the Subprime Bubble, which was in the background during the DotCom Bubble, finally burst as well.[1] But these pre-crisis statements remain important, because they indicate that, without the blinkers that neoclassical economic theory puts over the eyes of economists, the crisis now known as the Great Recession was not an unpredictable 'Black Swan' event, but an almost blindingly obvious certainty. The only question mark was over when it would occur, not if.

This brief chapter therefore provides excerpts from the first edition on the likelihood of a crisis as seen from the vantage point of non-neoclassical economics – and in particular, Minsky's 'Financial Instability Hypothesis' – in 2000 and early 2001. I hope these pre-crisis observations persuade you to reject the 'Nobody could have seen this coming' smokescreen. Rather than being a 'Black Swan', the Great Recession was a 'White Swan' made invisible to neoclassical economists because their theory makes them ignore the key factors that caused it: debt, disequilibrium, and time.

The destabilizing effect of neoclassical economics

The belief that a capitalist economy is inherently stabilizing is also one for which inhabitants of market economies may pay dearly in the future. As they were initially during the Great Depression, economists today may be the main force preventing the introduction of countervailing measures to any future economic slump. Economics may make our recessions deeper, longer and more intractable, when the public is entitled to expect economics to have precisely the opposite effect.

Fortunately for economists, the macroeconomy – at least in the United States – appeared to be functioning fairly well at the end of the year 2000. It is thus possible for economists to believe and preach almost anything, because they can bask in the entirely coincidental fact that the macroeconomy appears healthy.

However, this accidental success may not last long if the pressures which have been clearly growing in the financial side of the economy finally erupt (Keen 2001a: 213).

Possibility of debt deflation in the USA

If a crisis does occur after the Internet Bubble finally bursts, then it could occur in a milieu of low inflation (unless oil price pressures lead to an inflationary spiral). Firms are likely to react to this crisis by dropping their margins in an attempt to move stock, or to hang on to market share at the expense of their competitors. This behavior could well turn low inflation into deflation.

The possibility therefore exists that America could once again be afflicted with a debt deflation – though its severity could be attenuated by the inevitable increase in government spending that such a crisis would trigger. America could well join Japan on the list of the global economy's 'walking wounded' – mired in a debt-induced recession, with static or falling prices and a seemingly intractable burden of private debt (ibid.: 254).

The likelihood of a Japanese outcome for America after the crash

Only time will tell whether the bursting of the Internet Bubble will lead to as dire an outcome as the Great Depression. Certainly, on many indicators, the 1990s bubble has left its septuagenarian relative in the shade.

The price to earnings ratio peaked at over one and a half times the level set in 1929, the private and corporate debt to output ratio is possibly three times what it was prior to the Great Crash, and prices, though rising in some sectors, are generally quiescent. On all these fronts, Fisher's debt-deflation theory of great depressions seems a feasible outcome.

On the other hand, Minsky argued that 'Big Government' could stabilize an unstable economy, by providing firms with cash flow from which their debt commitments could be financed despite a collapse in private spending. Certainly, the US government of 2000 is 'big' when compared to its 1920s counterpart, and its automatic and policy interventions will probably attenuate any economic crash to something far milder than the Great Depression. What appears more likely for post-Internet America is a drawn-out recession like that experienced by Japan since its Bubble Economy collapsed in 1990 (ibid.: 256–7).

The impact of the Maastricht Treaty on Europe during a crisis

Macroeconomics is economic policy par excellence, but economic theory itself has virtually reached the position that there should be no macroeconomic policy. The clearest evidence of this is the Maastricht Treaty, which made restricting budget deficits to no more than 3 percent of GDP a condition for membership of the European Union. While some fudging has been allowed to make membership possible in the first place, when an economic crisis eventually strikes, Europe's governments may be compelled to impose austerity upon economies which will be in desperate need of a stimulus (ibid.: 212–13).

The Efficient Markets Hypothesis encouraging debt-financed speculation

[According to the Efficient Markets Hypothesis] The trading profile of the stock market should therefore be like that of an almost extinct volcano. Instead, even back in the 1960s when this [Sharpe] paper was written, the stock market behaved like a very active volcano. It has become even more so since, and in 1987 it did a reasonable, though short-lived, impression of Krakatau. In 2000, we saw 25 percent movements in a week. October 2000 lived up to the justified reputation of that month during bull markets;

heaven only knows how severe the volatility will be when the bubble finally bursts (ibid.: 232).

What can I say? By promulgating the efficient markets hypothesis, which is predicated on each investor having the foresight of Nostradamus, economic theory has encouraged the world to play a dangerous game of stock market speculation. When that game comes unstuck, America in particular will most likely find itself as badly hobbled by debt as Japan has been for the past decade. This speculative flame may have ignited anyway, but there is little doubt that economists have played the role of petrol throwers rather than firemen. When crisis strikes, conventional economists will be the last people on the planet who can be expected to provide sage advice on how to return to prosperity – unless, as often happens in such circumstances, they drop their theoretical dogmas in favor of common sense.

When the Great Crash of 1929 led to the Great Depression of the 1930s, many of the erstwhile heroes of the finance sector found themselves in the dock. It is unlikely that any particular economists will find themselves so arraigned, but there is little doubt that economic theory has been complicit in encouraging America's investing public to once again delude itself into a crisis (ibid.: 256).

Deregulation and crisis

Deregulation of the financial sector was not the sole cause of the financial instability of the past twenty years. But it has certainly contributed to its severity, by removing some of the limited constraints to cyclical behavior which exist in a regulated system.

These deregulations were mooted as 'reforms' by their proponents, but they were in reality retrograde steps, which have set our financial system up for a real crisis. I can only hope that, if the crisis is serious enough, then genuine reform to the finance sector will be contemplated. Reform, of course, cannot make capitalism stable; but it can remove the elements of our corporate system which contribute most strongly to instability.

The major institutional culprit has to be the finance sector itself, and in particular the elements of the stock market which lead to it behaving more like a casino than a place of reasoned calculation […]

Surely, when the Internet Bubble really bursts, it will be time to admit that one fundamental excess of the market as currently organized is its ability to allow sky-high valuations to develop (ibid.: 255–6).

The history of crises causing – and not causing – paradigm shifts in economics

This is far from the first book to attack the validity of economics, and it is unlikely to be the last. As Kirman commented, economic theory has seen off many attacks, not because it has been strong enough to withstand them, but because it has been strong enough to ignore them.

Part of that strength has come from the irrelevance of economics. You don't need an accurate theory of economics to build an economy in the same sense that you need an accurate theory of propulsion to build a rocket. The market economy began its evolution long before the term 'economics' was ever coined, and it will doubtless continue to evolve regardless of whether the dominant economic theory is valid. Therefore, so long as the economy itself has some underlying strength, it is a moot point as to whether any challenge to economic orthodoxy will succeed.

However, while to some extent irrelevant, economics is not 'mostly harmless'. The false confidence it has engendered in the stability of the market economy has encouraged policy-makers to dismantle some of the institutions which initially evolved to try to keep its instability within limits. 'Economic reform,' undertaken in the belief that it will make society function better, has instead made modern capitalism a poorer social system: more unequal, more fragile, more unstable. And in some instances, as in Russia, a naive faith in economic theory has led to outcomes which, had they been inflicted by weapons rather than by policy, would have led their perpetrators to the International Court of Justice.

But even such a large-scale failure as Russia seems to have little impact upon the development of economic theory. For economics to change, it appears that things have to 'go wrong' on a global scale, in ways which the prevailing theory believed was impossible. There have been two such periods this century.

The first and most severe was the Great Depression, and in that calamity, Keynes turned economic theory upside down. However, Keynes's

insights were rapidly emasculated, as Chapter 9 showed. 'Keynesian economics' became dominant, but it certainly was not the economics of Keynes.

The second was the 'stagflationary crisis' – the coincidence of low growth, rising unemployment and high inflation during the 1970s. That crisis led to the final overthrow of the emasculated creature that Keynesian economics had become, and its replacement by an economic orthodoxy which was even more virile than that against which Keynes had railed.

One step forward and two steps back – with the first step backwards being taken when the economy was doing well, in the aftermath of the Depression and WWII and hence when the ramblings of economists could comfortably be ignored.

That historical record is both comforting and disturbing. Change is possible in economics, but normally only when the fabric of society itself seems threatened; and change without crisis can involve the forgetting of recent advances.

It is possible, therefore, that economic theory may continue to function mainly as a surrogate ideology for the market economy, right up until the day, in some distant future, when society evolves into something so profoundly different that it no longer warrants the moniker 'capitalism.'

I hope, however, that events follow a different chain. I am not wishing an economic crisis upon the modern world – instead, I think one has been well and truly put in train by the cumulative processes described in chapters 10 and 11. If that crisis eventuates – one which neoclassical economic theory argues is not possible – then economics will once again come under close and critical scrutiny (ibid.: 311–12).

Public reactions to the failure of neoclassical economics

This time, the chances are much better that something new and indigestibly different from the prevailing wisdom will emanate from the crisis. As this book has shown, critical economists are much more aware of the flaws in conventional economics than they were during Keynes's day, non-orthodox analysis is much more fully developed, and advances in many other fields of science are there for the taking, if economics can be

persuaded – by force of circumstance – to abandon its obsession with equilibrium.

The first factor should mean that the lines will be much more clearly drawn between the old orthodoxy and the new. The latter two should mean that the techniques of the old orthodoxy will look passé, rather than stimulating, to a new generation of economists schooled in complexity and evolutionary theory.

But ultimately, schooling is both the answer and the problem. If a new economics is to evolve, then it must do so in an extremely hostile environment – the academic journals and academic departments of Economics and Finance, where neoclassic orthodoxy has for so long held sway.[2] The nurturing of a new way of thinking about economics could largely be left in the hands of those who have shown themselves incapable of escaping from a nineteenth-century perspective.

There are two possible palliatives against that danger. The first is the development, by non-orthodox economists, of a vibrant alternative approach to analyzing the economy which is founded in realism, rather than idealism. Such a development would show that there is an alternative to thinking about the economy in a neoclassical way, and offer future students of economics a new and hopefully exciting research program to which they can contribute.

The second is an informed and vigilant public. If you have struggled to the end of this book, then you now have a very strong grasp on the problems in conventional economic thought, and the need for alternative approaches to economics. Depending on your situation, you can use this knowledge as a lever in all sorts of ways.

If you are or you advise a person in authority in the private or public sectors, you should know now not to take the advice of economists on faith. They have received far too easy a ride as the accepted vessels of economic knowledge. Ask a few enquiring questions, and see whether those vessels ring hollow. When the time comes to appoint advisers on economic matters, quiz the applicants for their breadth of appreciation of alternative ways to ‘think economically,’ and look for the heterodox thinker rather than just the econometric technician.

If you are a parent with a child who is about to undertake an economics or business degree, then you're in a position to pressure potential schools to take a pluralist approach to education in economics. A quick glance through course structure booklets and subject outlines should be enough to confirm what approach they take at present.

If you are a student now? Well, your position is somewhat compromised: you have to pass exams, after all! I hope that, after reading this book, you will be better equipped to do that. But you are also equipped to 'disturb the equilibrium' of both your fellow students and your teachers, if they are themselves ignorant of the issues raised in this book.

You have a voice, which has been perhaps been quiescent on matters economic because you have in the past deferred to the authority of the economist. There is no reason to remain quiet.

I commented at the beginning of this book that economics was too important to leave to the economists. I end on the same note (ibid.: 312–13).

Postscript 2011

As these excerpts emphasize, the never-ending crisis in which the USA and much of the OECD is now ensnared was no 'Black Swan.' Its inevitability was obvious to anyone who paid attention to the level of debt-financed speculation taking place, and considered what would happen to the economy when the debt-driven party came to an end. The fact that the vast majority of economists pay no attention at all to these issues is why they were taken by surprise.

It may astonish non-economists to learn that conventionally trained economists ignore the role of credit and private debt in the economy – and frankly, it is astonishing. But it is the truth. Even today, only a handful of the most rebellious of mainstream 'neoclassical' economists – people like Joe Stiglitz and Paul Krugman – pay any attention to the role of private debt in the economy, and even they do so from the perspective of an economic theory in which money and debt play no intrinsic role. An economic theory that ignores the role of money and debt in a market economy cannot possibly make sense of the complex, monetary, credit-based economy in which we live. Yet that is the theory that has dominated economics for the

last half-century. If the market economy is to have a future, this widely believed but inherently delusional model has to be jettisoned.

2 | NO MORE MR NICE GUY

Why economics must undergo a long-overdue intellectual revolution

A decade ago, economics appeared triumphant. Though the spiritually inclined might have railed at its materialistic way of looking at the world, it nonetheless appeared that the materialistic road to riches was working. After decades of stagnation, significant sections of the developing world were in fact developing; a long-running boom in the USA had continued with only the slightest hiccup after the Nasdaq crash in April 2000; and in the USA and many other advanced nations, both inflation and unemployment were trending down in a process that leading economists christened 'The Great Moderation.'

It seemed that, after the turmoil of the period from the late 1960s till the recession of the early 1990s, economists had finally worked out how to deliver economic nirvana. To do so, they rejected many of the concepts that had been introduced into economics by the 'Keynesian Revolution' in the 1930s.

The resulting theory of economics was called Neoclassical Economics, to distinguish it from the 'Keynesian Economics' it had overthrown (though in a confusing twist, the major subgroup within neoclassical economics called itself 'New Keynesian').[1] In many ways, it was a return to the approach to economics that had been dominant prior to Keynes, and for that reason it was often referred to as 'the Neoclassical Counter-Revolution.'

At a practical level, neoclassical economics advocated reducing government intervention in the economy and letting markets – especially finance markets – decide economic outcomes unimpeded by politicians, bureaucrats or regulations. Counter-cyclical government budget policy – running deficits during downturns and surpluses during booms – gave way to trying to run surpluses all the time, to reduce the size of the government

sector. The only policy tool in favor was manipulation of the interest rate – by a politically independent central bank which itself was controlled by neoclassical economists – with the objective of controlling the rate of inflation.

At a deep theoretical level, neoclassical economics replaced many tools that Keynes and his supporters had developed to analyze the economy as a whole ('macroeconomics') with their own tools. Unlike the analytic tools of Keynesian macroeconomics, the new neoclassical macroeconomics toolset was derived directly from microeconomics – the theory of how the individual agents in the economy behave.

Purge

Not all academic economists joined in this overthrow of the previous Keynesian orthodoxy. Many fought against it, though ultimately to no avail, and academic economics eventually divided into roughly six camps: the dominant neoclassical school that represented perhaps 85 percent of the profession, and several small rumps called Post-Keynesian, Institutional, Evolutionary, Austrian and Marxian economics.

An outsider might have expected this situation to lead to vigorous debates within the academy. In fact, what eventually evolved was a mixture of both hostility and indifference. Neoclassical economists didn't pay any attention to what these rumps said, but they also gave up on early attempts to eliminate them. Try as they might, they could never get rid of the dissidents completely, for two main reasons.

First, some, like myself, had always been opposed to neoclassical economics, and were hard to remove because of impediments like academic tenure. Secondly, others would begin as neoclassical economists, but then undergo some personal epiphany that would lead them to abandon this approach and swap horses to one of the dissident streams.

So, though neoclassical economists dominated almost all academic economic departments, they were also forced to tolerate the odd critic within. But it was hardly peaceful coexistence.

In teaching, core courses on microeconomics, macroeconomics and finance were purged of non-neoclassical ideas. The odd non-neoclassical course continued as an option to give dissenters something to do, but

generally, non-neoclassical staff filled out most of their teaching time giving tutorials in subjects that taught neoclassical ideas with which they fundamentally disagreed. They toed the line in tuition and marking – though they would occasionally grumble about it, to encourage dissent in students who seemed more critical than the run of the mill.

In research, the purge was more complete, because neoclassical editors and referees could exclude the dissidents from the journals they edited. Up until the early 1970s, non-neoclassical authors were regularly published in the prestigious journals of the profession – for example, a major debate over the theories of production and distribution between neoclassical and non-neoclassical economists, known as the 'Cambridge Controversies,' largely occurred in the *American Economic Review* (*AER*), the *Economic Journal* (*EJ*), and the *Quarterly Journal of Economics* (*QJE*) – witness Joan Robinson's papers (Robinson 1971a, 1971b, 1972, 1975), including one entitled 'The second crisis of economic theory' in the *AER*. However, by the mid-1980s, these, their companion major journals the *Journal of Political Economy*, the *Journal of Economic Theory* and many other minor journals had become bastions of neoclassical thought. Papers that did not use neoclassical concepts were routinely rejected – frequently without even being refereed.

Non-neoclassical economists in general gave up on these citadels of orthodoxy, and instead established their own journals in which they communicated with each other, and vigorously criticized neoclassical theory. The *Journal of Post Keynesian Economics* (*JPKE*), founded in 1978 by Sidney Weintraub and Paul Davidson, was the first dedicated to non-neoclassical economics, and many others were subsequently established.

In public policy, as in the most prestigious journals, neoclassical economics reigned supreme. Few dissidents were ever appointed to positions of public influence,[2] and most bureaucratic positions were filled by graduates from the better colleges who – because of the purging of non-neoclassical ideas from the core curriculum – generally didn't even know that any other way of thinking about economics was possible. To them, neoclassical economics *was* economics.

Triumph

This purge within academia was aided and abetted by developments in the economy itself. Inflation, which had been as low as 1 percent in the early 1960s, began to rise in a series of cycles to a peak of 15 percent in 1980. Unemployment, which had in the past gone down when inflation went up, began to rise as well in the 1970s – in apparent contradiction of Keynesian doctrine.

As a result, the media and the public were clamoring for change, supporting the efforts of leading neoclassicals like Milton Friedman to overthrow their Keynesian overlords in the academy. The public policy focus shifted from the Keynesian emphasis upon keeping unemployment low – and tolerating higher inflation as a side effect – to keeping inflation low, in the belief that this would allow the private sector to 'do its thing' and achieve full employment.

The initial results were mixed – inflation plunged as Fed chairman Volcker pushed the cash rate[3] to 20 percent, but unemployment exploded to its post-war peak of almost 11 percent in 1983. But that painful crisis proved to be the worst under neoclassical management of economic policy. The next recession in the early 1990s had a peak unemployment rate of less than 8 percent. The one after that in 2003 had a peak unemployment rate of 6.3 percent.

Inflation had also come down, and fluctuated in a band between 1 and 4 percent, with occasional spikes up to 6 percent – far below the tumultuous level of the period from 1965 to 1985, when the *average* had been over 6 percent. Neoclassical economists enshrined the objective of keeping inflation low in the rules they set for central banks, which instructed them to manipulate the rate of interest to keep inflation in a narrow band between 1 and 3 percent.

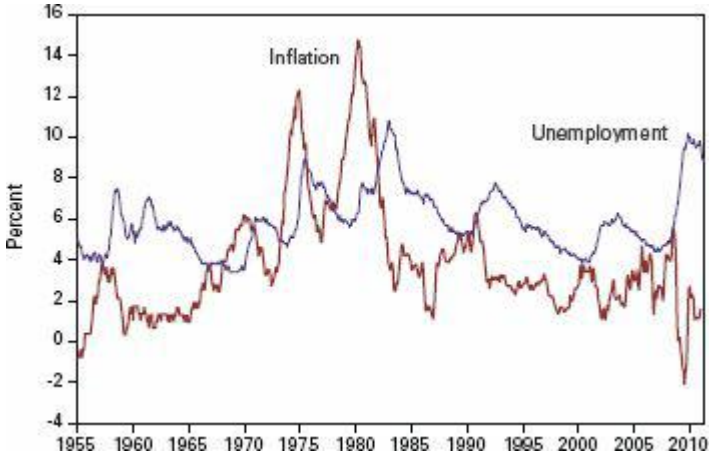
Looking back on how neoclassical economics had remodeled both economic theory and economic policy, the current US Federal Reserve chairman Ben Bernanke saw two decades of achievement. Writing in 2004, he asserted that there had been:

> not only significant improvements in economic growth and productivity but also a marked reduction in economic volatility, both in the United States and abroad, a phenomenon that has been dubbed 'the Great Moderation.'

> Recessions have become less frequent and milder, and quarter-to-quarter volatility in output and employment has declined significantly as well.
>
> The sources of the Great Moderation remain somewhat controversial, but as I have argued elsewhere, there is evidence for the view that improved control of inflation has contributed in important measure to *this welcome change in the economy*. (Bernanke 2004b; emphasis added)

The chief economist of the OECD, Jean-Philippe Cotis, was equally sanguine about the immediate economic prospects in late May of 2007:

> In its Economic Outlook last Autumn, the OECD took the view that the US slowdown was not heralding a period of worldwide economic weakness, unlike, for instance, in 2001. Rather, a 'smooth' rebalancing was to be expected, with Europe taking over the baton from the United States in driving OECD growth.
>
> Recent developments have broadly confirmed this prognosis. Indeed, *the current economic situation is in many ways better than what we have experienced in years*. Against that background, we have stuck to the rebalancing scenario. *Our central forecast remains indeed quite benign:* a soft landing in the United States, a strong and sustained recovery in Europe, a solid trajectory in Japan and buoyant activity in China and India. In line with recent trends, sustained growth in OECD economies would be underpinned by strong job creation and falling unemployment. (Cotis 2007: 7; emphases added)



2.1 US inflation and unemployment from 1955

Then, in late 2007, the 'Great Moderation' came to an abrupt end.

Crisis

Suddenly, everything that neoclassical economics said couldn't happen, happened all at once: asset markets were in free-fall, century-old bastions of finance like Lehman Brothers fell like flies, and the defining characteristics of the Great Moderation evaporated: unemployment skyrocketed, and mild inflation gave way to deflation.

Confronted by a complete disconnect between what they believed and what was happening, economists reacted in a very human way: they panicked. Suddenly, they threw their neoclassical policy rules out the window, and began to behave like 'Keynesian' economists on steroids. Having eschewed government intervention, budget deficits, and boosting government-created money for decades, at their command the government was everywhere. Budget deficits hit levels that dwarfed anything that old-fashioned Keynesians had ever run in the 1950s and 1960s, and government money flowed like water over the Niagara Falls. Ben Bernanke, as Federal Reserve chairman, literally doubled the level of government-created money in the US economy in five months, when the previous doubling had taken thirteen years. A long decay in the ratio of government-created money to the level of economic activity, from 15 percent of GDP in 1945 to a low of 5 percent in 1980, and 6 percent when the crisis began, was eliminated in less than a year as Bernanke's 'Quantitative Easing 1' saw the ratio rocket back to 15 percent by 2010.

2.2 Bernanke doubles base money in five months

The tenor of these times is well captured in Hank Paulson's *On the Brink*:

> 'We need to buy hundreds of billions of assets,' I said. I knew better than to utter the word trillion. That would have caused cardiac arrest. 'We need an announcement tonight to calm the market, and legislation next week,' I said.
>
> What would happen if we didn't get the authorities we sought, I was asked.
>
> 'May God help us all,' I replied. (Paulson 2010: 261)

As they threw their once-cherished neoclassical economic principles out the window, and ran about in panic like a coop full of Chicken Littles, the overwhelming refrain from the public was 'Why didn't you see this coming? And if you're experts on the economy and you were in control of it, why did the crisis happen in the first place?' The first question was famously put directly to academic economists by the Queen of England at the prestigious London School of Economics:

> During a briefing by academics at the London School of Economics on the turmoil on the international markets the Queen asked: 'Why did nobody notice it?'
>
> Professor Luis Garicano, director of research at the London School of Economics' management department, had explained

> the origins and effects of the credit crisis when she opened the £71 million New Academic Building.
>
> The Queen, who studiously avoids controversy and never gives away her opinions, then described the turbulence on the markets as 'awful'. (Pierce 2008)

The answer these economists later gave the Queen[4] was a popular refrain for a profession that, after decades of dominating economic and social policy around the world, suddenly found itself under concerted attack, with its opinions openly derided. It wasn't their fault, because 'No One Could Have Seen This Coming': though the risks to individual positions could be calculated, no one could have foreseen the risk to the system as a whole:

> the difficulty was seeing the risk to the system as a whole rather than to any specific financial instrument or loan. Risk calculations were most often confined to slices of financial activity, using some of the best mathematical minds in our country and abroad. But they frequently lost sight of the bigger picture. (Besley and Hennessy 2009: 1)

Balderdash. Though the precise timing of the crisis was impossible to pick, a systemic crisis was both inevitable and, to astute observers in the mid-2000s, likely to occur in the very near future. That is why I and a handful of other unconventional economists went public in the years leading up to the crisis, warning whenever and however we could that a serious economic calamity was imminent.

'No one saw this coming'

In a paper with the mocking title of '"No one saw this coming": understanding financial crisis through accounting models'[5] (Bezemer 2009, 2010, 2011), Dutch academic Dirk Bezemer trawled through academic and media reports looking for any people who had warned of the crisis before it happened, and who met the following exacting criteria:

> Only analysts were included who:

- provided some account of how they arrived at their conclusions.

- went beyond predicting a real estate crisis, also making the link to real-sector recessionary implications, including an analytical account of those links.
- the actual prediction must have been made by the analyst and available in the public domain, rather than being asserted by others.
- the prediction had to have some timing attached to it. (Bezemer 2009: 7)

Bezemer came up with twelve names: myself and Dean Baker, Wynne Godley, Fred Harrison, Michael Hudson, Eric Janszen, Jakob Brøchner Madsen and Jens Kjaer Sørensen, Kurt Richebächer, Nouriel Roubini, Peter Schiff, and Robert Shiller.

He also identified four common aspects of our work:

1 a concern with financial assets as distinct from real-sector assets,
2 with the credit flows that finance both forms of wealth,
3 with the debt growth accompanying growth in financial wealth, and
4 with the accounting relation between the financial and real economy. (Ibid.: 8)

TABLE 2.1 Anticipations of the housing crisis and recession

Analyst	Country	Capacity	Forecast
Dean Baker	USA	Co-director, Center for Economic and Policy Research	... plunging housing investment will likely push the economy into recession (2006)
Wynne Godley	USA	Distinguished Scholar, Levy Economics Institute of Bard College	The small slowdown in the rate at which US household debt levels are rising resulting from the house price decline will immediately lead to a ... growth recession ... before 2010 (2006); [will] start to rise significantly and does not come down again (2007)
Fred Harrison	UK	Economic commentator	The next property market tipping point is due at end of 2007 or early 2008 ... only way prices can be brought back to affordable levels is a slump or recession (2005)
Michael Hudson	USA	Professor, University of Missouri	Debt deflation will shrink the economy, drive down real wages, and push our debt-ridden economy into Japan-style stagnation or worse (2006)
Eric Janszen	USA	Investor and iTulip commentator	The US will enter a recession within years (2006); stock markets are likely to begin in 2008 to experience a 'Deflation Bear Market' (2007)
Stephen Keen	Australia	Associate Professor, University of Western Sydney	Long before we manage to reverse the current rise in debt, the economy will be in a recession. On current data, we may already be in one (2006)
Jakob Brechner Madsen and Jens Kjaer Sørensen	Denmark	Professor and graduate student, Copenhagen University	We are seeing large bubbles and if they bust, there is no backup. The outlook is very bad (2005); The bursting of this housing bubble will have a severe impact on the world economy and may even result in a recession (2006)
Kurt Richebächer	USA	Private consultant and investment newsletter writer	The new housing bubble – together with the bond and stock bubbles – will invariably implode in the foreseeable future, plunging the US Economy into a protracted, deep recession (2001); recession and bear market in asset prices are inevitable for the US economy ... All remaining questions pertain solely to speed, depth and duration of the economy's downturn (2006)
Nouriel Roubini	USA	Professor, New York University	Real home prices are likely to fall at least 30% over the next three years; itself this house price slump is enough to trigger a US recession (2006)
Peter Schiff	USA	Stockbroker, investment adviser and commentator	The United States economy is like the Titanic ... I see a real financial crisis coming for the United States (2006); will be an economic collapse (2007)
Robert Shiller	USA	Professor, Yale University	There is significant risk of a very bad period, with rising default and foreclosures, serious trouble in financial markets, and a possible recession sooner than most of us expected (2006)

If you have never studied economics before, this list may surprise you: don't *all* economists consider these obviously important economic issues?

As you will learn in this book, the answer is no. Neoclassical economic theory ignores all these aspects of reality – even when, on the surface, they might appear to include them. Bezemer gives the example of the OECD's 'small global forecasting' model, which makes forecasts for the global

economy that are then disaggregated to generate predictions for individual countries – it was the source of Cotis's statement 'Our central forecast remains indeed quite benign' in the September 2007 *OECD Economic Outlook*.

This OECD model apparently includes monetary and financial variables. However, these are not taken from data, but are instead derived from theoretical assumptions about the relationship between 'real' variables – such as 'the gap between actual output and potential output' – and financial variables. As Bezemer notes, the OECD's model lacks *all* of the features that dominated the economy in the lead-up to the crisis: 'There are no credit flows, asset prices or increasing net worth driving a borrowing boom, nor interest payment indicating growing debt burdens, and no balance sheet stock and flow variables that would reflect all this' (ibid.: 19).

How come? Because standard 'neoclassical' economic theory assumes that the financial system is rather like lubricating oil in an engine – it enables the engine to work smoothly, but has no driving effect. Neoclassical economists therefore believe that they can ignore the financial system in economic analysis, and focus on the 'real' exchanges going on behind the 'veil of money.'

They also assume that the real economy is, in effect, a miracle engine that always returns to a state of steady growth, and never generates any undesirable side effects – rather like a pure hydrogen engine that, once you take your foot off the accelerator or brake, always returns to a steady 3,000 revs per minute, and simply pumps pure water into the atmosphere.[6]

To continue the analogy, the common perspective in the approaches taken by the economists Bezemer identified is that we see finance as more akin to petrol than oil. Without it, the 'real economy' engine revs not at 3,000 rpm, but zero, while the exhaust fumes contain not merely water, but large quantities of pollutants as well.

As the financial crisis made starkly evident, neoclassical economists were profoundly wrong: the issues they ignored were vital to understanding how a market economy operates, and their deliberate failure to monitor the dynamics of private debt was the reason why they did not see this crisis coming – and why they are the last ones who are likely to work out how to end it.

Consequently, neoclassical economics, far from being the font of economic wisdom, is actually the biggest impediment to understanding how the economy actually works – and why, periodically, it has serious breakdowns. If we are ever to have an economic theory that actually describes the economy, let alone one that helps us manage it, neoclassical economics has to go.

Revisionism

Yet this is not how neoclassical economists themselves have reacted to the crisis. Bernanke, whose appointment as chairman of the US Federal Reserve occurred largely because he was regarded by his fellow neoclassical economists as *the* academic expert on the Great Depression, has argued that there is no need to overhaul economic theory as a result of the crisis. Distinguishing between what he termed 'economic science, economic engineering and economic management,' he argued that:

> the recent financial crisis was more a failure of economic engineering and economic management than of what I have called economic science […]
>
> Shortcomings of […] economic science […] were for the most part less central to the crisis; indeed, although the great majority of economists did not foresee the near-collapse of the financial system, economic analysis has proven and will continue to prove critical in understanding the crisis, in developing policies to contain it, and in designing longer-term solutions to prevent its recurrence. (Bernanke 2010: 3)

However, Bernanke's primary argument in defense of neoclassical economics is simply silly, because he defends modern economic theory by pointing to the work of theorists that most neoclassical economists would never have heard of: 'The fact that dependence on unstable short-term funding could lead to runs is hardly news to economists; it has been a central issue in monetary economics since Henry Thornton and Walter Bagehot wrote about the question in the 19th century' (ibid.: 6).

This might give non-economists the impression that the works of Thornton and Bagehot are routinely studied by today's economists – or that today's neoclassical economic toolkit is based, among other pillars, on such historically informed sources. However, a significant aspect of the Neoclassical Counter-Revolution was the abolition of courses on economic history and the history of economic thought, in which the works of Thornton and Bagehot would have occasionally featured.

Today, only rebel, non-neoclassical economists – or a central banker with a personal interest in monetary history like Bernanke – is likely to have read Thornton, Bagehot, or any analysis of any financial crises prior to this one. Core neoclassical courses on microeconomics and macroeconomics are devoid of any discussion of financial crises, let alone pre-twentieth-century analysis of them, while even specialist 'Money and Banking' courses teach neoclassical models of money and banking, rather than historical or pre-neoclassical analysis.[7]

One of the few textbook writers who has been trying – largely without success – to broaden the economic curriculum reacted similarly to Bernanke's paper.

> I find this justification very strange. In my view, the fact that Thornton and Bagehot provided useful insights into macroeconomic policy problems is an indictment of fundamental macroeconomic science as currently conceived. If it were fundamental science, it would be taught somewhere – ideally in the core macro courses. That doesn't happen. The core macroeconomic courses teach DSGE ['Dynamic Stochastic General Equilibrium'] modeling almost exclusively.
>
> Not only are the writings of Thornton or Bagehot missing, the writings of Keynes, Minsky, Hicks, Clower, Leijonhufvud, Gurley, Davidson, Goodhardt, Clower, or even Friedman, to mention just a few of those whose writings could also have contributed to a better understanding of the crisis, are missing as well. Most students who have graduated in the past twenty years would never have even heard of half of them, let alone read them.

> *If nobody reads them, and their ideas aren't part of the material that students study or learn, how can Bernanke consider them part of modern economic science?* (Colander 2011: 4–5; emphasis added)

In other words, defending modern economics by pointing to the work of pre-neoclassical economists is rather like rebutting criticisms of modern art by extolling the virtues of Leonardo Da Vinci. It is a fob-off, rather than a serious response to criticism.

Bernanke comes closer to engaging with reality when he admits that mainstream neoclassical models failed to predict the crisis: 'Standard macroeconomic models, such as the workhorse new-Keynesian model, did not predict the crisis, nor did they incorporate very easily the effects of financial instability' (Bernanke 2010: 16–17).

But rather than seeing this as a weakness that necessitated revision, Bernanke defended these models on the basis that they are appropriate for non-crisis times:

Do these failures of standard macroeconomic models mean that they are irrelevant or at least significantly flawed? I think the answer is a qualified no. Economic models are useful only in the context for which they are designed. *Most of the time, including during recessions, serious financial instability is not an issue. The standard models were designed for these non-crisis periods, and they have proven quite useful in that context.* Notably, they were part of the intellectual framework that helped deliver low inflation and macroeconomic stability in most industrial countries during the two decades that began in the mid-1980s. (Ibid.: 17; emphasis added)

The sheer naivety of this argument caused me pause when writing this chapter. How does one even begin to respond to such a blasé perspective on the role of economic theory, especially when expressed by someone of such reputed knowledge, and in a position of such responsibility, who surely should know better?

There are many tacks I could have taken. The defense of having models for good times would be valid only if there were also models for bad times – but neoclassical economics has no such models. The quaint belief that the

conditions prior to the crisis – the so-called Great Moderation – had no connection with the events that followed shows that he has no idea as to what caused the Great Recession.

Ultimately, the most apposite critique of Bernanke's defense of the indefensible is to compare his position with that of the post-Keynesian economist Hyman Minsky. Minsky argued that, since crises like the Great Depression have occurred, a crucial test for the validity of an economic theory is that it must be able to generate a depression as one of its possible states:

Can 'It' – a Great Depression – happen again? And if 'It' can happen, why didn't 'It' occur in the years since World War II? These are questions that naturally follow from both the historical record and the comparative success of the past thirty-five years. To answer these questions it is necessary to have an economic theory which makes great depressions one of the possible states in which our type of capitalist economy can find itself. (Minsky 1982: 5)

On this basis, Minsky rejected neoclassical economics for the very reason that Bernanke defends it above: in its core models, a depression is an impossibility. Therefore, the neoclassical model is an inadequate basis for modeling and understanding capitalism:

> The abstract model of the neoclassical synthesis cannot generate instability. When the neoclassical synthesis is constructed, capital assets, financing arrangements that center around banks and money creation, constraints imposed by liabilities, and the problems associated with knowledge about uncertain futures are all assumed away. For economists and policy-makers to do better we have to abandon the neoclassical synthesis. (Ibid.: 5)

Clearly, Bernanke shows no such inclination. Even in the aftermath of a financial crisis that took him and the vast majority of neoclassical economists completely by surprise, and which terrified them as much as it bewildered the public, Bernanke and his many neoclassical colleagues still cling to their belief in an economic theory that asserts that events like this could never happen.

Ignorance

A major reason for Bernanke's inability to accept that the core of neoclassical economics is 'irrelevant or at least significantly flawed' is that, in common with so many of his neoclassical peers, he innately believes that the neoclassical model of the economy is essentially correct – so much so that even the financial crisis could not shake his faith in it.

This faith emanates from the seductive nature of the neoclassical vision. It portrays capitalism as a perfect system, in which the market ensures that everything is 'just right.' It is a world in which meritocracy rules, rather than power and privilege as under previous social systems. This vision of a society operating perfectly without a central despotic authority is seductive – so seductive that neoclassical economists *want* it to be true.

This faith is maintained by a paradoxical, transcendental truth: *neoclassical economists don't understand neoclassical economics*. Their belief that it is a coherent, comprehensive theory of how a market economy operates is based on a profound ignorance of the actual foundations of the theory.[8]

In one sense, their ignorance is utterly justified, because they are behaving in the same way that professionals do in genuine sciences like physics. Most physicists don't check what Einstein actually wrote on the Theory of Relativity, because they are confident that Einstein got it right, and that their textbooks accurately communicate Einstein's core ideas. Similarly, most economists don't check to see whether core concepts like 'supply and demand microeconomics' or 'representative agent macroeconomics' are properly derived from well-grounded foundations, because they simply assume that if they're taught by the textbooks, then there must be original research that confirms their validity.

In fact, the exact opposite is the case: *the original research confirms that all these concepts are false*. Virtually every concept that is taught as gospel in the textbooks has been proved to be unsound in the original literature.

If they actually appreciated what the foundations were – and how utterly flawed they really are – then neoclassical economists would run a mile from their beliefs, and feel compelled to look for alternatives. But they have no knowledge of the actual state of neoclassical economics because their education shields them from it, right from their very first exposure to

economic theory (for the rest of the book, if I say 'economics' without qualification, I will normally mean 'neoclassical economics,' unless otherwise noted).

Educated into ignorance

If the real world were accurately described by economic textbooks, there would not now be a financial crisis – and nor would there ever have been one in the past either: the Great Depression would not have happened. The economy would instead be either in equilibrium, or rapidly returning to it, with full employment, low inflation, and sensibly priced assets.

Of course, the real world is nothing like that. Instead, it has been permanently in disequilibrium, and in near-turmoil, ever since the financial crisis began in 2007. So the textbooks are wrong. But there is a bizarre irony in this disconnect between reality and economic textbooks. *If those same textbooks gave an accurate rendition of the underlying theory, they would describe an economy that generated cycles, was in disequilibrium all the time, and was prone to breakdown.*

This is not because the theory itself envisages a turbulent, cyclical world – far from it. The underlying neoclassical vision of the market economy is of permanent equilibrium, just as the textbooks portray it. However, there are preconditions for that state of equilibrium to apply, and deep economic research has established that none of them holds.

These preconditions arise from the neoclassical practice of analyzing the economy from the point of view of individual 'agents,' where those agents can be consumers, firms, workers, or investors. Generally speaking, though the description of the individual itself can be criticized as stylized and barren, this analysis is internally consistent: if you accept the model's assumptions, then the conclusions about *individual* behavior flow logically from them.

However, to be a theory of economics rather than one of individual psychology, this model of the individual must be aggregated to derive a model of a market, where many individual consumers and sellers interact, or an entire economy where multiple markets interact with each other. The analysis of the individual must be aggregated somehow, to derive a theory of the aggregate entity called 'The Market' or 'The Economy.'

In literally every case, the attempt to move from the analysis of the individual to the aggregate failed – in the sense that results that were easily derived for the isolated individual could not be derived for the aggregate. But this failure to derive a coherent model of aggregate economic behavior was suppressed from the economics textbooks. Students were therefore taught a theory of how markets and economies behave which was strictly true only for isolated individuals, and was false for markets and economies themselves.

As I explain in the next chapter, this applies to the simplest and in many ways most fundamental concept in neoclassical economics – the 'downward-sloping demand curve' that is one half of its iconic 'supply and demand' analysis of markets. The theory proves that an individual's demand curve is downward-sloping – i.e. that an individual will buy more units of a commodity if its price falls – but the attempt to prove that the market demand curve also sloped downwards failed. However, textbooks writers are either truly ignorant of this failure, or delude themselves about the failure, or deliberately obfuscate it.

For example, the latest edition of Samuelson's textbook (whose first edition in 1948 set the neoclassical standard for economic instruction ever since) asserts that to derive a market demand curve, all you have to do is add together individual demand curves, and the resulting market demand curve will behave just like the individual demand curves from which it was derived: 'The market demand curve is found by adding together the quantities demanded by all individuals at each price. Does the market demand curve obey the law of downward-sloping demand? It certainly does' (Samuelson and Nordhaus 2010: 48).

That statement is provably false. The true situation is honestly stated in a leading research book, the *Handbook of Mathematical Economics*: 'market demand functions need not satisfy in any way the classical restrictions which characterize consumer demand functions […] The utility hypothesis tells us nothing about market demand unless it is augmented by additional requirements' (Shafer and Sonnenschein 1982: 671).

As I explain in the next chapter, the 'additional requirements' needed to ensure that a market demand curve slopes downwards are patently absurd. The realistic conclusion therefore is that market demand curves should have

any shape *except* the one that is drawn in the textbooks, and standard 'supply and demand' analysis becomes impossible.

However, economics students don't get to learn about this or any other aggregation failure. As the extract from Samuelson and Nordhaus illustrates, neoclassical textbooks present a sanitized, uncritical rendition of conventional economic theory, either ignoring problems with aggregation, or directly contradicting the results of advanced research. The courses in which these textbooks are used do little to counter this mendacious presentation. Students might learn, for example, that 'externalities' reduce the efficiency of the market mechanism. However, they will not learn that the 'proof' that markets are efficient is itself flawed.

Since this textbook rendition of economics is also profoundly boring, many students do no more than an introductory course in economics, and instead go on to careers in accountancy, finance or management – in which, nonetheless, many continue to harbor the simplistic notions they were taught many years earlier.

The minority which continues on to further academic training is taught the complicated techniques of neoclassical economic analysis, with little to no discussion of whether these techniques are actually intellectually valid. The enormous critical literature is simply left out of advanced courses, while glaring logical shortcomings are glossed over with specious assumptions. However, most students accept these assumptions because their training leaves them both insufficiently literate and insufficiently numerate.

Modern-day economics students are insufficiently literate because economic education eschews the study of the history of economic thought. Even a passing acquaintance with this literature exposes the reader to critical perspectives on conventional economic theory – but students today receive no such exposure. They are insufficiently numerate because the material which establishes the intellectual weaknesses of economics is complex. Understanding this literature in its raw form requires an appreciation of some quite difficult areas of mathematics – concepts which require up to two years of undergraduate mathematical training to understand.

Curiously, though economists like to intimidate other social scientists with the mathematical rigor of their discipline, most economists do *not* have

this level of mathematical education.

Instead, most economists learn their mathematics by attending courses in mathematics given by other economists. The argument for this approach – the partially sighted leading the partially sighted – is that generalist mathematics courses don't teach the concepts needed to understand mathematical economics (or the economic version of statistics, known as econometrics). This is quite often true. However, this has the side effect that economics has produced its own peculiar versions of mathematics and statistics, and has persevered with mathematical methods which professional mathematicians have long ago transcended. This dated version of mathematics shields students from new developments in mathematics that, incidentally, undermine much of economic theory.

One example of this is the way economists have reacted to 'chaos theory' (discussed in Chapter 9). Most economists think that chaos theory has had little or no impact – which is generally true in economics, but not at all true in most other sciences. This is partially because, to understand chaos theory, you have to understand an area of mathematics known as 'ordinary differential equations.'[9] Yet this topic is taught in very few courses on mathematical economics – and where it is taught, it is not covered in sufficient depth. Students may learn some of the basic techniques for handling what are known as 'second-order linear differential equations,' but chaos and complexity begin to manifest themselves only in 'third order nonlinear differential equations.'[10]

Economics students therefore graduate from master's and PhD programs with an uncritical and unjustified belief that the foundations of economic analysis are sound, no appreciation of the intellectual history of their discipline, and an approach to mathematics which hobbles both their critical understanding of economics, and their ability to appreciate the latest advances in mathematics and other sciences.

A minority of these ill-informed students themselves go on to be academic economists, and then repeat the process. Ignorance is perpetuated.

The attempt to conduct a critical dialogue within the profession of academic economics has therefore failed, not because neoclassical economics has no flaws, but because – figuratively speaking – neoclassical

economists have no ears. As Bernanke's reaction shows, even the global financial crisis wasn't enough to make them listen.

So then, 'No More Mr Nice Guy.' If economists can't be trusted to follow the Queensberry Rules of intellectual debate, then we critics have to step out of the boxing ring and into the streets.

Does economics matter?

Economists have been justly criticized for failing to anticipate the financial crisis, but if that had been their only failing, they would be no different to weather forecasters who failed to warn of a destructive storm. They could be at fault for failing to give the warning, but you couldn't blame them for the storm itself. Economics, on the other hand, has direct responsibility for the economic storm we are currently experiencing. This is not to say that capitalism is inherently stable – far from it. But the beliefs and actions of economists made this economic crisis far worse than it would have been without their interventions.

First, the naive theories they developed, especially in finance, encouraged reckless behavior in finance by their ex-students. More than a generation of business students were unleashed on the world who believed – or at least paid lip-service to – the fallacies that finance markets always price financial assets correctly, and that debt was good.

Secondly, economists also developed many of the tools of the financial trade that Warren Buffett so aptly described as 'weapons of financial mass destruction.' Options pricing models, 'value at risk' formulas and the like were all based on neoclassical economics, and many were developed by academic economists – some of whom received the Nobel Prize in Economics for their inventions.

Thirdly, probably their greatest negative contribution to human history was that, as regulators, they allowed the excesses of the finance sector to go on for perhaps two decades longer than would have occurred without their 'rescues.'

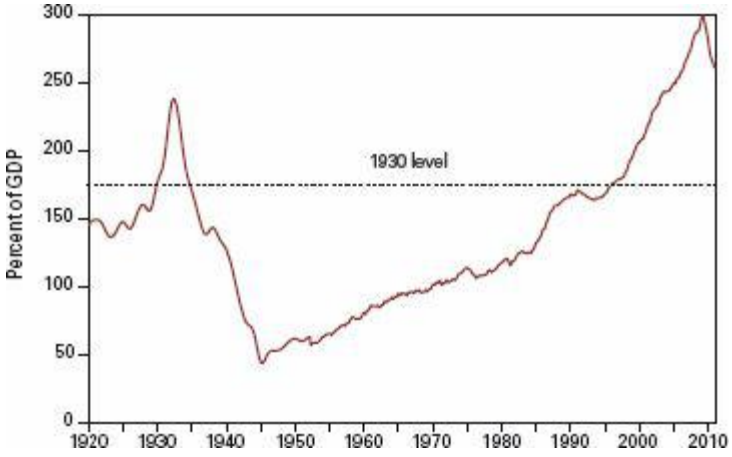
Here, pride of place goes to the central bankers – especially Alan Greenspan. In Chapter 12, I make the case that were it not for the extreme rescue efforts he initiated in 1987, the stock market crash of that year would have precipitated a serious recession, but one far milder than that we are

now experiencing. Instead, that rescue and the many others in the crises that followed – the Savings and Loans crisis, the Long Term Capital Management crisis, and finally the DotCom crisis – encouraged the speculative excesses of Wall Street to continue. The ultimate result was the subprime crisis, the fallout from which was so big that a further rescue was impossible.

The key indicator here – and the key reason that I and the others Bezemer identified as having predicted the crisis could tell that one was coming – is the ratio of private debt to national income (known as GDP, which stands for 'gross domestic product'). Every time the US Fed (and its counterparts in the rest of the OECD) rescued the financial sector from its latest folly, that sector continued doing what it is best at: creating debt.

If the Fed hadn't intervened in 1987, this process of escalating debt would probably have ended there, and America would have begun the painful but necessary process of deleveraging from a debt-to-GDP level of 160 percent – about 10 percent below the 175 percent level that precipitated the Great Depression – and in a milieu of moderate inflation.

Instead, rescued by the Fed, the financial sector lived to lend another day, and went through the veritable nine lives of the cat before the excesses of the Subprime Bubble brought Wall Street to its knees. By then, however, the debt ratio had risen to almost 300 percent of GDP – 1.7 times the 1930s level, and even 1.25 times the peak level of 235 percent of GDP achieved in 1932, when rampant deflation and plunging output drove the debt ratio higher even as Americans drastically reduced the nominal level of debt.



2.3 Private debt peaked at 1.7 times the 1930 level in 2009

By delaying the day of reckoning, neoclassical economists thus turned what could have been a 'run of the mill' financial crisis and recession into possibly the greatest capitalism will ever experience. The jury won't be in on the scale of 'The Great Recession' for several decades, but I expect that history will judge it to be more severe than the Great Depression – probably not in the depths of the downturn, but almost certainly in its duration and apparent intractability. It could not have got this bad without the assistance afforded by neoclassical economics.

Revolt

Bernanke's refusal to countenance that neoclassical economics could be flawed is indicative of the profession as a whole. The vast majority of neoclassical economists have sailed through the financial crisis and the Great Recession with their belief in neoclassical economics intact. If left to their own devices, economists will continue teaching that the economy is fundamentally stable, despite the abounding evidence that they are wrong.

The public could still afford to ignore economics if the discipline had the ability to correct its own excesses. But it does not. Despite its record at forecasting, despite the evidence that economic theories are not consistent, and despite the Great Recession that they have no choice but to admit they failed to foresee, the intellectual discipline of economics shows no tendency to reform itself. Instead, unsound theories continue to be taught to students as if they were incontrovertible. Economics cannot be trusted to reform its own house. Therefore, just as politics is too important to leave to the politicians, economics is too important to leave to the economists. The revolt against neoclassical economics has to go beyond the academic profession itself.

But it seems to make sense …

One of the great difficulties in convincing believers that neoclassical economics fundamentally misunderstands capitalism is that, at a superficial and individual level, it seems to make so much sense. This is one reason for the success of the plethora of books like *The Undercover Economist* (Harford 2005) and *Freakonomics* (Levitt and Dubner 2009) that apply economic thinking to everyday and individual issues: at an individual level,

the basic economic concepts of utility-maximizing and profit-maximizing behavior seem sound.

As I explain later, there are flaws with these ideas even at the individual level, but by and large they have more than a grain of wisdom at this level. Since they seem to make sense of the personal dilemmas we face, it is fairly easy to believe that they make sense at the level of society as well.

The reason this does not follow is that most economic phenomena at the social level – the level of markets and whole economies rather than individual consumers and producers – are 'emergent phenomena': they occur because of our interactions with each other – which neoclassical economics cannot describe – rather than because of our individual natures, which neoclassical economics seems to describe rather well.

The concept of emergent properties is a complex one, and I don't expect you to accept this argument right away; but as it happens, neoclassical economic theory provides an excellent example of an emergent phenomenon which I cover in Chapter 3 (and at the beginning of Chapter 10). Once you've read that, I think you'll understand why the fact that neoclassical economics seems sensible at the individual level has no bearing on whether it can make sense of capitalism itself.

Sincerity is no defense

Much – well, pretty much all – of what I have to say about neoclassical economics will be offensive to neoclassical economists.[11] Since this edition is far more likely than its predecessor to actually be read by some neoclassical economists, let me say now that I mean no personal offense. Pardon the cliché, but some of my best friends are neoclassical economists, and I've never for a second doubted the sincerity of most neoclassical economists. Though many in the public believe that neoclassical economists say what they say for personal gain, or to curry favor with the powers that be, the vast majority of neoclassical economists that I have met, or whose work I have read, are undoubtedly sincere in the belief that their work is intended to improve society as a whole, and not merely the situation of the powerful within it.

Unfortunately, as I learnt long ago, sincerity is no defense. A schoolteacher of mine put it this way in a discussion my class was having

about politics, when one student defended a particular politician with the statement 'Well, at least he's sincere!'

The class nodded sagely: yes, whatever we individually thought of this politician, we all had to concede that he was sincere. Our teacher, who normally let class discussions proceed unmonitored, suddenly piped up from the back of the room. 'Don't overrate sincerity,' he said. 'The most sincere person you'll ever meet is the maniac chasing you down the street with an ax, trying to chop your head off!'

I never did find out what personal experience led to that epiphany for Brother Gerard, but I've had many opportunities to reflect on its wisdom since: the most dangerous people on the planet are those who sincerely believe something that is false.

So while there is a mass of criticism of neoclassical economics – and of neoclassical economists for believing in it – I mean no offense to neoclassical economists as people. But as would-be scientists, their beliefs should not be provably false, as most of neoclassical economics is.

Debunking economics: a user's guide

Who is this book for? Interest in economics as an intellectual pursuit for its own sake has waned significantly over the last thirty years, and I have often heard academic economists lament this fact – especially since falling student enrollments have undermined their job security.

I am not at all amazed by this drop in interest: it is a predictable side effect of the very philosophy of life which neoclassical economists espouse.[12] They have told all and sundry that the world would be a better place if we all focused upon our own self-interest, and let the market take care of the common good. Why, then, is it surprising that students have swallowed this spiel, and decided to study subjects which more clearly lead to a well-paid job – business management, human resources, computing, etc. – rather than to study economics?

In its first incarnation in 2000, this book was directed at this audience, which economists once derided, and whose absence they now lament: people who are interested in 'the common good.' Its message, that the economic mantra ('individuals should pursue their own interests and leave society's overall interests to the market') is wrong, is not new. Many books

have made the same point in the past. What is new about this book is that it makes that point using economic theory itself.

In this second edition, I have an additional audience in mind: the professional economist who is honest enough to consider that perhaps the failure of the economics profession at large to anticipate the biggest economic event of the last seventy years could be due to deficiencies in the underlying theory itself. There will, I expect, be only a handful of such readers (and on current form, Ben Bernanke and Paul Krugman won't be among them), but if so they will be stunned at how much critical economic literature was omitted in their original education in economics. This book provides a compendium of that literature.[13]

I can guarantee that mainstream economists will hate the irreverent tone of this book. Nonetheless, I'd ask them to persevere with open – but skeptical – minds. I hope that exposure to the many published critiques of economics might explain to them why a theory which they accepted too uncritically was so manifestly unable to explain how a market economy actually behaves.

This book should also be useful to budding students of economics, in at least two ways. First, unless they are lucky enough to attend one of the few universities where pluralism rules, they are about to submit to an education in economics that is in reality an indoctrination. This book covers the issues which should form part of an education in economics, but which are omitted by the vast majority of textbooks.

Secondly, they should find that the explanations of economic theory in this book make it easier to pass exams in economics. I have found that one of the main barriers which new students face in learning economics sufficiently well to be able to pass exams in it is that they can't reconcile the theory with their own 'gut feelings' about economic issues. Once students realize that they should trust their gut feelings, and treat economic theory as irrelevant to the real economy, then suddenly it becomes much easier to pass exams. Just treat economics like a game of chess, play the games the exam questions require of you, and you'll pass easily (just don't mention the inconsistencies in the rules!).

If you are already a somewhat uncomfortable student of economics, but you lack confidence because you are surrounded by peers who can't

understand your disquiet, then this book should allay your fears. Normally, the journey from troubled student to informed critic is a difficult and lonely one. I hope to make that journey far less difficult, and less lonely. I hope it also gives you the confidence to confront your teachers if, while an economic crisis continues to rage about them in the real world, they continue teaching theories that argue that such things can't happen.

Similarly, I hope that professional critical economists will find this book a useful introductory compendium to those many critiques of economic theory that are currently scattered through dozens of books and hundreds of journal articles. While the arguments are not presented with the rigor of those formal critiques, the book provides an accessible and understandable introduction to that important and neglected literature. The curious student can be told to use this book as a guide before delving into the more difficult, formal literature.

Because it explains and debunks economic theory from first principles, this book will also be of use to anyone whose career makes them reliant upon the advice of economists. Hopefully it will encourage such people to look more widely for advice in future.

What's in this book? This book has been primarily written for people who are inclined to be critical of economics, but who are intimidated by its apparently impressive intellectual arsenal. I start from the premise that, though you might be familiar with the conclusions of economic theory, you are unfamiliar with how those conclusions were derived. You therefore don't have to have studied economics previously to be able to read this book.

I have also eschewed the use of mathematical formulas.[14] Though I frequently use mathematics in my own research, I'm well aware of the impact that mathematical symbols have on the intelligent lay reader (a Norwegian colleague calls it the MEGO effect: 'My Eyes Glaze Over.') Instead, where some mathematical concept is needed to understand a critique, I present it, as well as is possible, in verbal (and sometimes tabular) form.

Despite the absence of mathematics, this book will still require significant intellectual exertion by the reader. The arguments of economic theory are superficially appealing, as Veblen long ago observed. To

understand why they are nonetheless flawed requires thought at a deeper level than just that of surface appearances. I have attempted to make both economic theory and the flaws behind it relatively easy to comprehend, but there will be times when the difficulty of the material defeats my abilities as an expositor.

This problem is amplified by the fact that this book is effectively two books in one.

First, it provides a detailed exposition of the conventional theory, and takes none of the short cuts followed by the vast majority of conventional economic texts. As I noted above, one reason why economic instruction takes short cuts is because the foundations of conventional economics are not only difficult to grasp, but also profoundly boring. Economics should be an exciting, stimulating intellectual challenge, but conventional economics almost goes out of its way to be mundane. Unfortunately, I have to explain conventional economics in detail in order to be able to discuss the critiques of this theory. There are thus sections of this book which are inherently tedious – despite my attempts to lighten the discourse. This applies especially to the chapters on the neoclassical theories of consumption (Chapter 3) and production (Chapter 4).

Secondly, this book provides a detailed debunking of conventional theory. This is, I hope, rather more interesting than conventional theory itself – though nowhere near as interesting as an exposition of a truly relevant economics would be. But it is quite possible that the exposition of conventional theory which precedes each debunking may persuade you that the conventional economic argument makes sense. Your mind will therefore be tossed first one way and then the other, as you first grind through understanding the foundations of conventional economics, and then attempt to comprehend profound but subtle critiques of the superficially convincing conventional logic.

So, especially if you have never read a book on economic theory, you will undoubtedly find some sections very difficult. You may therefore find it easier to treat this book as a reference work, by reading Part 1 (Chapters 3–6) carefully, and then turning to the rest when you have some specific economic issue to explore. Alternatively, you can read the chapters in Parts 2 (Chapters 7–12) and 3 (Chapters 13–18) before you attempt the earlier, foundation ones. This is possible because in these later chapters I ‘cut

economics some slack,' and accept concepts which have in fact been debunked in the earlier chapters. After you've considered the failings of economics in these more interesting applied areas, you could then turn to the flaws in its foundations.

Whichever way you approach it, this book will be a difficult read. But if you are currently a skeptic of economics, and you wish to develop a deeper understanding of why you should be skeptical, I believe the effort will be worth it.

Not left versus right but right versus wrong One possible interpretation of this book – certainly one I expect to get from many economists – is that it is just a left-wing diatribe against rational economics. This common response to intellectual criticism – categorize it and then dismiss it out of hand – is one of the great sources of weakness in economics, and indeed much political debate.

It is probably true that the majority of those who criticize conventional economic theory are closer to the left than the right end of the political spectrum – though there are many profoundly right-wing critics of conventional economics. Only those occupying the middle of the political spectrum tend to espouse and implement conventional economics.

However, the critiques in this book are not based on politics, but on logic. No political position – left, right or middle – should be based on foundations which can easily be shown to be illogical. Yet much of conventional economic theory is illogical. Those who occupy the center stage of modern politics should find a firmer foundation for their politics than an illogical economic theory.

The same comment, of course, applies to those at the left-wing end of the political spectrum, who base their support for radical social change on conventional Marxian economics. As I argue in Chapter 17, conventional Marxism is as replete with logical errors as is neoclassical economics, even though Marx himself provides a far better foundation for economic analysis than did Walras or Marshall.

Escher without the panache One thing which sets economics apart from other social sciences, and which makes it hard for non-economists to understand economics, is the extent to which its arguments are presented in

the form of diagrams. Even leading economists, who develop their theories using mathematics, will often imagine their models in diagrammatic form.

These diagrams represent models which are supposed to be simplified but nonetheless accurate renditions of aspects of the real-world phenomena of production, distribution, exchange, consumption, and so on. When an economist talks of the economy behaving in a particular fashion, what he really means is that a model of the economy – and normally a graphical model – has those characteristics.

To learn economics, then, one has to learn how to read diagrams and interpret the models they represent. This applies to critics as much as believers, but the very act of learning the diagrams tends to separate one from the other. Most critical thinkers find the process tedious, and drop out of university courses in economics. Most of those who stay become seduced by the diagrams and models, to the point where they have a hard time distinguishing their models from reality.

The critical thinkers, who could not cope with the diagrammatic representation of economic reality, were fundamentally correct: economic reality cannot be shoehorned into diagrams. Consequently, these diagrams often contain outright fallacies, conveniently disguised by smooth but technically impossible lines and curves.

In other words, rather than being accurate renditions of the economy, the standard economic diagrams are rather like Escher drawings, in which the rules of perspective are used to render scenes which appear genuine – but which are clearly impossible in the real, three-dimensional world.

Whereas Escher amused and inspired with his endless staircases, eternal waterfalls and the like, economists believe that their models give meaningful insights into the real world. But they could only do so if the Escher-like assumptions economists make could apply in reality – if, metaphorically speaking, water could flow uphill. Since it cannot, economic models are dangerously misleading when used to determine real-world policy.

Obviously, therefore, I do not wish to encourage you to 'think diagrammatically,' since this mode of thought has helped to confuse economics rather than to inform it. However, to be able to understand where economics has gone wrong, you need to see what has led it astray. I have attempted to explain economic theory without diagrams, but it is still

probable that to be able to fully comprehend the fallacies in neoclassical economics, you will need to learn how to read diagrams – though not, I hope, to believe them (see 'Where are the diagrams?').

Blow by blow In most chapters, I take a key facet of economics, and first state the theory as it is believed by its adherents. I then point out the flaws in this superficially appealing theory – flaws that have been established by economists and, in most instances, published in economic journals. As I show, the effect of each flaw is normally to invalidate the theoretical point completely, yet in virtually every case, economics continues on as if the critique had never been made.

Economics is a moving target, and the outer edges of the theory sometimes bear little resemblance to what is taught at undergraduate level. Except in the case of macroeconomics, I concentrate upon the fare served up to undergraduates, rather than the rarefied extremities of new research – mainly because this is the level at which most economists operate, but also because much of the work done at the theoretical 'cutting edge' takes as sound the foundations learnt during undergraduate days. However, for some topics – notably macroeconomics – the difference between undergraduate and postgraduate economics is so extreme that I cover both topics.

The Great Recession has resulted in a much-expanded treatment of macroeconomics, and also two new chapters that 'break the mold' of the rest of the book by being expositions of my own approach to economics.

Chapter by chapter The book commences with two introductory chapters – which hopefully you have just read!:

- Chapter 1 ('Predicting the "unpredictable"') shows that the 'unpredictable' Great Recession was easily foreseeable almost a decade before it occurred.
- Chapter 2 ('No more Mr Nice Guy') gives an overview of the book.

Part 1, 'Foundations,' considers issues which form part of a standard education in economics – the theories of demand, supply, and income

distribution – and shows that these concepts have very rickety foundations. It has four chapters:

- Chapter 3 ('The calculus of hedonism') reveals that economics has failed to derive a coherent theory of consumer demand from its premise that people are no more than self-interested hedonists. As a result, economic theory can't justify a crucial and seemingly innocuous element of its analysis of markets – that demand for a product will fall smoothly as its price rises. Far from being innocuous, this failure cripples neoclassical theory, but neoclassical economists have both ignored this failure, and responded to it in ways that make a mockery of their claims to being scientific.
- Chapter 4 ('Size does matter') shows that the economic theory of 'the firm' is logically inconsistent. When the inconsistencies are removed, two of the central mantras of neoclassical economics – that 'price is set by supply and demand' and 'equating marginal cost and marginal revenue maximizes profits' are shown to be false. Economic theory also cannot distinguish between competitive firms and monopolies, despite its manifest preference for small competitive firms over large ones.
- Chapter 5 ('The price of everything and the value of nothing') argues that the theory of supply is also flawed, because the conditions which are needed to make the theory work are unlikely to apply in practice. The concept of diminishing marginal returns, which is essential to the theory, is unlikely to apply in practice, 'supply curves' are likely to be flat, or even downward-sloping, and the dynamic nature of actual economies means that the neoclassical rule for maximizing profit is even more incorrect than it was shown to be in the previous chapter.
- Chapter 6 ('To each according to his contribution') looks at the theory of the labor market. The theory essentially argues that wages in a market economy reflect workers' contributions to production. Flaws in the underlying theory imply that wages are not in fact based on merit, and that measures which economists argue would reduce unemployment may in fact increase it.

Part 2, 'Complexities,' considers issues which should be part of an education in economics, but which are either omitted entirely or trivialized in standard economics degrees. It has five chapters:

- Chapter 7 ('The holy war over capital') complements Chapter 5 by showing that the theory of capital is logically inconsistent. Profit does not reflect capital's contribution to output, and changing the price of capital relative to labor may have 'perverse' impacts on demand for these 'factors of production.'
- Chapter 8 ('There is madness in their method') examines methodology and finds that, contrary to what economists tell their students, assumptions do matter. What's more, the argument that they don't is actually a smokescreen for neoclassical economists – and especially journal editors, since they routinely reject papers that don't make the assumptions they insist upon.
- Chapter 9 ('Let's do the Time Warp again') discusses the validity of applying static (timeless) analysis to economics when the economy is clearly dynamic itself. The chapter argues that static economic analysis is invalid when applied to a dynamic economy, so that economic policy derived from static economic reasoning is likely to harm rather than help an actual economy.
- Chapter 10 ('Why they didn't see it coming') tracks the development of macroeconomics into its current sorry state, and argues that what has been derided as 'Keynesian' macroeconomics was in fact a travesty of Keynes's views. It explains the otherwise bizarre fact that the people who had the least inkling that a serious economic crisis was imminent in 2007 were the world's most respected economists, while only rebels and outsiders like myself raised the alarm.
- Chapter 11 ('The price is not right') deals with the economic theory of asset markets, known as the 'Efficient Markets Hypothesis'. It argues that the conditions needed to ensure what economists call market efficiency – which include that investors have identical, accurate expectations of the future, and equal access to unlimited credit –

cannot possibly apply in the real world. Finance markets cannot be efficient, and finance and debt do affect the real economy.

- Chapter 12 ('Misunderstanding the Great Depression and the Great Recession') returns to macroeconomics, and considers the dominant neoclassical explanation of the Great Depression – that it was all the fault of the Federal Reserve. The great irony of today's crisis is that the person most responsible for promoting this view is himself now chairman of the Federal Reserve.

Part 3, 'Alternatives,' considers alternative approaches to economics. It has six chapters:

- Chapter 13 ('Why I did see "It" coming') outlines Hyman Minsky's 'Financial Instability Hypothesis,' and my nonlinear and monetary models of it, which were the reason I anticipated this crisis, and why I went public with my warnings in late 2005.
- Chapter 14 ('A monetary model of capitalism') shows how a strictly monetary model of capitalism can be built remarkably simply, once all the factors that neoclassical theory ignores are incorporated: time and disequilibrium, and the institutional and social structure of capitalism.
- Chapter 15 ('Why stock markets crash') presents four non-equilibrium approaches to the analysis of asset markets, all of which indicate that finance destabilizes the real economy.
- Chapter 16 ('Don't shoot me, I'm only the piano') examines the role of mathematics in economic theory. It argues that mathematics itself is not to blame for the state of economics today, but instead that bad and inappropriate mathematics by economists has resulted in them persisting with an inappropriate static equilibrium analysis of the economy. The dynamic, non-equilibrium social system that is a market economy should be analyzed with dynamic, non-equilibrium tools.
- Chapter 17 ('Nothing to lose but their minds') dissects Marxian economics, arguing that this potential alternative to conventional economics is seriously flawed. However, much of the problem stems

from an inadequate understanding of Marx by not just his critics, but also his alleged friends.

- Finally, Chapter 18 ('There are alternatives') briefly presents several alternative schools in economics, and shows that viable if somewhat underdeveloped alternative ways to 'think economically' already exist.

There's even more on the Web This book does not begin and end with the chapters just mentioned. It is also intimately linked to one of my two websites, www.debunkingeconomics.com (my other website, www.debtdeflation.com, currently supports my blog on the financial crisis and ultimately will be the online companion to my next book, *Finance and Economic Breakdown*).

The website complements the book in several ways. First, sections of the argument have been placed on the Web. These are technically necessary, but somewhat tedious, and therefore could distract attention from key issues. These web entries are noted in the text with a comment like 'I've skipped explaining a concept called XX. Check the link More/XX if you want the full version,' which indicates both what has been placed on the Web, and where it is located.

Secondly, more lengthy discussion of some topics has been placed on the Web. For instance, the failure of the conventional theory of market demand means that alternative approaches must be developed. These, and additional critiques of conventional theory, are on the website and referred to under the heading 'But wait, there's more.' The locations of these additional discussions are given by comments like 'These and other issues are discussed on the Web. Follow the links to More/Hedonism.' These sections raise many issues which should be of interest to those critical of conventional economics.

Thirdly, while there are no mathematical formulas used in this book, the logic underlying many of the critiques is mathematical. The mathematically inclined reader can check the original logic by consulting the website. These links are indicated by a parenthetical statement such as '(follow the link Maths/Size/PC_eq_M for the maths).'

Fourthly, some related topics are not covered in the book. One obvious omission is the theory of international trade. The major reason for this omission is that, while sound critiques of international trade theory exist, what I regard as the most obvious and telling critique has not yet been formally developed (I outline this on the website at the link More/Trade, as well as discussing the formal critiques that have been published). Another reason is that the theory of international trade also depends on many basic concepts that are thoroughly debunked in this book.

Passing judgment on modern economics This book can be thought of as a critical report card on economics at the beginning of the third millennium. Economic theory, as we know it today, was born in the late nineteenth century in the work of Jevons, Walras, Menger and (somewhat later) Marshall. I have a reasonably high regard for these founders of what has become mainstream economics. They were pioneers in a new way of thinking, and yet, in contrast to their modern disciples, they were often aware of possible limitations of the theory they were trying to construct. They expected their heirs to extend the boundaries of economic analysis, and they expected economics to develop from the precocious but hobbled child to which they gave birth into a vibrant and flexible adult.

Instead, economics today is ridden with internal inconsistencies: an economic model will start with some key proposition, and then contradict that proposition at a later stage. For example, the theory of consumer demand begins with the proposition that each consumer is unique, but then reaches a logical impasse which it sidesteps by assuming that all consumers are identical.

This raises an important general point about scientific theories. Any theory will have some starting point, which can be chosen in any of a number of ways. Newtonian physics, for example, began with the starting point that any object subject to a force (in a vacuum) will accelerate; Einsteinian physics began with the starting point that the speed of light (also in a vacuum) sets an absolute speed limit for any material object.

Clearly the starting point of a theory can be challenged, but the basis of such a critique is normally what we might term ‘external consistency.’ That is, since the theory is supposed to describe some objective reality, it must be

possible to show significant consistency between the predictions of the theory and that objective reality.

Here the degree of proof often comes down to some statistical measure of accuracy. Using the example of physics again, it is obvious that, at the speeds which humans could impart to a physical body during the nineteenth century, the Newtonian vision was extremely accurate.

Internal consistency, on the other hand, requires that everything within the theory must legitimately follow from its starting point. Here, statistical accuracy is not good enough: the fit of the theory with the starting point from which it is derived must be exact. If a theory at some point requires a condition which contradicts its starting point, or any other aspect of itself, then the theory is internally inconsistent and therefore invalid. It is possible to criticize much of economics on the basis that 'reality isn't like that' – and this is occasionally done in the subsequent chapters. However, in general I take two allegedly related aspects of economic theory – the theory of individual consumption and the theory of the market demand curve, for example – and show that to get from one to the other, a clearly contradictory condition must be imposed.

A theory cannot survive with such contradictions – or rather, it should not. They are clear signals that something is fundamentally wrong with the starting position of the theory itself, and that real progress involves radically revising or even abandoning that starting point. Even some of the most committed economists have conceded that, if economics is to become less of a religion and more of a science, then the foundations of economics should be torn down and replaced. However, if left to its own devices, there is little doubt that the profession of academic economics would continue to build an apparently grand edifice upon rotten foundations.

The founding fathers of modern economics would, I expect, be surprised to find that a manner of thinking they thought would be transitional has instead become ossified as the only way one can do economics and be respectable. They would, I hope, be horrified to find that the limitations of economic theory have been soundly established, and that most 'respectable' economists nevertheless transgress these limits without conscience, and often without knowledge.

Respectability be damned. Like the populace watching the parade of the emperor, respectability has led us to kowtow to a monarch in fine cloth,

when an unindoctrinated child can see that the emperor has no clothes. It's time to expose the nakedness of neoclassical economics.

PART 1 | FOUNDATIONS

THE LOGICAL FLAWS IN THE KEY CONCEPTS OF CONVENTIONAL ECONOMICS

The belief that price and quantity are jointly determined by the interaction of supply and demand is perhaps the most central tenet of conventional economics. In Alfred Marshall's words, supply and demand are like the two blades of a pair of scissors: both are needed to do the job, and it's impossible to say that one or the other determines anything on its own. Demand for a commodity falls as its price rises, supply rises as price rises, and the intersection of the two curves determines both the quantity sold and the price.

This argument still forms the core of modern instruction in economics, and much of economic policy is directed at allowing these twin determinants to act freely and unfettered, so that economic efficiency can be at its maximum. But both mainstream and dissident economists have shown that the real world is not nearly so straightforward as Marshall's famous analogy. The next four chapters show that the 'blades of supply and demand' cannot work in the way economists believe.

3 | THE CALCULUS OF HEDONISM

Why the market demand curve is not downward-sloping

Maggie Thatcher's famous epithet that 'There is no such thing as society' succinctly expresses the neoclassical theory that the best social outcomes result from all individuals looking after their own self-interest: if individuals consider only their own well-being, the market will ensure that the welfare of all is maximized. This hedonistic, individualistic approach to analyzing society is a source of much of the popular opposition to economics. Surely, say the critics, people are more than just self-interested hedonists, and society is more than just the sum of the individuals in it?

Neoclassical economists will concede that their model does abstract from some of the subtler aspects of humanity and society. However, they assert that treating individuals as self-interested hedonists captures the essence of their economic behavior, while the collective economic behavior of society can be derived by summing the behavior of this self-interested multitude. The belief that the economic aspect of society is substantially more than the sum of its parts, they say, is misguided.

This is not true. Though mainstream economics began by assuming that this hedonistic, individualistic approach to analyzing consumer demand was intellectually sound, it ended up proving that it was not. The critics were right: society is more than the sum of its individual members, and a society's behavior cannot be modeled by simply adding up the behaviors of all the individuals in it. To see why the critics have been vindicated by economists, and yet economists still pretend that they won the argument, we have to take a trip down memory lane to late eighteenth-century England.

The kernel

Adam Smith's famous metaphor that a self-motivated individual is led by an 'invisible hand' to promote society's welfare asserts that self-centered behavior by individuals necessarily leads to the highest possible level of welfare for society as a whole. Modern economic theory has attempted, unsuccessfully, to prove this assertion. The attempted proof had several components, and in this chapter we check out the component which models how consumers decide which commodities to purchase.

According to economic theory, each consumer attempts to get the highest level of satisfaction he can from his income, and he does this by picking the combination of commodities he can afford which gives him the greatest personal pleasure. The economic model of how each individual does this is intellectually watertight.[1]

However, economists encountered fundamental difficulties in moving from the analysis of a solitary individual to the analysis of society, because they had to 'add up' the pleasure which consuming commodities gave to different individuals. Personal satisfaction is clearly a subjective thing, and there is no objective means by which one person's satisfaction can be added to another's. Any two people get different levels of satisfaction from consuming, for example, an extra banana, so that a change in the distribution of income which effectively took a banana from one person and gave it to another could result in a different level of social well-being.

Economists were therefore unable to prove their assertion, unless they could somehow show that altering the distribution of income did not alter social welfare. They worked out that two conditions were necessary for this to be true: (a) that all people have to have the same tastes; (b) that each person's tastes remain the same as his income changes, so that every additional dollar of income was spent exactly the same way as all previous dollars – for example, 20 cents per dollar on pizza, 10 cents per dollar on bananas, 40 cents per dollar on housing, etc.

The first assumption in fact amounts to assuming that there is only one person in society (or that society consists of a multitude of identical drones) – since how else could 'everybody' have the same tastes? The second amounts to assuming that there is only one commodity – since otherwise spending patterns would necessarily change as income rose. These 'assumptions' clearly contradict the case economists were trying to prove,

since they are necessarily violated in the real world – in fact, they are really a 'proof by contradiction' that Adam Smith's invisible hand doesn't work. Sadly, however, this is not how most economists have interpreted these results.

When conditions (a) and (b) are violated, as they must be in the real world, then several important concepts which are important to economists collapse. The key casualty here is the vision of demand for any product falling as its price rises. Economists can prove that 'the demand curve slopes downward in price' for a single individual and a single commodity. But in a society consisting of many different individuals with many different commodities, the 'market demand curve' can have any shape at all – so that sometimes demand will rise as a commodity's price rises, contradicting the 'Law of Demand.' An essential building block of the economic analysis of markets, the market demand curve, therefore does not have the characteristics needed for economic theory to be internally consistent.

The roadmap

The chapter opens with an outline of Jeremy Bentham's philosophy of utilitarianism, which is the philosophical foundation for the economic analysis of individual behavior. The conventional economic analysis is outlined. The chapter's punchline is that economic theory cannot derive a coherent analysis of market demand from its watertight but ponderous analysis of individual behavior. In addenda, I show that this analysis is only a toy model anyway – it can't apply to actual human behavior, and experimentally, it has been a failure.

Pleasure and pain

The true father of the proposition that people are motivated solely by self-interest is not Adam Smith, as is often believed, but his contemporary, Jeremy Bentham. With his philosophy of 'utilitarianism,' Bentham explained human behavior as the product of innate drives to seek pleasure and avoid pain. Bentham's cardinal proposition was that

> Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point

> out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne. They govern us in all that we do, in all we say, in all we think; every effort we can make to throw off our subjection, will serve but to demonstrate and confirm it. In a word a man may pretend to abjure their empire; but in reality he will remain subject to it all the while. (Bentham 1948 [1780])

Thus Bentham saw the pursuit of pleasure and the avoidance of pain as the underlying causes of everything done by humans, and phenomena such as a sense of right and wrong as merely the surface manifestations of this deeper power. You may do what you do superficially because you believe it to be right, but fundamentally you do it because it is the best strategy to gain pleasure and avoid pain. Similarly, when you refrain from other actions because you say they are immoral, you in reality mean that, for you, they lead to more pain than pleasure.

Today, economists similarly believe that they are modeling the deepest determinants of individual behavior, while their critics are merely operating at the level of surface phenomena. Behind apparent altruism, behind apparent selfless behavior, behind religious commitment, lies self-interested individualism.

Bentham called his philosophy the 'principle of utility' (ibid.), and he applied it to the community as well as the individual. Like his Tory disciple Maggie Thatcher some two centuries later, Bentham reduced society to a sum of individuals:

> The community is a fictitious body, composed of the individual persons who are considered as constituting as it were its members. The interests of the community then is, what? – the sum of the interests of the several members who compose it. It is in vain to talk of the interest of the community, without understanding what is in the interest of the individual. (Ibid.)

The interests of the community are therefore simply the sum of the interests of the individuals who comprise it, and Bentham perceived no

difficulty in performing this summation: 'An action then may be said to be conformable to the principle of utility when the tendency it has to augment the happiness of the community is greater than any it has to diminish it' (ibid.).

This last statement implies measurement, and Bentham was quite confident that individual pleasure and pain could be objectively measured, and in turn summed to divine the best course of collective action for that collection of individuals called society.[2] Bentham's attempts at such measurement look quaint indeed from a modern perspective, but from this quaint beginning economics has erected its complex mathematical model of human behavior. Economists use this model to explain everything from individual behavior, to market demand, to the representation of the interests of the entire community. However, as we shall shortly see, economists have shown that the model's validity terminates at the level of the single, solitary individual.

Flaws in the glass

In most chapters, the critique of conventional theory has been developed by critics of neoclassical economics, and neoclassical economists are unaware of it because, in general, they cope with criticism by ignoring it.

This isn't the case with this first critique because, ironically, it was an 'own goal': the people who proved that the theory was flawed were themselves leading neoclassical economists, who were hoping to prove that it was watertight.

It is not. While economics can provide a coherent analysis of the individual in its own terms, it is unable to extrapolate this to an analysis of the market.

Since this critique was developed by neoclassical economists themselves, many mainstream academic economists are aware of it, but they either pretend or truly believe that this failure can be managed with a couple of additional assumptions. Yet, as you'll see shortly, the assumptions themselves are so absurd that only someone with a grossly distorted sense of logic could accept them. That twisted logic is acquired in the course of a standard education in economics.

This ‘education’ begins with students being taught conclusions which would apply if the theory had no logical flaws. Students normally accept that these conclusions have been soundly derived from the basic economic propositions of individual behavior, and they are in no position to believe otherwise, since the basic building blocks of this analysis are not taught at the introductory level because they are ‘too hard.’ This abbreviated induction is sufficiently boring to dissuade the majority of business students from pursuing further economics, and they graduate in some other discipline. However, a minority find the game intriguing, and continue on to another year.

In later undergraduate years, they finally encounter indifference curves and the derivation of the individual demand curve. The mildly relevant ‘Engel curves’ and the complete chimera of the ‘Giffen good’ are explored as apparent applications of the theory. Market demand curves, and sometimes the basic concepts of ‘general equilibrium’ (the conditions under which many markets will simultaneously be in equilibrium), are discussed – again, without considering whether the step from the individual to the aggregate is valid.

Most economics graduates seek employment in the private sector, and parts of the public sector, where they normally champion the neoclassical perspective. However, a minority of this minority pursues further study, to seek employment as academic economists – and in search of education rather than remuneration, since academic salaries are far lower than private and even public sector ones. Once they have embarked upon this road to ordination as an economist, most students are fully inculcated in the neoclassical way of thinking.

Finally, in honors, master’s or PhD courses, they study the full exposition given below, and finally learn that the aggregation of individual demand is valid only under patently absurd conditions. However, by this time the indoctrination into the neoclassical mindset is so complete that most of them cannot see the absurdity. Instead, they accept these conditions as no more than simple devices to sidestep pesky but minor problems, so that ‘rational’ economic analysis can be undertaken.

It would be easy to accede to a simplistic conspiracy theory to explain why economic education takes such a convoluted route on this issue.

However, I believe the explanation is both more mundane and more profound.

At the mundane level, the proposition that individual behavior is motivated by utility maximization, the concept of a downward-sloping demand curve, and the vision of society as simply an aggregate of individuals are easier to grasp than the many qualifications which must be applied to keep these notions intact. Academic economists therefore instruct their students in the easy bits first, leaving the difficult grist for higher-level courses.

At the profound level, it reflects the extent to which economists are so committed to their preferred methodology that they ignore or trivialize points at which their analysis has fundamental weaknesses. Were economics truly worthy of the moniker 'social science' these failures would be reason to abandon the methodology and search for something sounder.

Whatever the reasons, this lazy pedagogy trifurcates economics students into three camps. The vast majority study a minimum of economics in a business degree, and graduate unaware of any flaws in the glass. Members of the second, much smaller group go on to professional academic careers, and treat the flaws as marks of a fine crystal, rather than clear evidence of a broken vessel. The third, a handful, become critics within the profession, who aspire to build more realistic theories and, sometimes, try to make the second group see the cracks in their beloved but broken goblet. These sentiments may appear extreme now, but I doubt that they will appear so by the time you have read this chapter.

Now pour yourself a strong cup of coffee – or any other appropriate stimulant. The next few sections are crucial to understanding both economic theory and its weaknesses, but they can't help but be boring.

'The sum of the interests'

Bentham's statement that 'The community is a fictitious body […] The interests of the community then is [sic] the sum of the interests of the several members who compose it' is no more than an assertion. To turn this into a theory, economists had to achieve two tasks: to express Bentham's analysis mathematically, and to establish mathematically that it was possible to derive social utility by aggregating individual utility.

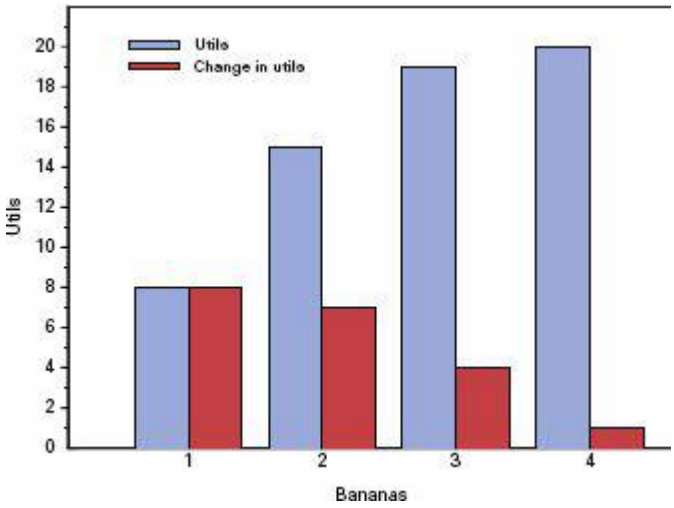
One century after Bentham, the founders of neoclassical economics accomplished the first task with relative ease. Over time, the representation of these concepts matured from simple but flawed notions to arcane but watertight models of individual behavior.

The individual consumer as represented by economic theory In keeping with the notion that beneath all individual actions lie the motivations of pleasure-seeking and pain avoidance, early attempts to use utility theory to explain behavior – by which economists meant almost exclusively the consumption of commodities[3] – postulated that each unit consumed of any commodity yielded a certain number of underlying units of satisfaction, called 'utils.' Additional units of a given commodity resulted in a smaller number of additional utils. The picture is as shown in Table 3.1.

TABLE 3.1 'Utils' and change in utils from consuming bananas

Bananas	Utils	Change in utils
1	8	8
2	15	7
3	19	4
4	20	1

For example, one unit of a commodity – say, a banana – yields 8 'utils' of satisfaction to the consumer. Two bananas yield 15 utils, so that the second banana has contributed seven additional utils to the consumer's satisfaction: one less than the first banana, but still a positive quantity. Three bananas yields 19 utils, so that the change in utils from consuming the third banana is 4 utils.



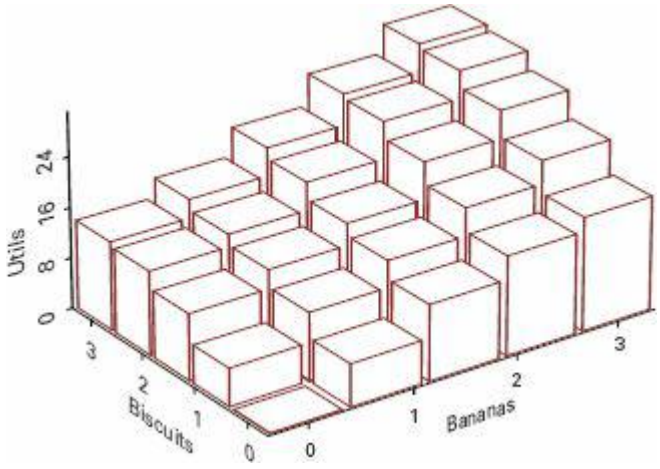
3.1 Rising total utils and falling marginal utils from consuming one commodity

This concept, that a consumer always derives positive utility from consuming something, but that the rate of increase in utility drops as more units of the commodity are consumed, is the key concept in the economic analysis of human behavior. The change in total utility is known as 'marginal utility,' and the essential belief that this falls as the level of consumption rises is known as the 'law of diminishing marginal utility.' This 'law' asserts that marginal utility is always positive, but always falling: more is always better, but each additional unit consumed gives less satisfaction than previous units.

Obviously, utility is derived from consuming more than just one commodity. Economists assume that the law of diminishing marginal utility applies across *all* commodities, so that additional units of any commodity give the consumer positive but falling amounts of utility. This is shown in Table 3.2, where the first commodity is bananas, and the second, biscuits. Each number in the table shows how many utils the consumer garnered from each combination of bananas and biscuits. Graphically, this yields a set of 3D bars, with the bars getting ever higher as more biscuits and bananas are consumed.

TABLE 3.2 Utils arising from the consumption of two commodities

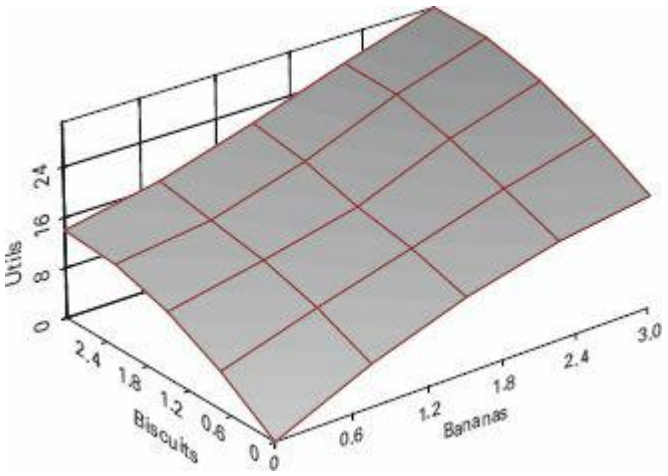
		BISCUITS			
		0	1	2	3
BANANAS	0	0	8	13	14
	1	9	13	15	18
	2	15	19	24	25
	3	18	25	30	31



3.2 Total utils from the consumption of two commodities

However, this representation is already clumsy. For a start, while it is possible to show the absolute number of utils given by any combination of bananas and biscuits, it is a cumbersome way to show the change in the number of utils caused by going from any one combination of biscuits and bananas to any other. Since marginal utility is a key concept, this was a major technical failing of this approach. It is also impossible to provide a geometric picture for more than two commodities.

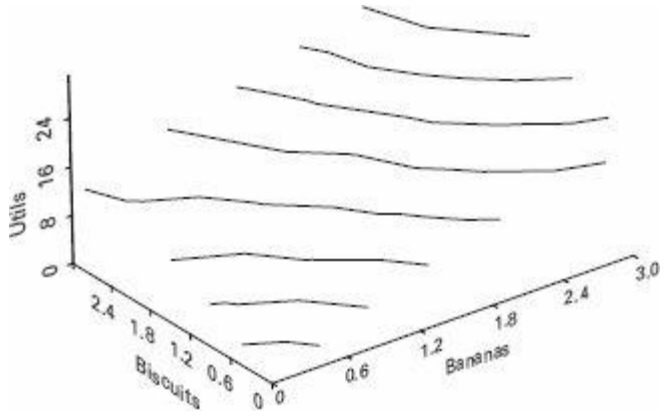
However, there is another, more obvious shortcoming. By postulating an objective measure of utility, it mooted an apparently impossible degree of precision and objectivity in the measurement of something so intrinsically subjective as personal satisfaction. As a result, the 'cardinal' concept of objectively measurable utility gave way to an 'ordinal'[4] notion, where all that could be said is that one combination of commodities gave more or less satisfaction than another combination.[5]



3.3 Total 'utils' represented as a 'utility hill'

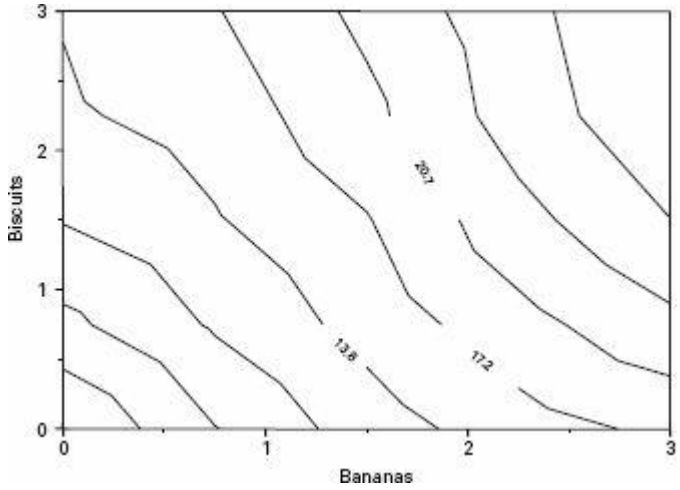
Metaphorically, this treated utility as a mountain, and the consumer as a mountain-climber whose objective was to get as high up this mountain as possible. The mountain itself was a peculiar one: first, it started at 'sea level' – zero consumption gave you zero utility – and then rose sharply because the first units consumed give you the highest 'marginal utility'; and secondly, it went on for ever – the more you consumed, the higher you got. The 'utility mountain' would get flatter as you consumed more, but it would never become completely flat since more consumption always increased your utility.[6]

The final abstraction en route to the modern theory was to drop this '3D' perspective – since the actual 'height' couldn't be specified numerically anyway – and to instead link points of equal 'utility height' into curves, just as contours on a geographic map indicate locations of equal height, or isobars on a weather chart indicate regions of equal pressure.[7]



3.4 The contours of the 'utility hill'

This representation enabled a conceptual advance which basically gave birth to modern consumer theory – however, as we shall see later, it also introduced an insurmountable intellectual dilemma. Since consumers were presumed to be motivated by the utility they gained from consumption, and points of equal utility height gave them the same satisfaction, then a consumer should be 'indifferent' between any two points on any given curve, since they both represent the same height, or degree of utility. These contours were therefore christened 'indifference curves.'



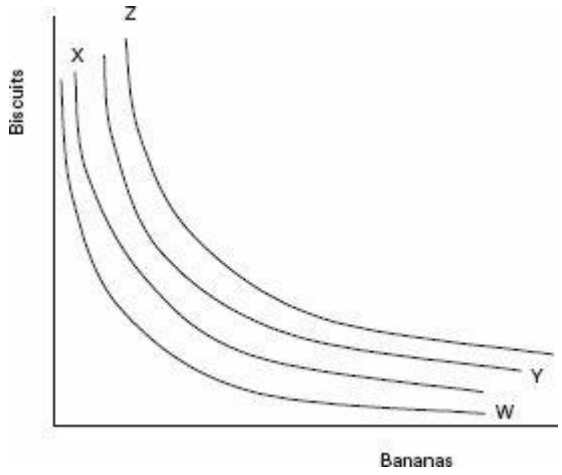
3.5 Indifference curves: the contours of the 'utility hill' shown in two dimensions

Since indifference curves were supposed to represent the innate preferences of a rational utility-maximizing consumer, economists turned their minds to what properties these curves could have if the consumer could be said to exhibit truly rational behavior – as neoclassical economists perceived it. In 1948, Paul Samuelson codified these into four principles:

- Completeness: If presented with a choice between two different combinations of goods, a consumer can decide which he prefers (or can decide that he gets the same degree of satisfaction from them, in which case he is said to be indifferent between them).
- Transitivity: If combination A is preferred to combination B, and B to C, then A is preferred to C.

- Non-satiation: More is always preferred to less. If combination A has as many of all but one commodity as B, and more of that one than B, then A is necessarily preferred to B.
- Convexity: The marginal utility a consumer gets from each commodity falls with additional units, so that indifference curves are convex in shape (shaped like a 'slippery dip').

This meant that indifference curves had a very specific shape: they had to look like a slippery dip that was steepest at its start, and always sloped downwards: the more of a good was consumed, the flatter the curve became, but it never became completely horizontal. And there were a multitude of such curves stacked on top of each other, with each higher one representing a higher degree of utility than the ones below. Economists then used these curves to derive the consumer's demand curve.



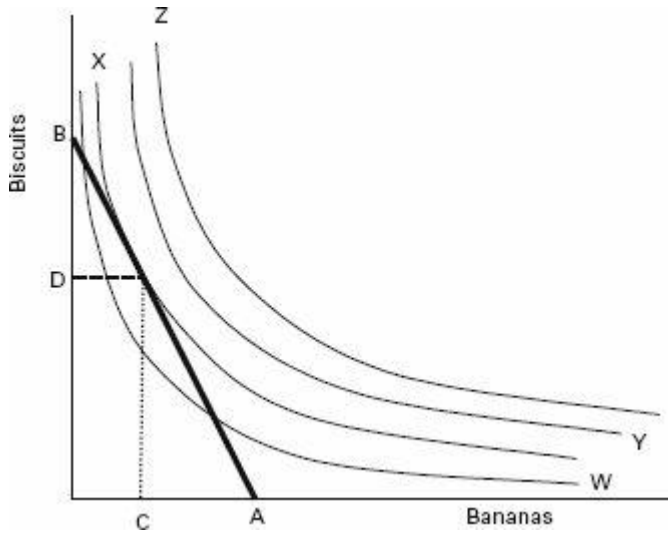
3.6 A rational consumer's indifference map

Deriving the individual demand curve

Obviously, since utility rises as more is consumed, the consumer would eat an infinite number of bananas and biscuits (yes, I know this is absurd) if not constrained by some other factors. The constraints are the consumer's income, and the prices of bananas and biscuits, so the next step in the economic saga is to notionally combine indifference curves with a consumer's income and prices to determine what a consumer will buy.

In terms of the 'utility mountain' analogy, this amounts to slicing the base of the mountain at an angle, where the slope of the slice represents the prices of biscuits and bananas, and cutting the mountain off at a distance that represents the consumer's income. There is now an obvious peak to the mountain, representing the highest point that the consumer can climb to.

In the '2D' model that economists actually use, the consumer's income is shown by a straight line which connects the quantity of bananas he could buy if he spent all his income on bananas, and the quantity of biscuits he could buy if he spent all his income on biscuits. If the consumer's income was $500, and biscuits cost 10 cents each, then he could purchase 5,000 biscuits; if bananas cost $1 each, then he could purchase 500 bananas. The budget line then connects these two points in a straight line – so that another feasible combination is 4,000 biscuits and 100 bananas.



3.7 Indifference curves, the budget constraint, and consumption

According to economists, a rational consumer would purchase the combination of biscuits and bananas which maximized his utility. This combination occurs where the budget line just touches a single indifference curve – in the 3D analogy, it's reaching the edge of the cliff at its highest point. If the consumer purchased any other feasible combination of biscuits and bananas using his income, then he would be forgoing some utility, which would be 'irrational.'[8]

Have you started to fall asleep yet? Sorry, but as I warned, this stuff is boring. But have some more coffee and stay tuned; after a few more

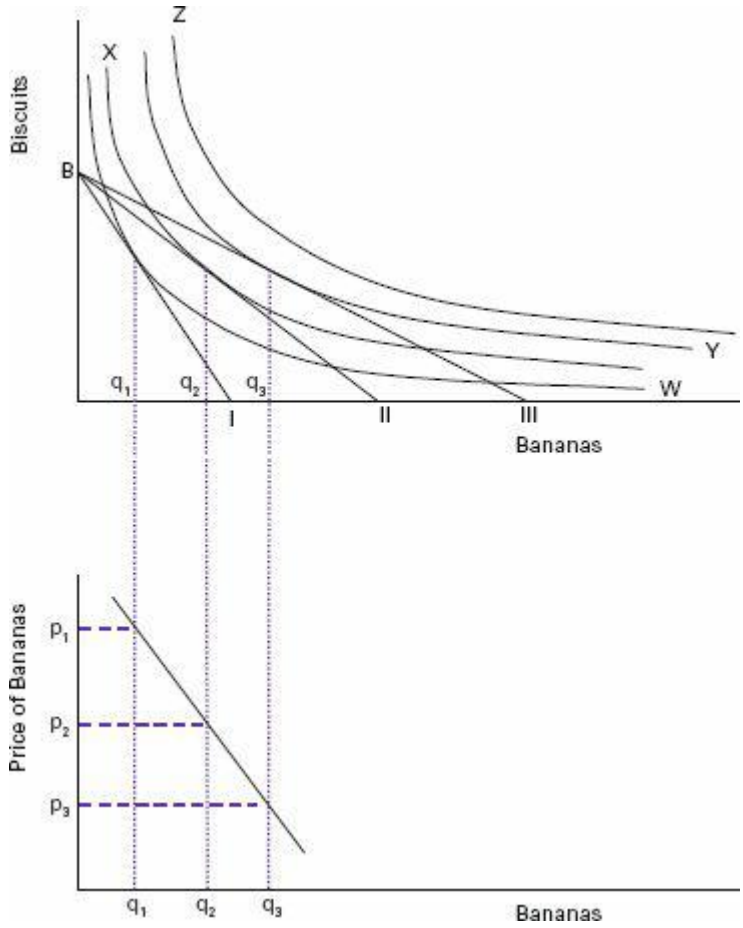
introductory bits, things start to get interesting.

The impact of changing prices on consumer demand

At this point, we have presented only the economic explanation of how a consumer will determine the consumption of any one bundle of commodities, given a fixed income and fixed prices. But what interests economists is what they call a 'demand curve,' which shows how demand for a commodity changes as its price changes, *while the consumer's income remains constant*.

This last condition is crucial – and as we'll see shortly, it is where the whole enterprise comes unstuck. Economists are trying here to separate how a consumer's behavior changes when prices change, from how behavior changes when incomes change. To do this, they have to assume that a change in prices won't change the consumer's income. This is OK if we're considering an isolated consumer who makes a living from, say, producing clothing: changing the price of bananas will have precious little impact on the income he makes from producing clothing.

If we consider a lower price for bananas, then the number of bananas the consumer can buy rises. If at the same time his income and the price of biscuits remain constant, then the budget line moves farther out on the bananas axis, but remains in the same spot on the biscuits axis. In the 3D analogy, this is like cutting a slice through the utility mountain at a different angle. The maximum point in the biscuits direction remains the same, but the maximum point in the bananas direction rises, and the overall hill is larger too – the consumer's maximum utility has risen because he can buy more.

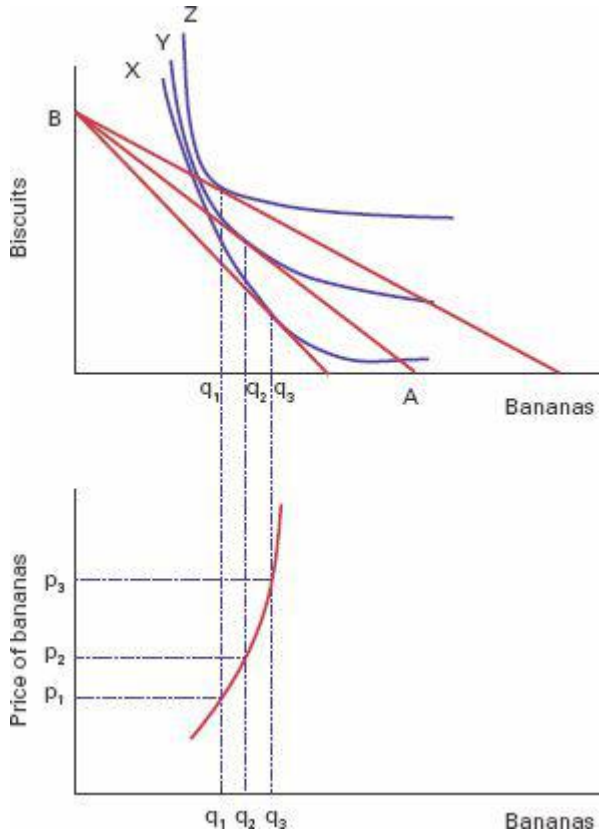


3.8 Deriving the demand curve

Economic theory then repeats this process numerous times – each time considering the same income and same price for biscuits, but a lower and lower price for bananas. Each time, there will be a new combination of biscuits and bananas that the consumer will buy, and the combination of the prices and quantities of bananas purchased is the consumer's demand curve for bananas. We finally have a demand curve, which *normally* slopes downwards as economists predicted. But it doesn't have to – there is still one wrinkle left. This is because, when the price of one good falls, and your income remains fixed, it's possible to increase the consumption of *all* goods – not just the one that has become cheaper.

It is even possible that your consumption of the good that has become cheaper could actually fall as its price falls, if it is so undesirable that you consume it simply because you are poor. Economists call such commodities 'Giffen Goods,' and their favorite alleged example is potatoes during the

potato famine in Ireland in the nineteenth century. They argue that as the price of potatoes rose during the famine, the Irish could no longer afford to buy more palatable goods like pork, so their consumption of potatoes actually rose as the famine continued and the price of potatoes also rose.



3.9 Upward-sloping demand curve

How's that coffee cup going? Empty? Then it's time you got a refill! There are two more tedious sections to come before the punchline that makes this banal trudge worthwhile.

Income and substitution effects and the 'Law of Demand'

The fact that a fall in price actually lets you consume more of everything can mean that it's possible for the demand curve for a given good to slope upwards at some points – to show the consumer consuming *less* as its price falls (and therefore more of it as its price rises!). This anomaly occurs because when the price of a commodity falls, the consumer's real income in effect increases.

This can be seen in our bananas and biscuits example: if the price of bananas falls while income and all other prices remain constant, then the consumer can buy more bananas without reducing his purchases of any other commodities. Therefore he is materially better off, even though his income hasn't changed.

This in turn can lead to perverse effects if one item in his shopping basket is relatively undesirable compared to more expensive alternatives – say, instant coffee rather than freshly ground beans – and it plays a large role in his budget. If the price of this commodity falls, it is possible the consumer could respond to the effective increase in income by consuming *less* of this product, even though it has become cheaper.

The increase in overall well-being due to the price of a commodity falling is known as the 'income effect.' It can lead you to consume more of the product, or it can lead you to consume less – it depends on the commodity. The pure impact of a fall in price for a commodity is known as the 'substitution effect.' So long as we are dealing with 'goods' – things which increase the consumer's utility – then the substitution effect is always going to be in the opposite direction to the change in price.

For this reason, economists say that the substitution effect is always negative. They don't mean that substitution is a bad thing, but that price and quantity move in opposite directions: if price falls, consumption rises. The income effect can be negative too – so that you consume *more* of a good as the *fall* in its price effectively increases your real income. But it can also be positive: you can consume *less* of a good when the *fall* in its price effectively increases your real income.

The always negative substitution effect is the phenomenon economists are trying to isolate with the demand curve, to establish what they call the 'Law of Demand' – that demand always increases when price falls. This 'law' is an essential element of the neoclassical model of how prices are set, which says that in competitive markets, supply will equal demand at the equilibrium price. For this model to work, it's vital that there is only one price at which that happens, so it's vital for the model that demand always increases as price falls (and similarly that supply always rises as price rises).

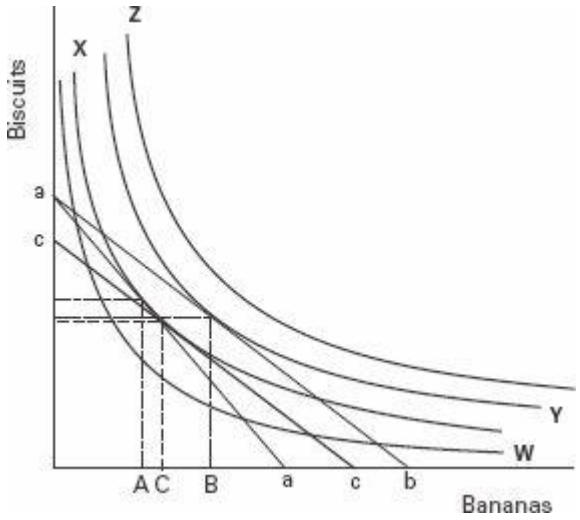
However the income effect can get in the way.

Economists thus found it necessary to search for a way to divide the impact of any change in price into the income effect and the substitution effect. If the income effect could be subtracted from a price change, this would leave the substitution effect as the pure impact on consumption of a change in relative prices. The problem is, though, that neither the 'income effect' nor the 'substitution effect' is directly observable: all we actually see is a consumer's purchases changing as the price of a commodity changes.

Economists dreamt up a way of at least notionally subtracting the income effect from a price change, using indifference curves. The clue is that, with income fixed and price falling, the lower price lets a consumer enjoy a higher effective standard of living – which in their model was manifested by the consumer reaching a higher indifference curve.

Since, to an economist, the real object of individual behavior is utility maximization, and since any point on a single indifference curve generates the same utility as any other point, then in utility terms the consumer's 'psychic income' is constant along this curve.

The substitution effect of a price fall could thus be isolated by 'holding the consumer's utility constant' by keeping him to the same indifference curve, and rotating the budget constraint to reflect the new relative price regime. This amounts to reducing the consumer's income until such time as he can achieve the same level of satisfaction as before, but with a different combination of biscuits and bananas. Then the budget constraint is moved out to restore the consumer's income to its actual level and, *voilà*, we have separated the impact of a price change into the substitution and income effects.



3.10 Separating out the substitution effect from the income effect

The demand curve derived from neutralizing the income effect is known as the 'Hicksian compensated demand curve,' after both the person who first dreamed it up (the English economist John Hicks) and the procedure used. It finally establishes the 'Law of Demand' for a single, isolated consumer: the demand for a commodity will rise if its price falls.

The dissident Australian economist Ted Wheelwright once described this hypothesized activity as 'tobogganing up and down your indifference curves until you disappear up your own abscissa,' and it's easy to see why.

Nonetheless, the end result is that desired by economists: increasing a product's price will reduce a consumer's demand for that product: an individual's demand curve slopes downwards. The 'Law of Demand' holds for a single consumer. There will be the odd commodity where a positive income effect outweighs the negative substitution effect, but these can be regarded as 'the exceptions that prove the rule' and safely ignored.

OK, take one more swig of coffee for the final tedious bit of detail – how economists consider the impact of changes in income on demand.

How rising income affects demand

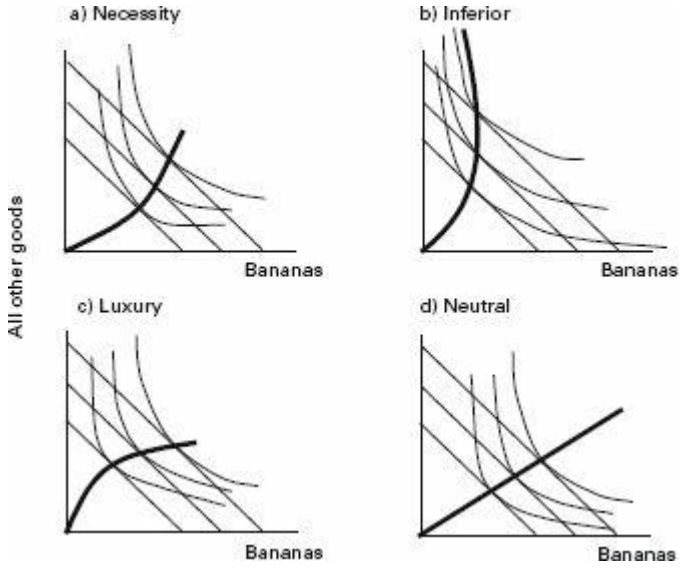
As with all other issues, economic theory uses indifference curves to handle this topic. The relevant commodity is placed on the horizontal axis, all other commodities on the vertical, and the budget constraint is 'moved out' (Figure 3.11). This represents an increase in income with relative prices held constant – unlike a pivot, which represents a change in relative

prices with income held constant. Economists say that the resulting plot – known as an 'Engel curve' – shows a consumer maximizing his utility as his income rises.

One point that is essential to the approaching critique is that Engel curves can take almost any shape at all. The shapes show how demand for a given commodity changes as a function of income, and four broad classes of commodities result: necessities or 'inferior goods,' which take up a diminishing share of spending as income grows; 'Giffen goods,' whose actual consumption declines as income rises; luxuries or 'superior goods,' whose consumption takes up an increasing share of income as it increases; and 'neutral' or 'homothetic' goods, where their consumption remains a constant proportion of income as income rises.

Necessities include such things as, for example, toilet paper. Your purchases of toilet paper will fall as a percentage of your total spending as you get wealthier (though you may buy more expensive paper). Some products that are substitutes for better-quality products when you are very poor – baked beans, perhaps – will disappear altogether from your consumption as you get wealthier, and economists refer to these as Giffen goods. Luxuries range from, for example, tourism to original works of art. Spending on holidays rises as income rises, and artworks are definitely the province of the rich.

I can't provide an example of a 'neutral good,' because strictly speaking, there are none. Spending on such a commodity would constitute the same percentage of income as a person rose from abject poverty to unimaginable wealth, and there is simply no commodity which occupies the same proportion of a homeless person's expenditure as it does of a billionaire's. But economists nonetheless have termed a word for someone whose preferences look like this: they call this pattern of consumption 'homothetic' (I call it 'neutral' in Figure 3.11d).



3.11 Engel curves show how spending patterns change with increases in income

Strictly speaking, no one could have homothetic preferences, and society in general would not display 'homothetic preferences' either: as income rose, the pattern of consumption of both individuals and society would change. Poor individuals and societies spend most of their money on staples (such as rice) while rich individuals and societies spend most of theirs on discretionary items (like the latest high-tech gadgets).

It may seem like explaining the obvious to say this, but the point is crucial to the approaching critique: as you age and as your income (hopefully) rises, your consumption pattern will change, as will the consumption pattern of a society as it gets richer. Thus any consumer is going to have lots of necessities and luxuries in his consumption, but no 'homothetic' goods.

Two is a crowd

The 'Law of Demand' has thus been proved – but only for a single consumer. Is it possible to generalize it so that it applies at the level of the market as well? In a nutshell, the answer is no. In the first of the many 'aggregation fallacies' that plague neoclassical economics, what applies when one consumer is isolated from all others does not apply when there is more than one consumer: what is true of Robinson Crusoe, so to speak, is not true of the society consisting of Robinson Crusoe and Man Friday.

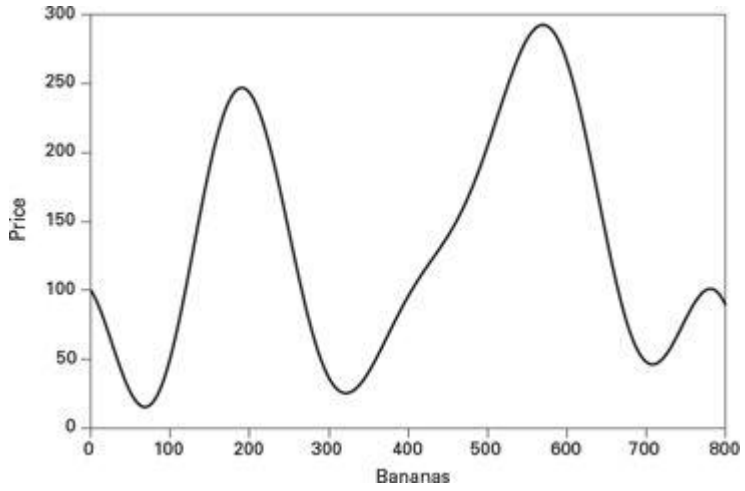
With Crusoe alone on his island, the distribution of income doesn't matter. But when Man Friday turns up, the distribution of income does matter, in ways that completely undermine everything involved in deriving an individual's demand curve.

One condition for deriving an individual's 'Hicksian compensated' demand curve for bananas was that changing the price of bananas didn't *directly* alter that individual's income.[9] That condition fails when you move from a one-person, two-commodity model to a two-person, two-commodity world – let alone anything more complicated – because changing the price of bananas (relative to biscuits) will alter the incomes of both individuals.

Unless they're clones of each other, one individual will earn more than the other from selling bananas – so an increase in the price of bananas makes the banana producer – let's call him Crusoe – richer, while making Friday poorer. This means that Crusoe is capable of buying more biscuits when the price of bananas rises. It's no longer possible to change the price of bananas while keeping constant the number of biscuits that the consumer can buy.

The complications don't stop there. Since the theory of the supply curve – which we'll encounter in the next two chapters – assumes that an increase in demand will drive up the price, the budget 'line' can't be a line: *it must be a curve*. In the isolated consumer example, not only did we assume that changing prices didn't alter the consumer's income, we also assumed that the consumer's purchases didn't affect the market price. This assumption is also invalid once we consider more than one consumer, which we must do to construct a market demand curve.

When Friday purchases the first banana, he pays a low price; the second banana costs more to produce (because of 'diminishing marginal productivity,' which we encounter in the next two chapters), so as well as his income changing, the price for bananas rises as his consumption of them rises. Each additional banana that Friday buys will therefore be more expensive than the previous one. The budget *curve* might start at the same point as the 'line' did (with an isolated consumer) when consumption is zero, but it must slope more steeply than the line as the consumer's consumption rises above zero.



3.12 A valid market demand curve

The situation is no better when we consider the demand that Crusoe has for the bananas he produces himself: his income rises as price rises, increasing his income, and his demand for bananas still drives the cost up – because according to the theory of the supply curve, the cost of production rises owing to falling productivity as output rises. There is no way to know which effect will dominate.

What was a straightforward exercise when each consumer was considered in isolation is therefore an unholy mess when we consider more than one individual, which we must do to derive a market demand curve. You can still derive points of tangency between these moving budget curves and the fixed indifference curves for each individual, and thus derive an individual demand curve, but it will no longer necessarily obey the 'Law' of Demand – and you can no longer easily separate the income and substitution effects either, since you cannot control incomes independently of prices anymore.

Finally, the market demand curve that is produced by summing these now poorly behaved individual demand curves will conflate these wildly varying influences: increasing price will favor the producer (thus increasing his demand) while disadvantaging the consumer (thus decreasing his demand); rising income for the luxury-good producer will increase his income while decreasing that of the necessity producer. As the sum of these tendencies, the market demand curve will thus occasionally show demand rising as price falls, but it will also occasionally show demand falling as

price falls. It will truly be a curve, because, as the neoclassical economists who first considered this issue proved (Gorman 1953), *it can take any shape at all* – except one that doubles back on itself.

Crucially, it can disobey the so-called 'Law of Demand': the quantity demanded can rise as the price rises. This has nothing to do with snob value, or price signaling quality, or any of the behavioral wrinkles that critics often throw at the assumptions that neoclassical economists make. The wavy demand curve shown in Figure 3.12 can be generated by ordinary, everyday commodities as soon as you move beyond the isolated individual.

This result – known as the 'Sonnenschein-Mantel-Debreu [SMD] conditions' – proves that the 'Law' of Demand does not apply to a market demand curve. If the market demand curve can have any shape at all, then there can be two or more possible demand levels for any given price, even if all consumers are rational utility maximizers who individually obey the Law of Demand. If only neoclassical economists had stated the result that honestly and accurately when it was first derived almost sixty years ago, economics today might be very different.

Instead, because the result was found by neoclassical economists who wished to prove the opposite of what they had in fact discovered, the result has been buried by a degree of obfuscation and evasion that makes the average corporate cover-up look tame by comparison.

Cut off at Pythagoras' pass

This result was first derived by neoclassical economists who had posed the question 'under what conditions will the market demand curve have the same properties as the individual demand curve?', and they were hardly pleased with their discovery. Though technically the analysis was a 'tour de force' – the sort of technical prowess that wins you awed respect from your peers – practically they clearly wished that they had proved the opposite result: that, despite the conundrums in moving from an isolated individual to multiple consumers, the Law of Demand still held.

They found themselves in the same situation as the ancient Pythagorean mathematicians, who believed that all numbers could be expressed as the ratio of two integers. The discovery that this was not the case 'destroyed

with one stroke the belief that everything could be expressed in integers, on which the whole Pythagorean philosophy up to then had been based' (Von Kurt 1945: 260).

Today, we're all familiar with the fact that if you draw two lines at right angles that are precisely one inch long, and draw a line between them, that line's length will be the square root of two inches long, which is an irrational number – a number that can't be expressed as the ratio of two integers. The fact that combining two rational numbers according to the laws of geometry generates an irrational number is now common knowledge. Neither mathematics nor the world has collapsed as a result – in fact both mathematics and the world are far richer for this discovery and the many that followed on from it.

However, the initial reaction of Pythagorean mathematicians to this discovery was brutal: they allegedly drowned Hippasus of Metapontum, who was the first to discover that irrational numbers existed. But to their credit, they subsequently embraced the existence of irrational numbers, and mathematics developed dramatically as a result.

Economists could have reacted intelligently to their discovery too. They had proved that if you take two consumers whose individual demand curves obey the Law of Demand, and add them together to get a market demand curve, that curve does not necessarily obey the Law of Demand. So adding two or more 'rational' consumers together generates an 'irrational' market. Therefore, market analysis has to transcend the simple rules that seemed to work for isolated consumers, just as mathematicians had to transcend the rules that apply when mathematical operations on rational numbers return only rational numbers.

Such a reaction by economists could have led to a far richer vision of economics than the simplistic one in which the Law of Demand applies, and in which all markets are assumed to be in equilibrium. Unfortunately, the way that they did react made the irate Pythagoreans who drowned Hippasus look like amateurs. Rather than drowning the discoverer of the result, neoclassical economists drowned the result itself.

Having proved that in general the 'Law of Demand' did not apply at the level of the market, they looked for the conditions under which it would apply, *and then assumed that those conditions applied to all markets*. It's as if Pythagoreans, on discovering that the square root of two was an irrational

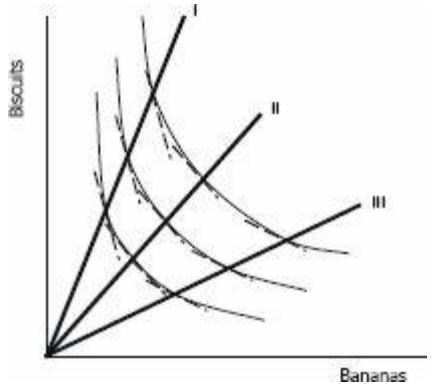
number, forbade for evermore the drawing of equal-sided right-angled triangles.

The Pythagorean analogy continues to apply here, because the conditions that were needed to 'ensure' that the Law of Demand applied at the market level are in fact a 'proof by contradiction' that it can't apply. Proof by contradiction is a venerable mathematical technique, and it can be used to establish that the square root of two is an irrational number. Not knowing the answer to a question – 'Is the square root of two a rational number?' – you assume that the answer is 'Yes,' and then follow through the logic of your assumption. If you generate a contradiction, you then know that the correct answer is 'No: the square root of two is not a rational number.'[10]

The two 'conditions' that economists found were necessary to guarantee that the 'Law of Demand' applied to the market demand curve were:

a) that all Engel curves are straight lines; and
b) that the Engel curves of all consumers are parallel to each other.

The first condition means that all commodities have to be neither luxuries nor necessities nor inferior goods, but 'neutral' or 'homothetic.' Therefore your ratios in which you consume different goods would have to remain fixed regardless of your income: if on an income of $100 a week, you spent $10 on pizza, then on an income of $100,000 a week you would have to spend $10,000 on pizza.



3.13 Straight-line Engel 'curves'

Clearly this is nonsense: as incomes rise, your consumption pattern would alter. There is only one situation in which this wouldn't apply: if there was only one commodity to consume. That is the real meaning of condition (a): there is only one commodity.

Condition (b) is just as absurd. For all consumers to have parallel Engel curves, all consumers have to have identical tastes. Clearly this is also nonsense: different consumers are identifiable by the very fact that they do have different tastes.

Even saying that the Engel curves of different consumers are parallel to each other is an obfuscation – it implies that two consumers could have parallel *but different* Engel curves, just as two lines that are parallel to each other but separated by an inch are clearly different lines. However, as anyone who has studied geometry at school knows, parallel lines that pass through the same point *are the same line*. Since a consumer with zero income consumes zero goods in neoclassical theory,[11] all Engel curves pass through the point 'zero bananas, zero biscuits' when income is zero. Therefore condition (b) really is that 'the Engel curves of all consumers are identical.'

There is only one situation in which this could apply: if there was only one consumer.

That is the real meaning of these two conditions: *the Law of Demand will apply if*, and only if, *there is only one commodity and only one consumer*. But in such a situation, the very idea of a 'Law of Demand' makes no sense. The whole purpose of the Law of Demand is to explain how relative prices are set, but if there is just one commodity and one consumer, then there can be no relative prices. We have a contradiction: we start from assuming that the Law of Demand applies, and then find that for this to be true, there can be only one commodity and one consumer – a situation in which the Law of Demand has no meaning.

These conditions are thus a proof by contradiction that the Law of Demand does not apply to the market demand curve: market demand does not necessarily increase when price falls, even if individual demand does.

This discovery is thus akin to the Pythagorean discovery of irrational numbers: adding together 'rational' consumers can result in an 'irrational' market. This discovery should have had an equally revolutionary – and

ultimately beneficial – impact upon economic theory. The simple parables of intersecting demand and supply curves would have had to give way to a more complicated but necessarily more realistic theory, in which prices would not be in equilibrium and the distribution of income would alter as prices alter.

If only.

Drowning the result

The economist who first discovered this result – the Hippasus of neoclassical economics – was William Gorman. As noted earlier, Hippasus was (allegedly) drowned for his trouble. Gorman, on the other hand, drowned his own result. He proved the result in the context of working out whether there was an economy-wide equivalent to an individual's indifference curves: 'we will show that there is just one community indifference locus through each point *if, and only if*, the Engel curves for different individuals at the same prices are parallel straight lines' (Gorman 1953: 63; emphasis added).

He then concluded, believe it or not, that these conditions were 'intuitively reasonable': 'The necessary and sufficient condition quoted above is intuitively reasonable. It says, in effect, that an extra unit of purchasing power should be spent in the same way no matter to whom it is given' (ibid.: 64).

'*Intuitively reasonable*'? As I frequently say to my own students, I couldn't make this stuff up! Far from being either intuitive or reasonable, Gorman's rationalization is a denial of one of the fundamental issues that most non-economists think economists must understand: the distribution of income. If the distribution of income changes, then surely the consumption pattern of society will change. I regard Gorman's statement here as the economic equivalent of the remark attributed to Marie Antoinette on being told that the peasants had no bread: 'Let them eat cake.'[12]

Gorman's original result, though published in a leading journal, was not noticed by economists in general – possibly because he was a precursor of the extremely mathematical economist who became commonplace after the 1970s but was a rarity in the 1950s. Only a handful of economists would have been capable of reading his paper back then. Consequently the result

was later rediscovered by a number of economists – hence its convoluted name as the 'Sonnenschein-Mantel-Debreu conditions.'

These economists were far less sanguine than Gorman about the 'conditions' needed for the Law of Demand to apply to a market demand curve. However, they still failed to make the logical leap to realize that they had disproved a core belief of neoclassical economics, and their statements of the result were, if anything, even more obtuse than was Gorman's: 'Can an arbitrary continuous function […] be an excess demand function for some commodity in a general equilibrium economy? […] we prove that every polynomial […] is an excess demand function for a specified commodity in some *n* commodity economy […] every continuous real-valued function is approximately an excess demand function' (Sonnenschein 1972: 549–50).

Translating this into English, a polynomial is a function consisting of constants and powers of some variable. The most well-known polynomials are the equation for a straight line, which is a polynomial of order one, and a parabola (a polynomial of order two). Any smooth curvy line that doesn't cross over itself can be fitted by a polynomial of sufficiently high order, so what Sonnenschein is saying here is that a demand curve can take any shape at all, except one that intersects with itself.[13] Therefore the 'Law of Demand' does *not* apply to the market demand curve. His joint summary of this result with Shafer for the encyclopedic *Handbook of Mathematical Economics* (Arrow et al. 1981–93) was more aware of the absurdity of the conditions, but still didn't connect the dots to comprehend that the conditions were a proof by contradiction that the Law of Demand is false:

> First, when preferences are homothetic and the distribution of income (value of wealth) is independent of prices, then the market demand function (market excess demand function) has all the properties of a consumer demand function […]
>
> Second, with general (in particular non-homothetic) preferences, even if the distribution of income is fixed, market demand functions need not satisfy in any way the classical restrictions which characterize consumer demand functions […]

> The importance of the above results is clear: strong restrictions are needed in order to justify the hypothesis that a market demand function has the characteristics of a consumer demand function. Only in special cases can an economy be expected to act as an 'idealized consumer.' The utility hypothesis tells us nothing about market demand unless it is augmented by additional requirements. (Shafer and Sonnenschein 1993)

As opaque as those statements might be, if they had been clearly passed on to economics students, the realization that the simple parables of supply and demand had to be replaced by something more sophisticated could have developed.

If only.

Don't tell the children

We now confront what will become a common theme in this book: the mendacious nature of economic textbooks. In the hands of economics textbook writers, the opaque but accurate statements of the SMD conditions above either disappear completely, or are portrayed in such a way that their significance will be perceived only by hypercritical students – like yours truly when I suffered through these courses while doing my Master's.

For many years, the leading text for Honors, Master's and PhD programs was Hal Varian's *Microeconomic Analysis* (Varian 1992). Varian 'summarized' this research so opaquely that it's no surprise that most PhD students – including those who later went on to write the next generation of undergraduate textbooks – didn't grasp how profoundly it challenged the foundations of neoclassical theory.

Varian started with the vaguest possible statement of the result: 'Unfortunately […] The aggregate demand function will in general possess no interesting properties […] Hence, the theory of the consumer places no restrictions on aggregate behavior in general.'

The statement 'no interesting properties' could imply to the average student that the market demand curve didn't differ in any substantive way from the individual demand curve – the exact opposite of the theoretical result. The next sentence was more honest, but rather than admitting

outright that this meant that the 'Law of Demand' didn't apply at the market level, he immediately reassured students that there was a way to get around this problem, which was to: 'Suppose that all individual consumers' indirect utility functions take the Gorman form [… where] the marginal propensity to consume good j *is independent of the level of income of any consumer and also constant across consumers* […] This demand function can in fact be *generated* by a representative consumer' (ibid.: 153–4; emphases added. Curiously the innocuous word 'generated' in this edition replaced the more loaded word 'rationalized' in the 1984 edition.)

Finally, when discussing aggregate demand, he made a vague and reassuring reference to more technical work: 'it is sometimes convenient to think of the aggregate demand as the demand of some "representative consumer" […] The conditions under which this can be done are rather stringent, but a discussion of this issue is beyond the scope of this book […]' (Varian 1984: 268).

It's little wonder that PhD students didn't realize that these conditions, rather than merely being 'rather stringent,' undermined the very foundations of neoclassical economics. They then went on to build 'representative agent' models of the macroeconomy in which the entire economy is modeled as a single consumer, believing that these models have been shown to be valid. In fact, the exact opposite is the case.

The modern replacement for Varian is Andreu Mas-Colell's hyper-mathematical – but utterly non-empirical – *Microeconomic Theory* (Mas-Colell, Whinston et al. 1995). At one level, this text is much more honest about the impact of the SMD conditions than was Varian's. In a section accurately described as 'Anything goes: the Sonnenschein-Mantel-Debreu Theorem,' Mas-Colell concludes that a market demand curve can have any shape at all, even when derived from consumers whose individual demand curves are downward-sloping:

> Can [… an arbitrary function] coincide with the excess demand function of an economy for every *p* [price …] Of course [… the arbitrary function] must be continuous, it must be homogeneous of degree zero, and it must satisfy Walras' law. But for any [arbitrary function] satisfying these three conditions, it turns out that the answer is, again, 'yes.' (Ibid.: 602)

But still, the import of this result is buried in what appear to the student to be difficult problems in mathematics, rather than a fundamental reason to abandon supply and demand analysis. Earlier, when considering whether a market demand curve can be derived, Mas-Colell begins with the question: 'When can we compute meaningful measures of aggregate welfare using […] the welfare measurement techniques […] for individual consumers? (ibid.: 116).

He then proves that this can be done when there is 'a fictional individual whose utility maximization problem when facing society's budget set would generate the economy's aggregate demand function' (ibid.: 116). However, for this to be possible, there must also exist a 'social welfare function' which: 'accurately expresses society's judgments on how individual utilities have to be compared to produce an ordering of possible social outcomes. We also assume that social welfare functions are increasing, concave, and whenever convenient, differentiable' (ibid.: 117).

This is already a case of assuming what you wish to prove – any form of social conflict is assumed away – but it's still not sufficient to generate the result Mas-Colell wants to arrive at. The problem is that the actual distribution of wealth and income in society will determine 'how individual utilities are compared' in the economy, and there is no guarantee that this will correspond to this 'social welfare function.'

The next step in his 'logic' should make the truly logical – and the true believers in economic freedom – recoil in horror, but it is in fact typical of the sorts of assumptions that neoclassical economists routinely make to try to keep their vision of a perfectly functioning market economy together. To ensure that the actual distribution of wealth and income matches the social welfare function, *Mas-Colell assumes the existence of a benevolent dictator who redistributes wealth and income prior to commerce taking place*: 'Let us now hypothesize that there is a process, *a benevolent central authority* perhaps, *that*, for any given prices p and aggregate wealth function w, *redistributes wealth in order to maximize social welfare*' (ibid.: 117; emphases added).

So free market capitalism will maximize social welfare *if, and only if, there is a benevolent dictator who redistributes wealth prior to trade*???

Why don't students in courses on advanced microeconomics simply walk out at this point?

I surmise that there are three main reasons, the first of which is banal. Mas-Colell's book is huge – just short of 1,000 pages – and lecturers would cherry-pick the sections they teach. I doubt that most students are exposed to this statement by their instructors, and few are likely to read parts that aren't required reading for pleasure alone.

Secondly, the entire text is presented as difficult exercises in applied mathematics. Students are probably so consumed with deriving the required answers that they gloss over English-language statements of these assumptions which make it blatantly obvious how insane they are.

Thirdly, by the time students get to this level – normally in PhD programs – they are so locked into the neoclassical 'assumptions don't matter' mindset that I discuss in Chapter 8 that they don't even worry if an assumption is insane.

From this bizarre point on, Mas-Colell, like Varian before him, encourages students to build models of the macroeconomy in which all agents have 'the Gorman form' of utility function – i.e. models of the macroeconomy in which there is one commodity and one consumer – so that students believe that the entire economy can be modeled as a single representative agent. Mas-Colell cautions that this involves a special assumption, but that caution is probably lost in the mist that envelops the mind of a budding neoclassical economist:

> If there is a normative representative consumer, the preferences of this consumer have welfare significance and the aggregate demand function can be used to make welfare judgments by means of the techniques [used for individual consumers]. In doing so however, *it should never be forgotten that a given wealth distribution rule* [imposed by the 'benevolent central authority'] *is being adhered to and that the 'level of wealth' should always be understood as the 'optimally distributed level of wealth.'* (Ibid.: 118; emphasis added)

These high-level texts, though, are at least honest that there is a problem in aggregating from the individual consumer to the market demand curve.

Undergraduate students instead are reassured that there is no problem. Paul Samuelson's iconic undergraduate textbook makes the following didactic statement about how a market demand curve is derived, and whether it obeys the 'Law of Demand,' which flatly contradicts the SMD results:

> The market demand curve is found by adding together the quantities demanded by all individuals at each price. *Does the market demand curve obey the law of downward-sloping demand? It certainly does*.
>
> If prices drop, for example, the lower prices attract new customers through the substitution effect. In addition, a price reduction will induce extra purchases of goods by existing consumers through both the income and the substitution effects. Conversely, a rise in the price of a good will cause some of us to buy less. (Samuelson and Nordhaus 2010: 48; emphasis added)

The leading undergraduate textbook today, by Gregory Mankiw, is equally misleading. It also implies that all that is needed to derive a market demand curve is to horizontally sum individual demand curves: 'The table in Figure 2 shows the demand schedules for ice cream for the two individuals in this market – Catherine and Nicholas […] The market demand at each price is the sum of the two individual demands […] Notice that we sum the individual demand curves horizontally to obtain the market demand curve […]' (Mankiw 2008: 68).

Other undergraduate textbooks either ignore the issue completely, or make similarly false statements. Who, then, can blame undergraduate economics students for believing that all is well with the underlying theory? The blame instead lies with textbook writers, and the question this raises is, *do they know they are at fault?* Did they knowingly conceal this advanced result from their students, or were they themselves ignorant of it?

Samuelson was certainly aware of Gorman's result, though he may not have followed the subsequent work of Sonnenschein and others because he believed he had proved that the Law of Demand *does* apply to the market demand curve (Samuelson 1956. And so he had – but using an assumption which shows how utterly unrealistic even the most famous of neoclassical

economists can be. He began quite sensibly, by noting that it was absurd to model an entire country as a single utility-maximizing individual:

> What defense do we make when challenged on the use of community indifference curves for a country or group of individuals? I suppose one of the following:
>
> (a) We may claim that our country is inhabited by Robinson Crusoe alone and claim only to show how trade between such single person countries is determined. This is admittedly not very realistic.
>
> (b) In order to give the appearance of being more realistic, we may claim that our country is inhabited by a number of identical individuals with identical tastes; they must also have identical initial endowments of goods if this artifice of examining what happens to the representative individual's indifference curves is to give us a true description of the resulting market equilibrium. This case, too, is not very realistic, though it may seem a slight improvement over Robinson Crusoe […]. (Ibid.: 3)

He then noted that most shopping is done by families, and since these consist of separate individuals, it is impossible even to construct a 'family indifference curve,' so that consumption by a family will also violate the foundations of the Law of Demand (the so-called Axioms of Revealed Preference, which are discussed in the addendum to this chapter).

However, he next surmised that *if, within the family, optimal transfers of income are undertaken*, then a family indifference curve *can* be constructed which has all the properties of an individual indifference curve.

> Since blood is thicker than water, the preferences of the different members are interrelated by what might be called a 'consensus' or 'social welfare function' which takes into account the deservingness or ethical worths of the consumption levels of each of the members. The family acts as if it were maximizing their joint welfare function […] *Income must always be reallocated among the members of our family society so as to keep the*

> *'marginal social significance of every dollar' equal*. (Ibid.: 10–11; emphasis added)

Finally, he hypothesized that if the entire nation behaves like one big happy family, and optimally reallocates income between its members prior to consumption, then society will also have 'well-behaved' indifference curves that obey the 'Law of Demand':

> The same argument will apply to all of society *if optimal reallocations of income can be assumed to keep the ethical worth of each person's marginal dollar equal*. By means of Hicks's composite commodity theorem and by other considerations, a rigorous proof is given that the newly defined social or community indifference contours have the regularity properties of ordinary individual preference contours (nonintersection, convexity to the origin, etc.). (Ibid.: 21; emphasis added)

Words fail me. Samuelson had 'proved' that social indifference curves exist – and therefore that market demand curves behave just like individual ones – by assuming that in a capitalist society, incomes are continuously adjusted so that an ethical distribution of income is achieved. *Did he even live in the United States?*[14] Yet on this basis, he confidently flourishes to his students that the market demand curve 'certainly does […] obey the law of downward-sloping demand.'

Samuelson's reason for perpetuating a falsehood is thus similar to Gorman's, who was capable of holding the equally delusional view that the proposition that 'an extra unit of purchasing power should be spent in the same way no matter to whom it is given' is 'intuitively reasonable.' So Samuelson, in a bizarre way, 'knew' what he was doing.

But in general I expect that the reason that undergraduate textbooks (written by lesser lights than Samuelson and Gorman) are so misleading is that *the authors themselves are unaware of this critical literature*.

This may seem bizarre: surely textbook writers must know the economic literature thoroughly in order to write a textbook in the first place? And haven't they done Master's and PhD courses, where they would at least have to read Varian or Mas-Colell on this topic?

Maybe. However, as I've pointed out above, the advanced textbooks present this result in such an obtuse way that it would be possible for a Mankiw to read this material, pass exams on it, and never even contemplate its true import. He might remember the 'Gorman form' limitation that had to be imposed to make aggregation possible, but he would probably regard this as just too difficult to teach to undergraduates. Undergraduate economic textbooks themselves have been 'dumbed down' so much in the last thirty years that even indifference curves – an essential element in this farce – are no longer taught in first-year courses. So the basics needed to even explain why there might be a problem are no longer part of the introductory pedagogy. Also, I expect that the Mankiws of the economics profession haven't read the original papers by Sonnenschein, Mantel and so on – and as I've noted, in a way they can't be criticized for this. Academics are accustomed to not having to read the original literature in their discipline, because they rely on their textbooks to accurately portray the key results of fundamental research. This belief is justified in physics – where even introductory texts point out that quantum mechanics and relativity can't be reconciled – but it is a false belief in economics.

Finally, in stark contrast to how a true science develops, this entire literature was developed not to explain an empirically observed phenomenon, but to examine the logical coherence of an utterly abstract, non-empirical model of consumer behavior. Downward-sloping demand curves were therefore not an empirical regularity for which a theory was needed, but a belief that economists had about the nature of demand that the vast majority of them took for granted. Most of them continue to hold this belief, unaware that mathematically erudite economists have shown that it is false. Since the underlying discipline is non-empirical, there is no disconnect between theory and reality that might warn them that something is wrong with the theory.

Worse still, the rationalization of a 'representative consumer' permeates modern economics – it has even taken over macroeconomic analysis, so that economists model an entire economy as if there is only one person in it (which they describe by the more general term of 'representative agent'). Many academic economists doubtless believe that the representative agent has been shown to be a valid abstraction. Yet far from being valid, it is in

fact a fudge, devised to get around the failure to prove that society can be reduced to the sum of its constituent individuals.

Following the madding crowd

There are many other reasons why economists did not recoil from the patent absurdities outlined above, and search for a sounder approach to economic theory than Bentham's individualistic calculus.

One is that economics has been wedded to the vision of society as simply a sum of utility-maximizing individuals since the inception of neoclassical economics in the 1870s. When the proof came, one century later, that this vision was internally inconsistent, the commitment to the vision was too strong to break. Better to search for special conditions which could let the theory survive – however ludicrous they might be – than to admit failure.

A second reason is that the peculiar language and mathematics used to derive these results makes it difficult to see just how absurd the assumptions needed to sustain the aggregation process are. It sounds much more highbrow to say that 'preferences are assumed to be homothetic and affine in income' than it does to say 'we assume all consumers are identical and never change their spending habits as their incomes increase.'

A third reason, perhaps the key one, is the division of mainstream economists into effective 'castes,' with only a tiny but exalted subset of the profession undertaking the detailed mathematical work needed to discover the weaknesses in the theory. The vast majority of economists believe that this high caste, the mathematical economists, did their work properly, and proved that the theory is internally consistent. The caste has indeed done its work properly, but it has proved precisely the opposite: that the theory is consistent only under the most restrictive and specious of assumptions.

However, rather than taking the next logical step, and acknowledging that the foundations of economics are unsound and must therefore be changed, most mathematical economists are so wedded to this way of thinking, and so ignorant of the real world, that they instead invent some fudge to disguise the gaping hole they have uncovered in the theory.

The majority of economists, blithely unaware of this state of affairs, then accept this fudge by the Brahmins of the profession as faithfully as

devout Hindus accept the cleansing properties of the Ganges river. As a result, the fudge then turns up in more mundane areas of economics, such as 'macroeconomics' (discussed in Chapter 10), where economists today analyze the economy as if it consisted solely of a single representative agent.

Consequently, these supposedly more practical theories can provide zip guidance in the serious business of managing a market economy. You would do as well to consult a Ouija board as an economist who rigorously follows economic theory when giving advice.

The Sonnenschein-Mantel-Debreu result is one of many that have effectively split the caste of mathematical economists into two sects. One pretends that business as usual can continue, despite the presence of this (and many other) fallacies in the creed. The other is dabbling in alternative religions – such as complexity theory, or evolutionary economics.

Sadly, the uninformed majority of the profession believes that the first sect is the bearer of the true religion, and that the members of the second sect have betrayed the faith. A more accurate analogy is that the dabblers in alternative religions are experiencing the first flushes of adolescence, while the majority of the profession remains mired in infancy. Clearly, the Benthamite ambition to portray society as simply an aggregate of its individual members is a failure. The whole is more than the sum of the parts.

The neoclassical rejoinder The great irony of this particular critique of economics is that it was constructed by its supporters. There is, as a result, no articulate rejoinder. Instead there are rationalizations, such as the 'representative agent' – which, as in Varian (1984), are often openly described as such.

If a defence were to be given of this practice, it would probably be what Samuelson termed 'the F-twist': that the assumptions of a theory don't matter; instead all that counts is how accurately a theory predicts reality. This popular but clearly invalid methodological defense is debunked in Chapter 8.

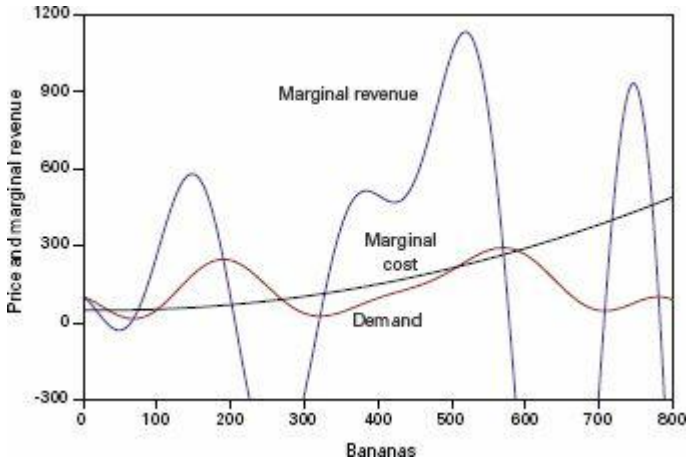
So what?

It might seem strange to make such a song and dance about whether market demand curves slope downwards. While economic theory clearly fails to prove that market demand falls smoothly as price rises, there are some sound reasons why demand might generally be a negative function of price. For example, a rise in the price of a commodity can force poorer consumers to substitute some cheaper alternative – or go without. So why does it matter that economists can't prove this?

First, it matters because economists had hoped to prove that a market economy necessarily maximizes social welfare. The SMD conditions establish that there is no measure of social welfare that is independent of the existing distribution of income, and that the distribution of income is not based solely on merit – it also reflects consumption patterns as well, since a change in consumption will alter the distribution of income.

Secondly, if we take the SMD conditions seriously, economic theory cannot rule out demand curves with a shape like that of Figure 3.12. Aesthetics aside, one of the many problems which such a curve presents for economic theory is that the resulting marginal revenue curve is even more volatile, and it can intersect the marginal cost curve (which we confront in the next chapter) in more than one place. This possibility undermines one of the key articles of the neoclassical faith, that 'everything happens in equilibrium.' If there are multiple points of intersection between marginal cost and marginal revenue, there will be multiple points where 'everything happens.' How then can you determine which will prevail in practice, let alone decide whether any one equilibrium is better or worse than any other?

These dilemmas flow from what appeared at the time to be a conceptual advance – dropping the fiction that utility could be measured in units akin to those we use to gauge weight, etc. While this was indeed more realistic, its interaction with two other aspects of economic theory made it impossible to aggregate the utility of two or more individuals.



3.14 Economic theory cannot rule out the possibility that a market demand curve may have a shape like this, rather than a smooth, downward-sloping curve

The culprits are the highly subjective nature of the concept of utility, and the belief that the price system determines income distribution. Since a change in relative prices will change the distribution of income, it therefore changes who consumes what, and hence the 'sum' of the subjective utility of all individuals. Since utility is subjective,[15] there is no way to determine whether one distribution of income generates more or less aggregate utility than any other.

Economists originally used this aspect of their theory to argue against social reformers who wished to redistribute income from the rich to the poor. They argued that such a redistribution might actually reduce social welfare by taking a unit of a commodity from a rich person who derived a great deal of utility out of it, and giving it to a poor person who derived very little utility from it.

It is ironic that this ancient defense of inequality ultimately backfires on economics, by making it impossible to construct a market demand curve which is independent of the distribution of income. If the market demand curve depends upon the distribution of income, if a change in prices will alter the distribution of income, and if this does not result in a single equilibrium between marginal revenue and marginal cost, then economics cannot defend any one distribution of income over any other. A redistribution of income that favors the poor over the rich cannot be formally opposed by economic theory – in fact, economic theory requires such a redistribution before it can even derive a market demand curve!

Finally, this failure rehabilitates the approach of classical economics to analyzing the economy. Classical economists such as Smith, Ricardo and Marx divided society into social classes, and considered how different policies might favor one social class over another. The notion of class has been expunged from economics by the concept of the indifference curve and its 'one size fits all' treatment of everyone from the poorest Somali to the richest American. Yet because the preferences of different individuals cannot be meaningfully aggregated, this concept is invalid for the analysis of anything more than an isolated individual.

But the conditions under which aggregation is valid – when tastes are identical and unaffected by changes in income – are at least reasonable as first approximations when the analysis splits society into different social classes. It is not too unreasonable to lump all workers, all landlords, and all capitalists together, as Smith, Ricardo and Marx used to do. Incomes within a class vary substantially less than incomes between classes, and tastes are far more likely to be common within classes than between them. A model with both Robinson Crusoe and Friday is at least slightly more reasonable than a model with Robinson Crusoe alone.

Leading mathematical economists have made very similar musings to this. Alan Kirman made one of the strongest such statements in his provocatively titled paper 'The intrinsic limits of modern economic theory: the emperor has no clothes.'[16] After discussing these and other theoretical failures of neoclassical economics, Kirman concluded that

> If we are to progress further we may well be forced to theories in terms of groups who have collectively coherent behavior. Thus demand and expenditure functions if they are to be set against reality must be defined at some reasonably high level of aggregation. The idea that we should start at the level of the isolated individual is one which we may well have to abandon. (Kirman 1989: 138)

In the end, then, the one benefit of neoclassical economics may be to have established why classical economists were correct to reason in terms of social class in the first place.

Addendum: an anti-empirical theory

There is one striking empirical fact about this whole literature, and that is that there is not one single empirical fact in it. The entire neoclassical theory of consumer behavior has been derived in 'armchair philosopher' mode, with an economist constructing a model of a hypothetical rational consumer in his head, and then deriving rules about how that hypothetical consumer must behave.

The aim of this armchair theorizing was to derive a watertight proof of market rationality from an underlying set of principles of rational individual behavior. The fact that this endeavor failed – that rational individual behavior can lead to an 'irrational' market – therefore means that the entire endeavor has been a waste of time. But many economists cling to this 'utility-maximizing' vision of how consumers behave because it seems so intuitively reasonable to them as a description of individual behavior.

Fittingly, this armchair theory has been proved to be empirically false by an experimental study. The experiment, by the German economist Reinhard Sippel, attempted to test the 'Axioms of Revealed Preference' that were developed by Paul Samuelson (Samuelson 1938a, 1938b) – one of the truly dominant figures in the development of neoclassical economics – as a way to derive a theory of consumer behavior in which utility did not need to be explicitly considered. Though this was not Samuelson's main intention, it also incidentally allowed the theory of utility maximizing behavior to be tested.

Samuelson defined a 'rational consumer' on the basis of how that consumer would behave when confronted with choices between bundles of goods, and he devised four rules to distinguish rational behavior from irrational: Completeness, Transitivity, Non-satiation and Convexity.

- *Completeness* meant that a rational consumer was able to compare different bundles of commodities – shopping trolleys containing different selections of goods from a supermarket – and decide which bundle he preferred. There were three possible outcomes: given a choice between the selection of goods in shopping trolley A and shopping trolley B, a rational consumer should be able to say that (a)

he preferred trolley A to trolley B; (b) that he preferred B to A; or (c) that he was indifferent between the two.

- *Transitivity* meant that if the consumer said he preferred trolley A to trolley B, and he also preferred trolley B to trolley C, then he necessarily had to prefer trolley A to trolley C.
- *Non-satiation* means that more is preferred to less. So if trolley B has the same contents as trolley A plus one additional chocolate bar, trolley B must be preferred to trolley A.
- Finally, the most complex property was *Convexity*, which is a mathematical expression of the concept of diminishing marginal utility. It argues that if you have two very different shopping trolleys, A and B, then any linear combination of the contents of these two trolleys should be preferred to the trolleys themselves. For example, imagine that trolley A contains ten chocolate bars and nothing else, while trolley B contains ten packs of chips and nothing else. Ten other shopping trolleys could be constructed by swapping one chocolate bar for one pack of chips, each of which would be more desirable than trolleys A and B.

These rules sound reasonable to most people when first explained to them – like many concepts in neoclassical economics, they are superficially appealing – but Sippel's experiment concluded that, if obeying these rules makes one rational, then the vast majority of us are irrational.

Sippel tested the theory in a very systematic way. He gave his student subjects a set of eight commodities from which to choose (see Table 3.3), a budget line, and a set of relative prices. This was repeated ten times, with each of the ten different price and budget line combinations being designed to test various aspects of Revealed Preference. Subjects were given as much time as they liked to make their choices, and after the ten tests, they got to consume one of the bundles they had selected.

I expect that Sippel conducted the experiment in order to confirm the theory. I would not be surprised to find that his intention was to use the results to derive 'indifference curves' for each of his subjects, and thus confirm that economic theory accurately described their behavior. But the results were a surprise: eleven of his twelve subjects failed the test of

rationality! He repeated it with a larger group of thirty – to find that twenty-two of these were also 'irrational' according to Samuelson's definition of rational behavior.

Sippel then tried to rescue the theory in a number of ways, none of which worked. One of the most ingenious methods was to hypothesize that real-world consumers can't as easily distinguish the utility they get from different bundles of goods, by assuming that indifference curves were 'thicker' than the thin lines drawn in neoclassical textbooks. This did indeed reduce the number of violations of the 'Axioms of Revealed Preference'; but it also had the undesirable impact that it made random choice – simply choosing what to consume by rolling dice – appear more rational than the consumption decisions of his students!

TABLE 3.3 The commodities in Sippel's 'Revealed Preference' experiment

Good	Max. amount (if all budget spent on one good)
Video clips	30–60 minutes
Computer games	27.5–60 minutes
Magazines	30–60 minutes
Coca-Cola	400ml–2 liters
Orange juice	400ml–2 liters
Coffee	600ml–2 liters
Candy	400gm–2 kilos
Pretzels, peanuts	600gm–2 kilos

To his great credit, Sippel concluded with an understated but accurate reflection on the implications of his experiment for economic theory:

> We conclude that the evidence for the utility maximization hypothesis is at best mixed. While there are subjects who appear to be optimizing, the majority of them do not. The high power of our test might explain why our conclusions differ from those of other studies where optimizing behavior was found to be an almost universal principle applying to humans and non-humans as well. In contrast to this, we would like to stress the diversity of individual behavior and call the universality of the maximizing principle into question […]

> We find a considerable number of violations of the revealed preference axioms, which contradicts the neoclassical theory of the consumer maximizing utility subject to a given budget constraint. We should therefore pay closer attention to the limits of this theory as a description of how people actually behave, i.e. as a positive theory of consumer behavior. Recognizing these limits, we economists should perhaps be a little more modest in our 'imperialist ambitions' of explaining non-market behavior by economic principles. (Sippel 1997: 1442–3)

Sippel did not speculate as to what his subjects were actually doing if they weren't in fact maximizing their utility, but it is fairly easy to show that these subjects were behaving rationally in the face of a real-world phenomenon of which armchair economic theorists are blithely unaware: the 'curse of dimensionality.'

Rational behavior and the curse of dimensionality

The neoclassical definition of rational behavior argues that a rational person, when confronted with a set of options, will attempt to choose the best option available. It appeared to Sippel that this was exactly what his subjects were doing:

> A closer look at the actual demand data corroborates the view that the subjects did not choose randomly. Every subject showed a marked preference for some of the goods while other goods were not chosen at all, even at low prices. Some subjects' demand was quite price inelastic, whereas others substituted cheaper goods for their more expensive counterparts, e.g. Coke for orange juice, sometimes to the extent that they always switched from one to the other, depending upon which was cheaper in the particular situation. There can be no doubt that the subjects tried to select a combination of goods that came as close as possible to what they really liked to consume given the respective budget constraints. (Ibid.: 1439)

However, despite this intention to choose the best option, they failed to do so rationally according to Samuelson's rules. So what's at fault – human behavior, or the neoclassical model of rationality?

The latter, of course. It is a 'toy' model that looks OK on paper, but fails completely when one takes even a tiny step into the real world – as Sippel's experiment did.

Let's look at what his subjects were being asked to do more closely. Sippel gave them a choice between eight different commodities, and let them choose any amount of them that they could afford with their budget. How many different 'shopping trolleys' could this mean they were looking at – each containing a different combination of goods?

Unfortunately, the answer is 'an infinite number of shopping trolleys,' so let's simplify it and imagine that students considered their choices in discrete units – say 5-minute segments for the videos and computer games (30 minutes, 35 minutes, and so on out to 60 minutes), 250ml units of drinks (400ml, 650ml, out to 2 liters), and 250 gram units of sweets (400 grams, 650 grams, out to 2 kilos). This means roughly eight different quantities for each of the eight goods. How many different shopping trolleys does that give us?

The answer will probably surprise you: *you could fill over 16.7 million shopping trolleys with different combinations of these eight goods*. Sixty-four would contain varying amounts of only one good – from 30 to 60 minutes of video, from 400 grams to 2 kilos of candy. The other 16.7 million-plus would have varying combinations of all the goods available.

This is a consequence of the real-world phenomenon that computer scientists have dubbed 'the curse of dimensionality.' The standard neoclassical 'toy' model of consumption shows you choosing between two different commodities. Most of these drawings don't show quantities on their axes, but if the quantities being considered were between zero and ten units of each good, then there would be 121 different combinations you could choose: zero units of both ([0,0]), ten units of both ([10,10]), and another 119 combinations in addition to that ([0,1], [1,0] right out to [10,9] and [9,10]).

The general rule for choices involving many commodities is that the number of different combinations equals one plus the number of units that

you could buy of each commodity,[17] raised to the power of the number of commodities you are considering. In the simple two-commodity case, this results in 11-squared choices – or 121. Your budget might allow you to rule out 90 percent of these, leaving just 10 or so choices to consider.

In Sippel's experiment, however, this resulted in 8 raised to the power of 8 – or in longhand 8 by 8 by 8 by 8 by 8 by 8 by 8 by 8 by 8, which equals 16.7 million.[18] Many of these 16.7 million combinations would be ruled out by the budget – the trolley containing the maximum amount of each item is clearly unattainable, as are many others. But even if the budget ruled out 99.99 percent of the options – for being either too expensive or too cheap compared to the budget – there would still be over 1,600 different shopping trolleys that Sippel's subjects had to choose between every time.

The neoclassical definition of rationality requires that, when confronted with this amount of choice, the consumer's choices are consistent every time. So if you choose trolley number 1355 on one occasion when trolley 563 was also feasible, and on a second occasion you reversed your choice, then according to neoclassical theory, you are 'irrational.'

Nonsense. The real irrationality lies in imagining that any sentient being could make the number of comparisons needed to choose the optimal combination in finite time. The weakness in the neoclassical vision of reality starts with the very first principle of 'Completeness': it is simply impossible to hold in your head – or any other data storage device – a complete set of preferences for the bewildering array of combinations one can form from the myriad range of commodities that confront the average Western shopper. With this principle being impossible, any sane person's shopping behavior will certainly also violate the neoclassical rules of Transitivity and Convexity (and probably Non-satiation as well). But it will be because the neoclassical principles themselves are irrational, not because the shopper is.

Consider, for example, your regular visit to a supermarket. The typical supermarket has between 10,000 and 50,000 items, but let's segment them into just 100 different groups. How many different shopping trolleys could you fill if you limited your decision to simply whether to buy or not buy one item from each group?

You would be able to fill two to the power of one hundred shopping trolleys with different combinations of these goods: that's 1,267,650,600,228,229,401,496,703,205,376 trolleys in total, or in words over 1,000 million trillion trillion shopping trolleys. If you could work out the utility you gained from each trolley at a rate of 10 trillion trolleys per second, it would take you 100 billion years to locate the optimal one.

Obviously you don't do that when you go shopping. Instead, what you do is use a range of commonplace heuristics to reduce the overwhelming array of choices you face to something manageable that you can complete in less than an hour. You partition your choices into a few basic groups, rather than looking at every separate product; and within the groups you use habit to guide your purchases – if you normally have muesli for breakfast, you ignore cornflakes. Truly rational behavior is therefore not choosing the best option, but *reducing the number of options you consider so that you can make a satisfactory decision in finite time*.

This is a commonplace observation in computer science, which unlike economics has built its knowledge of how decisions are made from experimentation and experience. What are sometimes called the 'Laws of Computational Theory' put front and center – in a rather paradoxical way – the fact that most real-world problems have so many potential solutions that an optimum cannot be found:

1 You cannot compute nearly all the things you want to compute.

2 The things you can compute are too expensive to compute. (Ballard 2000: 6)

The first law reflects research by Turing which established that most logical problems cannot be solved by a computer program. The second states that for the minority of problems that can be solved, the 'Curse of Dimensionality' means that an optimum solution cannot be found in finite time, no matter how much computing power is thrown at it. Computer scientists are much more informed than economists about the capacity of any reasoning system to solve even the simplest problems, and they are much more cautious as a result.

Economists should respect their greater knowledge, and accept that individual behavior will be 'satisficing' in nature rather than optimizing, as

the behavioral economist Herbert Simon put it (Simon 1996).

Conclusion

There are of course reasonable grounds to expect that, for many commodities, demand will rise as price falls. One, given by the marketer and statistician Andrew Ehrenberg, was that consumers allocated a fairly constant percentage of their spending to different classes of commodities (shelter, food, clothing, etc.) and a fall in the price of any item within a class resulted in an increase in the purchases of it, though very little change in the aggregate amount spent on that class of commodities overall (Ehrenberg 1975: 275–9).

This empirical reality cannot, however, rescue the neoclassical theory of consumer behavior from its result that market demand curves derived from consumers having 'rational' preferences (as neoclassical theory defines them) can have any shape at all. As I note later, this is an example of what is known in complexity theory as 'Emergent Behavior' – that the behavior of the sum of a set of isolated individuals cannot be deduced from the behavior of any of them in isolation.

This could imply that the best research strategy to develop economics is to abandon the model of rational behavior – as neoclassical economics defines it – and adopt the behavioral perspective of satisficing or bounded rationality instead.

I am more inclined to take Alan Kirman's lead here: that the failure of the endeavor to derive market rationality from individual rationality implies that the whole agenda of trying to derive systemic economic laws from the analysis of the isolated individual – known as 'methodological individualism' – is a waste of time. Instead, as Kirman put it, 'If we are to progress further we may well be forced to theories in terms of groups who have collectively coherent behavior' (Kirman 1989: 138). This implies that the old classical economics focus on social classes as the ideal level of analysis was correct – even if many of the conclusions derived from that in the 1800s were false.

That is the approach I take to macroeconomic modeling – as you will see in Chapters 13 and 14. I am inclined to leave studies of satisficing behavior to the psychologists.

So one half of the iconic ‘supply and demand’ model is unsound: what about the other half, the supply curve?

4 | SIZE DOES MATTER

Why there is no supply curve

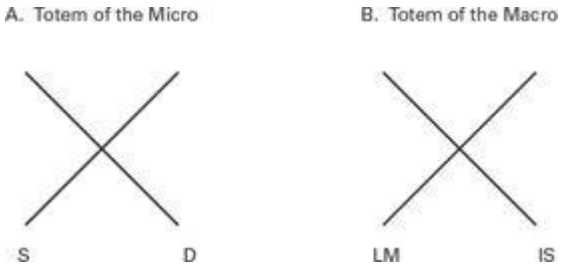
The image of one downward-sloping line intersecting with another upward-sloping one to determine an equilibrium is so iconic to neoclassical economics that a renowned wit once described it as the 'Totem' of economics. In a wonderful satire entitled 'Life among the Econ,' Swedish economist Axel Leijonhufvud imagined himself as an anthropologist investigating academic economists, whom he portrayed as a tribe living in the cold and arid Arctic: 'The Econ tribe occupies a vast territory in the far North. Their land appears bleak and dismal to the outsider, and travelling through it makes for rough sledding; but the Econ, through a long period of adaptation, have learned to wrest a living of sorts from it' (Leijonhufvud 1973: 327).

The Econ, he noted, were xenophobic towards the neighboring PolScis and the Sociogs tribes, obsessed with the building of 'modls,' and sharply divided into castes, the most numerous of which were the Micro and the Macro. The castes distinguished themselves from each other using Totems that were, to the outsider, remarkably similar. The 'Totem of the Micro' was a pair of lines labeled 'S' and 'D,' while (when Leijonhufvud wrote the paper in 1973) the totem of the Macro was a pair of intersecting lines labeled 'IS' and 'LM':

The Totems are easily drawn, but deriving them logically from the underlying theory is another matter altogether. As we saw in the previous chapter, a demand curve derived in accordance with the underlying theory can have any shape at all – it will more often look like a snake in a hurry than the simple downward-sloping line drawn here.

The supply curve suffers an even worse fate: it doesn't exist.

Economists attempt to derive the supply curve from their theory of how profit-maximizing firms decide how much output to produce. One essential step in this derivation is that firms must produce so that the price they are paid for their output equals what is known as the 'marginal cost' of production – the additional expense incurred in producing one more unit of output. Unless this condition is met, a supply curve cannot be drawn.



4.1 Leijonhufvud's 'Totems' of the Econ tribe

This explains the extreme hostility that neoclassical economists have towards monopolies. It's not only because they can abuse the power that being a monopoly can confer: it's also because, according to neoclassical theory, a monopoly will set its price above the marginal cost of production. If monopolies were the rule, then there could be no supply curve, and standard neoclassical microeconomic analysis would be impossible.

Conversely, neoclassical economists love the market structure they call 'perfect competition,' because it guarantees that profit-maximizing behavior will cause firms to produce an output at which marginal cost equals price.

Only it won't. The manner in which neoclassical economics derives the result that profit-maximizing behavior by competitive firms means that they will produce where marginal cost equals price commits one of the simplest mathematical mistakes possible: it confuses a very small quantity – an 'infinitesimal,' as mathematicians describe it – with zero.

When that error is corrected, it is easily shown that a competitive market will also set price above marginal cost, and therefore a supply curve that is independent of the demand curve can't be drawn. The other half of the 'Totem of the Micro' disappears.

The kernel

Try this party trick: convince someone that the world is flat, starting from the premise that it is a sphere.

The argument is simple. If you take a small enough segment of the world – say, the two feet your victim is standing on – then the curvature of that segment is so small that it is, to all intents and purposes, flat. Then consider the segment you're standing on – it is also so small that it is effectively flat.

Next, consider the angle between the two segments: it too will be so small that it is effectively zero. So these two small segments are effectively flat.

Finally, extrapolate your argument from these two tiny segments and the angle between them up to the level of the entire globe. If you consider the segment your victim occupies and the segment behind him, that pair is also effectively flat. Keep on going, and the entire world is flat.

The fallacy in the argument, clearly, is that while it will do as an approximation to treat your immediate surroundings as effectively flat, it will not do to ignore those imperceptible but non-zero angles if you move from the scale of one or two segments to the entire globe.

Yet this fallacy lies at the heart of the economic preference for small, competitive firms over large monopolistic ones. At crucial stages of the economic argument, an imperceptibly small quantity is treated as zero, and then all these zeros are added up to yield zero at the scale of an entire market. This is intellectually and mathematically unsound. When the correct position is imposed – that something which is extremely small is nonetheless not zero – the economic argument against monopolies and in favor of small competitive firms collapses.

Oh, and if your party trick convinces your victim? Then he is either stoned, or an economist.

Prelude: the War over Perfect Competition

Most of this book explains flaws in neoclassical economic theory that have been known for decades, but have been ignored by neoclassical economists. When I first wrote *Debunking Economics*, I thought that the argument presented in this chapter was a new critique.

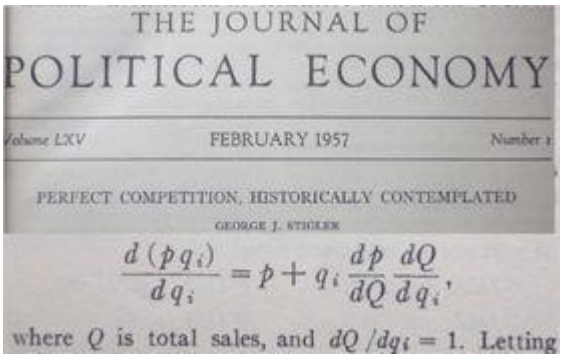
As I found out shortly after the book was published in 2001, it wasn't: the same key point had been made forty-four years earlier, and not by a critic of neoclassical economics but by one of the most strident defenders, George Stigler. In his paper 'Perfect competition, historically contemplated' (Stigler 1957: 8, n. 31), Stigler applied one of the most basic rules of mathematics, the 'Chain Rule,' to show that the slope of the demand curve facing the competitive firm was exactly the same as the slope of the market demand curve – see Figure 4.2.

If you haven't yet studied economics, then the importance of that result won't yet be obvious to you. But if you have, this should shock you: a central tenet of your introductory 'education' in economics is obviously false, and has been known to be so since at least 1957.

Stigler's mathematics deconstructed the demand curve for the individual firm into two components:

- the slope of the market demand curve; multiplied by
- how much market output changes given a change in the output of a single firm.

Neoclassical theory assumes that the slope of the market demand curve is negative: a fall in price will cause demand to increase. So the demand curve for the individual firm can only be zero if the second component is zero: the amount that industry output changes given a change in output by a single firm.


THE JOURNAL OF
POLITICAL ECONOMY

Volume LXV FEBRUARY 1957 Number 1

PERFECT COMPETITION, HISTORICALLY CONTEMPLATED

GEORGE J. STIGLER

$$\frac{d(pq_i)}{dq_i} = p + q_i \frac{dp}{dQ}\frac{dQ}{dq_i},$$

where Q is total sales, and $dQ/dq_i = 1$. Letting

4.2 Stigler's proof that the horizontal firm demand curve is a fallacy

However, Stigler very correctly stated that this second component is not zero, but instead equals one. In the basic 'Marshallian' theory of the firm that is taught to undergraduates, individual firms are assumed not to react strategically to what other firms do or might do. Therefore if one firm changes its output by ten units, there is no instantaneous reaction to this by the other firms, so that industry output also changes by ten units (though it might alter afterwards as other firms adjust to the new market price). The ratio of the change in industry output to the change in output by a single firm is therefore 1.

As a consequence, *the slope of the demand curve for the individual competitive firm equals the slope of the market demand curve*. Far from the individual firm's demand curve being horizontal, it has the same negative slope as the market demand curve.

I was stunned. This had been known for over four decades, and yet economic textbooks everywhere continued to mouth the fallacy that the individual competitive firm had a horizontal demand curve?

Even by the standards of mendacity that I had come to expect of economic textbooks, this surprised me. Most critiques of neoclassical theory involve complicated concepts – like the critique of the 'Law of Demand' outlined in the previous chapter, or the disputes over the nature of capital in Chapter 7. Frequently, when I have criticized textbooks for not discussing these issues, I have been hit with the rejoinder that this material is just too complicated for undergraduates to understand: better leave it for more advanced courses.[1] But this error in the theory is so simple that it can be explained in a few lines of English (and one line of calculus).

Neoclassical economists ignored most of this book, but vigorously attacked this chapter. As I responded to their attacks, the critique grew in depth and complexity. Attempts to get it into neoclassical journals failed, but it was published in a range of non-neoclassical outlets, including the journal of interdisciplinary physics *Physica A* (Keen and Standish 2006), *A Guide to What's Wrong with Economics* (Keen and Fullbrook 2004), the *Handbook of Pluralist Economics Education* (Keen 2009a), and the *Real-World Economics Review* (Keen and Standish 2010).[2]

Though nothing in that war with neoclassical economists challenged the accuracy of the case I first made in 2000, I have made extensive changes to

this chapter to focus on the key challenge it makes to neoclassical orthodoxy: that a 'supply curve' cannot be drawn. I have also added new material, including the key advance over the case made in 2000: a proof that the alleged profit-maximizing formula ('set marginal cost and marginal revenue equal to maximize profits') does not maximize profits. I derive another formula that does maximize profits, given the assumptions of neoclassical theory.

The roadmap

In this chapter I outline the neoclassical analysis of monopolies on the one hand, and 'perfect competition' on the other, and point out that the sole difference between them is that a monopolist is shown to face falling marginal revenue, whereas the competitive firm faces constant marginal revenue which is equal to the market price. From this proposition alone flows the crucial result, for the neoclassical approach to economics, that a supply curve can be derived that is independent of the demand curve.

I then show that this proposition leads to logical fallacies: a quantity that economists assume is zero actually has to be minus one; firms that are allegedly profit maximizers must produce more than the amount which maximizes profits; zero amounts at the individual level must somehow aggregate to negative amounts at the aggregate.

A careful analysis of what is implied by this proposition that marginal revenue equals price for competitive firms shows that it is based on a simple mathematical error. Once this is corrected, it is obvious that a competitive market with profit-maximizing firms that faces the same cost conditions as a monopoly will produce the same amount at the same price.

It follows that the amount supplied by a competitive industry is not determined by the aggregate marginal cost curve alone, but instead depends on conditions of demand as well, as with a monopoly. A supply curve that is independent of the demand curve therefore cannot be derived.

Economic perfection

Pejorative expressions abound in economics, despite its claim to be a value-free science, and 'perfect competition' is possibly the most value-laden of all. To economists, however, the word 'perfect' has a very precise

meaning: it is a market in which the competitively set price equals the marginal cost of production.

This is 'perfect' because, according to economic theory, it achieves the maximum possible gap between community welfare and the cost of providing it. Community welfare is maximized when the gap between total benefit to society from consuming a given product and the total cost of providing that benefit is as big as it can be. Given the shape that economists assume that these benefits and costs take – the benefit of consumption rising but at a decreasing rate, the cost of production rising at an increasing rate – the gap between the two is highest when the rate of change of total benefit equals the rate of change of total cost.

The demand curve (which we deconstructed in the last chapter) represents the rate of change of the total benefit, while the supply curve represents the rate of change of total cost. Therefore the benefit to society is maximized where these two rates of change – one rising, the other falling – are equal.

Producers are trying to maximize the benefit to them – their profits – not society's benefits. These two interests – consumers aiming to get the maximum benefit out of consumption, producers trying to get the maximum profit out of production – only coincide if the price equals the change in revenue that producers get from selling an extra unit, which economists call 'marginal revenue.' This is because the price – the amount that consumers are willing to pay – tells you the 'marginal utility' they get from the last item consumed. Only if this also equals the 'marginal revenue' that the producer gets from selling this very last unit of output will the benefits to society also equal the individual gain for the producer who sells it. This can only occur if the 'marginal revenue' for producing this last item sold equals its price.

Only perfect competition guarantees this outcome, because, economists believe, only then does marginal revenue always equal price.

Perfect competition is also 'perfect' because a supply curve exists if, and only if, price equals marginal cost. Without perfect competition, though a marginal cost curve can still be drawn, this will not be the supply curve, and as we shall see, the amount supplied to the market will be less than the amount that will maximize social welfare.

This concept of economic perfection relies upon downward-sloping market demand curves, which we already know is invalid. However, even if we accept, for the sake of argument, that the market demand curve is smoothly downward sloping and represents community welfare, the neoclassical argument for the superiority of the perfectly competitive market over the monopoly firm is still internally flawed. To establish this, we'll first consider the market form least favored by economics: monopoly.[3]

Monopoly

A monopoly has the entire market demand curve to itself. If the market demand curve is smoothly downward sloping, the price at which its output can be sold decreases as the quantity it tries to sell increases. In this chapter I'll work with a hypothetical example in which the market price is assumed to start at $1,000 for the first unit sold, and then to drop by five cents for every additional unit (see Table 4.1).

TABLE 4.1 Demand schedule for a hypothetical monopoly

Quantity	Price	Total revenue	Marginal revenue
1	1,000	1,000	1,000.00
2	999.95	1,999.90	999.90
3	999.90	2,999.70	999.80
10	999.55	9,995.50	999.10
11	999.50	10,994.50	999.00
2,001	900.00	1,800,900	800.00
2,002	899.95	1,801,700	799.90
4,001	800.00	3,200,800	600.00
6,001	700.00	4,200,700	400.00
8,001	600.00	4,800,600	200.00
10,001	500.00	5,000,500	0.00
12,001	400.00	4,800,400	–200.00
14,001	300.00	4,200,300	–400.00
16,001	200.00	3,200,200	–600.00
18,001	100.00	1,800,100	–800.00
20,000	0.05	1,000	–999.90

This may seem silly if you've never read an economics textbook before – why not simply use some real data on a real firm instead? – and you are right! The reason, as I explain in Chapter 5, is that *there are no such data*: the revenue and costs of real firms are nothing like those assumed by neoclassical economists. As a result, they always use made-up number in

their examples. To critique their theory, I have to do the same. So here we go …

Since the firm can sell one unit of output for $1,000.00, its total revenue is $1,000.00, and its 'marginal revenue' – the change in total revenue from zero dollars for zero units sold, to $1,000.00 for one unit sold – is also $1,000.00. So price equals marginal revenue at this level; but as soon as another unit is sold, price and marginal revenue diverge. Two units can only be sold if the firm drops the price for all units by 5 cents, so the market price becomes $999.95, the total revenue is $1,999.90, and the marginal revenue for the firm – the change in revenue from selling one unit to selling two – is $999.90.[4]

The interests of the firm therefore diverge from those of society, since the marginal benefit to it (the marginal revenue) is less than the marginal benefit to society as a whole (the price).

The price consumers are willing to pay drops smoothly as the quantity supplied rises, so that for an output of 10 units, the sale price has to drop to $999.55 per unit. The total revenue for selling 10 units is $9,995.50. If 11 units were to be sold, the monopolist would have to drop the price per unit by 5 cents, to $999.50 each. Total revenue would be $10,994.50 (eleven times $999.50), and marginal revenue would be $999.00.

The same process continues indefinitely, so that if output were 2,001 units, then sale price would have to drop to $900. Total revenue would be $1,800,900, and marginal revenue – the amount of additional revenue added by selling the 2,002nd unit – would be $800.00.

Eventually, the point is reached at which any further increase in output requires a price cut which reduces, rather than increases, total revenue. In this example, this occurs at an output of 10,001 units, where the sale price is $500. The sale of the 10,001st unit adds nothing to total revenue, and any increase in sales past this point actually reduces total revenue – marginal revenue has become negative.

That covers the revenue side of the analysis. The picture is completed by the analysis of costs, which I'll cover extensively in Chapter 5. Briefly, the firm has two types of costs: fixed costs, which apply no matter what the level of output is, and variable costs, which depend directly on how many units are produced.

Fixed costs are just that – fixed – so that the fixed cost *per unit of output* will fall as output rises. One fixed cost is the design of a product, and if this was, say, $10 million, then that component of the fixed costs per unit would be $1 million per unit when output was 10 units, and $1 per unit when output was 10 million units.

Variable costs depend on how many units are produced. One obvious variable cost is labor, and clearly you will need more labor to produce 10 million units than to produce 10. Neoclassical economics also assumes that, eventually, the productivity of the variable inputs such as labor will fall as output rises (we explore this assumption in Chapter 5). Therefore the variable costs to produce the 10 millionth unit will be much higher than those for the 10th unit. In my example, fixed costs are $10,000, and variable costs are defined by an equation in which they start at just over $15 each, fall for a while but then ultimately rise (see Table 4.2).[5]

Variable costs fall for a while because the firm experiences 'rising marginal productivity' as the ratio of the variable factors of production to fixed factors approaches the ideal level. This means that, for a while, the additional cost involved in producing the next unit of output falls. In my example, while it cost an additional $15 to go from producing zero units of output to producing one unit, it cost only an additional $8.80 to go from producing 2,001 units to 2,002 units.

This change in the cost of production resulting from producing one more unit is a very important concept in neoclassical economics, called the 'marginal cost of production.' As you can see from this example, marginal cost depends only on the change in variable costs – since fixed costs are the same no matter what level of output you produce – and it changes only because of changes in productivity that in turn reflect how many variable inputs are being used (workers) relative to the fixed inputs (machines).

TABLE 4.2 Costs for a hypothetical monopoly

Quantity	Fixed cost	Total cost	Marginal cost
1	1,000,000	1,000,015	15.00
2	1,000,000	1,000,030	15.00
3	1,000,000	1,000,045	14.99
10	1,000,000	1,000,150	14.96
11	1,000,000	1,000,165	14.95
2,001	1,000,000	1,022,419	8.80
2,002	1,000,000	1,022,428	8.80
4,001	1,000,000	1,042,433	13.41
6,001	1,000,000	1,086,464	33.62
8,001	1,000,000	1,190,514	74.23
10,001	1,000,000	1,400,190	140.05
12,001	1,000,000	1,770,696	235.86
14,001	1,000,000	2,366,836	366.48
16,001	1,000,000	3,263,017	536.70
18,001	1,000,000	4,543,241	751.32
20,000	1,000,000	6,300,100	1,015.01

Common sense, and earlier theories of economics like Ricardo's theory of rent, might consider that maybe the productivity of the individual inputs changes. Ricardo, for example, assumed that the cost of producing food rose as population rose because farmers started off using the most productive land, and had to use less fertile land as population increased. Common sense might suggest that as a firm demands more workers, it affects the wage at which workers can be hired, thus driving its costs per worker higher.

But neoclassical economists rule both these effects out, by assuming first that all inputs are homogeneous, and secondly that, while the monopoly has its own market to itself, it is a small player in the labor market and can hire as many workers as it likes at the going wage. The only source of changes in marginal cost that they allow arises from changing the ratio of variable inputs to the fixed inputs.

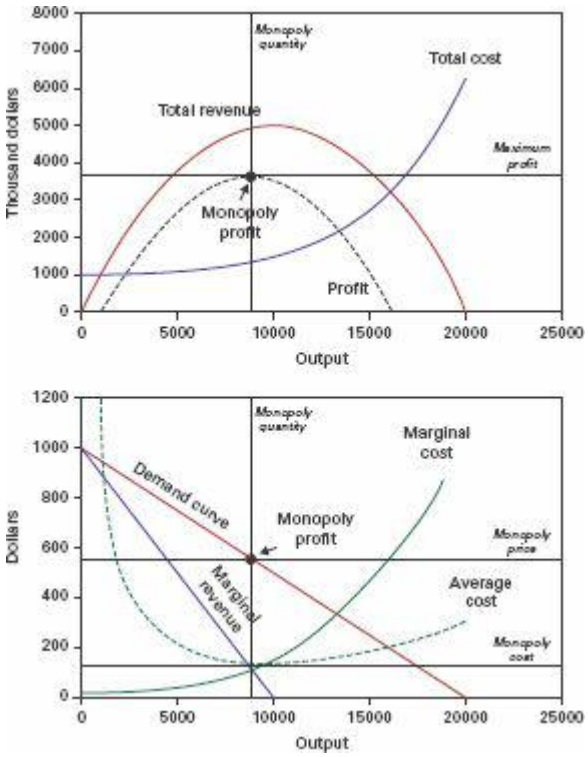
Consider a road construction firm, whose fixed costs include a number of jackhammers – say 100 of them. At a very low level of production, it will have only one worker and 100 jackhammers, so the worker will be very inefficient (*please read the footnote here*).[6] However, as the number of workers rises the firm will approach the ideal ratio of one worker per jackhammer, at which point maximum efficiency will be reached. But once the firm hits the ideal ratio, additional workers will add to output at a diminishing rate. Marginal productivity will fall, and therefore marginal costs will rise.

Table 4.3 combines the revenue information from Table 4.1 with the cost information from Table 4.2, and indicates the role of marginal revenue and marginal cost in identifying the point of maximum profit. For a while, each additional unit sold adds much more to revenue than it causes the total cost of production to rise: marginal revenue exceeds marginal cost, and therefore the final column in the table, which shows marginal revenue minus cost, is positive. But once marginal revenue and marginal cost are equal, profit is maximized.

The precise point at which this occurs lies between 8,973 and 8,974 units in this table, but the firm can't sell a fraction of a unit, so it will produce the lower amount of 8,973 units, at which the marginal cost is $102.77 and its profit will be $3,671,679.

The second column tells us that the market is willing to pay a price of $551.40 per unit if total supply is 8,973 units – so the sale price is $448.63 higher than the marginal cost of production (and $409.19 above the average cost).[7] Thus to maximize its profits, the firm produces where marginal cost equals marginal revenue, and sells the output at a much higher price.

As well as substantially exceeding the average cost of production, the market price exceeds the marginal cost of producing the last unit sold. This means, in economic welfare terms, that the marginal benefit of the last unit sold exceeds the marginal cost of producing it. Society would therefore benefit from an increased level of production, since additional units of output would increase social welfare. But the monopolist has no incentive to produce more: in fact producing any more would reduce his profits. Therefore, according to economists, monopolies reduce social welfare.



4.3 Profit maximization for a monopolist: marginal cost equals marginal revenue, while price exceeds marginal cost

Crucially for the way neoclassical economists prefer to model the economy, a supply curve can't be derived for a monopoly. Instead, if monopolies were the rule, then three curves – price, marginal revenue, and marginal cost – would be needed for a complete 'Totem of the Micro.' The intersection of the marginal revenue curve with the marginal cost curve would determine the amount the firm produced, and the market price would then depend on this quantity. In place of the simple mantra that 'prices are set by supply and demand,' the minimum statement of the 'Creed of the Micro' would be 'price is set by the demand curve, given the quantity set by marginal cost and marginal revenue.'

TABLE 4.3 Sales and costs determine the level of output that maximizes profit

Quantity	Price	Total revenue	Total cost	Profit	Average cost	Marginal revenue	Marginal cost	Marginal revenue minus marginal cost
1	1,000	1,000	1,000,015	−999,015	1,000,015	1,000	15	985
2	999.95	2,000	1,000,030	−998,030	500,015	999.90	15.00	984.90
3	999.90	3,000	1,000,045	−997,045	333,348	999.80	14.99	984.81
10	999.55	9,996	1,000,150	−990,154	100,015	999.10	14.96	984.14
11	999.50	10,995	1,000,165	−989,170	90,924	999.00	14.95	984.05
8,973	551.40	4,947,712	1,276,033	3,671,679	142.21	102.80	102.77	0.03
8,974	551.35	4,947,815	1,276,136	3,671,679	142.20	102.70	102.80	−0.10
10,001	500.00	5,000,500	1,400,190	3,600,310	140.01	0.00	140.05	−140.05
13,456	327.25	4,403,476	2,177,950	2,225,526	161.86	−345.50	327.16	−672.66
13,457	327.20	4,403,130	2,178,277	2,224,853	161.87	−345.60	327.23	−672.83
14,001	300.00	4,200,300	2,366,836	1,833,464	169.05	−400.00	366.48	−766.48
16,001	200.00	3,200,200	3,263,017	−62,817	204	−600.00	536.70	−1,136.70
18,001	100.00	1,800,100	4,543,241	−2,743,141	252	−800.00	751.32	−1,551.32
20,000	0.05	1,000	6,300,100	−6,299,100	315	−999.90	1,015.01	−2,014.90

It's no wonder, then, that, despite all the criticisms leveled at it, neoclassical economists cling to the model of the 'perfect' competitive market. In a competitive market, since marginal revenue equals price, profit-maximizing behavior leads to an output level at which price equals marginal cost. This is the embodiment of Smith's 'invisible hand' metaphor

about the capacity of market economy to reconcile private interest and public virtue, and that is the real message of the 'Totem of the Micro.'[8]

Perfect competition

The main distinguishing feature of the perfectly competitive market is the number of firms in it. Whereas a monopoly has just one firm – which therefore has the entire market demand curve to itself – a perfectly competitive market has many little firms, each competing for a tiny slice of total demand.

In the standard 'Marshallian' model that economists teach in undergraduate courses, these firms are assumed to be profit maximizers who behave in an 'atomistic' way: they neither know of, nor react in any way to, what other firms do or may hypothetically do – they simply respond to the market price.[9] In addition, it is assumed that entry into and exit from a competitive industry is 'free,' or more accurately, not subject to any barriers. Therefore firms outside the industry can move in at any time to take advantage of any above-normal profits if they exist.[10]

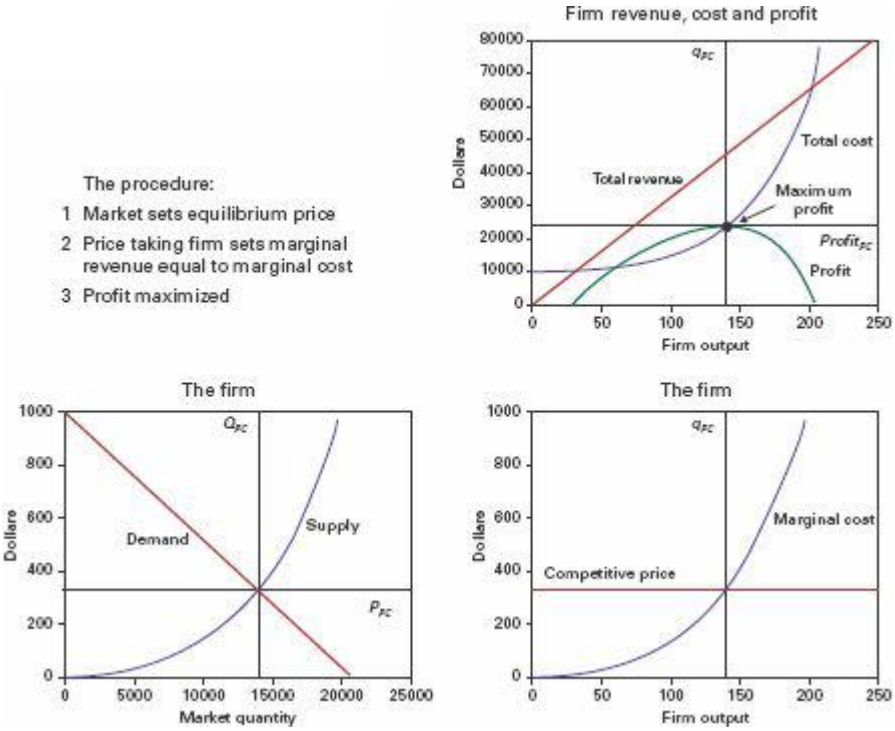
All firms are assumed to produce a product that is homogeneous from the consumers' point of view, so that there is no brand loyalty. All firms are therefore 'price-takers': they cannot influence the market price, but instead must take price as given.

At the market level, demand is still a negative function of price. Therefore, total market revenue will initially be a rising and then a falling function of price, and marginal revenue at the market level will be less than price (because to increase overall sales, the average price must fall).

However, economists argue that for each price-taking firm, marginal revenue and price are identical. The argument is that since they are each so small, no single firm can influence the market price. As a result, if any firm increases its price above the market equilibrium price, it will lose all its customers; while if any firm decreases its price below the market equilibrium, it will suddenly be swamped by all customers for that commodity. Therefore, the firm effectively sees a horizontal demand curve (set by the intersection of supply and demand at the level of the market).[11]

Given the assumption that they can sell as much as they like at the price set by the market, then as profit maximizers they will produce until the

marginal cost of producing this amount equals the marginal revenue from doing so. Since price is a constant for them, marginal revenue equals price, so they produce at the point where marginal cost equals price. In a 100-firm industry whose costs are identical to the monopoly I discussed previously, this results in the representative firm producing about 135 units.[12] This then results in a profit of $22,255.26 for the firm, or $2,225,526 dollars for the industry in total.



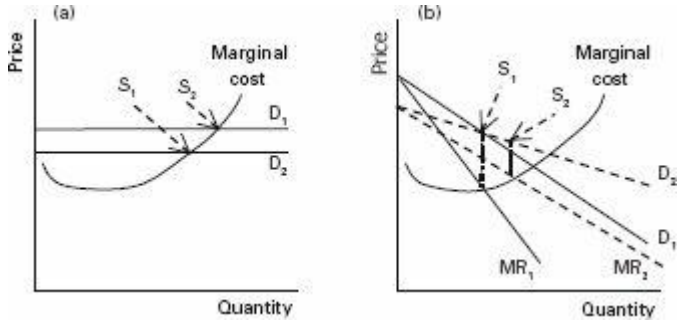
4.4 Profit maximization for a perfectly competitive firm: marginal cost equals marginal revenue, which also equals price

Since the total revenue for a perfectly competitive firm is simply a constant price times the number of units it sells, increasing its sales has no effect on its price, so its marginal revenue is constant. This in turn is why a supply curve can be derived for perfect competition, but not for monopoly.

The amount a monopolist will supply depends both on the firm's marginal cost function, and the market's demand function. Since both are needed to determine supply, and since many different demand curves can be drawn through the same point – each with a different slope and therefore different marginal revenue implications – it is impossible to derive a curve which shows how much a monopolist will supply at each price level (all

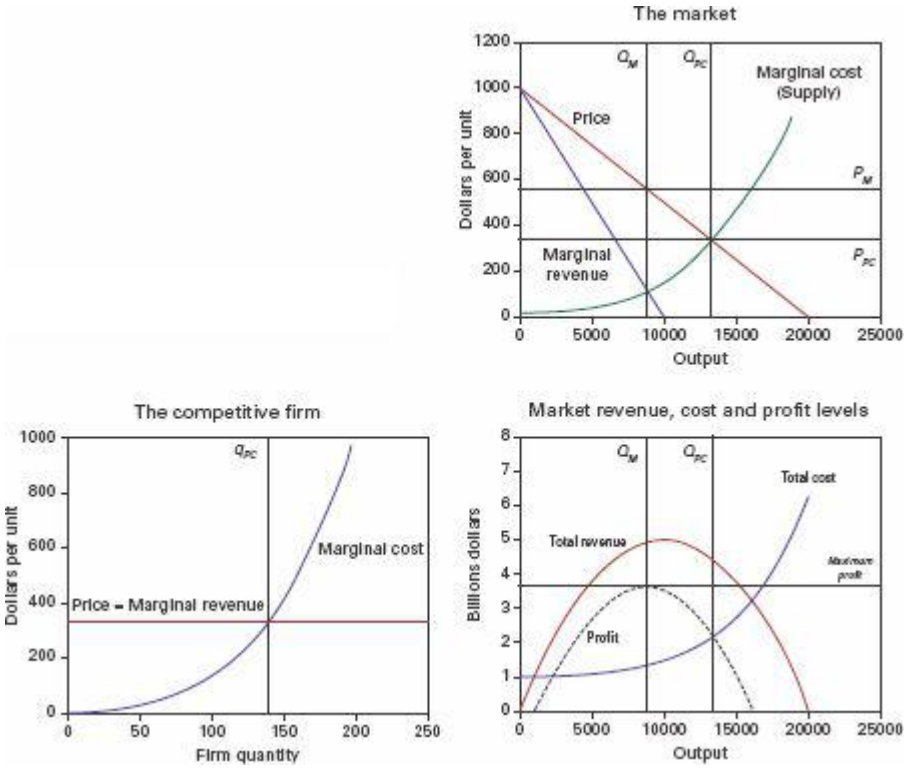
you can do is consider specific examples of hypothetical demand curves, as I did to generate Table 1).

However, for the perfectly competitive firm, since price equals marginal revenue, the amount the firm will produce corresponds in every case to its marginal cost curve. The supply curve of a single firm in a perfectly competitive market is thus its marginal cost curve.



4.5 A supply curve can be derived for a competitive firm, but not for a monopoly

The supply curve for a perfectly competitive industry is constructed simply by adding up the amounts that each firm is willing to supply at a given price. This amounts to summing up their marginal cost curves, so that the supply curve for the industry represents the marginal cost of producing output. Since demand equals supply in equilibrium, the marginal benefit for the last unit consumed equals its marginal cost of production, and social utility is maximized. This results in both a higher level of output and a lower price than would occur if the industry were a monopoly.



4.6 A competitive industry produces a higher output at a lower cost than a monopoly

Checking our sums

This argument normally convinces economics students, and it explains much of the hostility economists in general have towards monopolies, or any market in which firms have some market power by virtue of their size. This 'social radicalism' is unusual for a profession which is normally perceived as socially conservative. It is also curiously at odds with the real world, where it's fairly obvious that industries have a clear tendency to end up being dominated by a few large firms – why fight the real world?

Economists argue that their opposition to large firms, and their allegedly uncharacteristic radicalism on this issue, is based on sound analysis. But is it? Let's check, after first seeing how moving from a monopoly to a perfectly competitive industry would benefit society in my example.

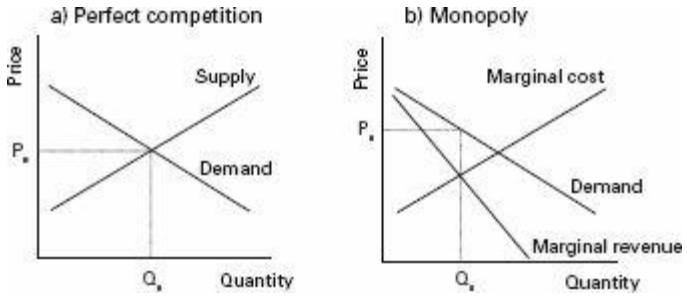
Table 4.4 adds up the costs and revenues of all the competitive firms, to show the aggregate outcome for a competitive industry with 100 firms. Note that the output of this industry (in the rows shown in italic) is higher

than the monopoly's output – roughly 13,456 units, versus 8,973 – and its price is lower – roughly $327.25 per unit, versus $551.40 for the monopoly.

Economists therefore put forward three reasons to prefer a competitive industry to a monopoly:

- the competitive industry produces where marginal cost equals price, thus maximizing social welfare;
- it produces a higher level of output than a monopoly; and
- it sells this higher output at a lower price.

However, the key reason why neoclassical economists themselves prefer perfect competition to monopoly is that perfect competition is the only market structure in which price and quantity are set by the intersection of the supply curve and the demand curve.



4.7 The standard 'supply and demand' explanation for price determination is valid only in perfect competition

Well, that's the theory. Now we will consider a subtle but profound set of problems which invalidate this entire analysis.

Calculus 101 for economists: infinitesmals ain't zero

Throughout the economic analysis of perfect competition, the assumption is made that the perfectly competitive firm is so small, relative to the overall market, that its impact on the market can be treated as zero. As I intimated earlier in this chapter, this kind of logic is OK when you are dealing with local approximations – such as whether you can regard the ground on which you stand as either flat or curved – but it will not do when those local approximations are aggregated together. When we insist that

infinitesimally small amounts are not in fact zero, the apparently watertight logic behind the comparison of monopoly and perfect competition falls apart.

TABLE 4.4 Cost and revenue for a 'perfectly competitive' industry identical in scale to hypothetical monopoly

Quantity	Price	Total revenue	Total cost	Profit	Average cost	Marginal cost	Price minus marginal cost
1	1,000	1,000	1,000,015	−999,015	1,000,015	15	985
2	999.95	2,000	1,000,030	−998,030	500,015	15.00	984.95
3	999.90	3,000	1,000,045	−997,045	333,348	14.99	984.91
10	999.55	9,996	1,000,150	−990,154	100,015	14.96	984.59
11	999.50	10,995	1,000,165	−989,170	90,924	14.95	984.55
8,973	551.40	4,947,712	1,276,033	3,671,679	142.21	102.77	448.63
8,974	551.35	4,947,815	1,276,136	3,671,679	142.20	102.80	448.55
10,001	500.00	5,000,500	1,400,190	3,600,310	140.01	140.05	359.95
13,456	*327.25*	*4,403,476*	*2,177,950*	*2,225,526*	*161.86*	*327.16*	*0.09*
13,457	*327.20*	*4,403,130*	*2,178,277*	*2,224,853*	*161.87*	*327.23*	*−0.03*
14,001	300.00	4,200,300	2,366,836	1,833,464	169.05	366.48	−66.48
16,001	200.00	3,200,200	3,263,017	−62,817	204	536.70	−336.70
18,001	100.00	1,800,100	4,543,241	−2,743,141	252	751.32	−651.32
20,000	0.05	1,000	6,300,100	−6,299,100	315	1,015.01	−1,014.96

Too small to matter? An essential part of the argument for perfect competition is that each firm is so small that it can't affect the market price – which it therefore takes as given. Consequently the demand curve, as perceived by each firm, is effectively horizontal at the market price. The firms are also so small that they do not react to any changes in behavior by other firms: in the language of economic theory, their 'conjectural variation' – how much all other firms change their output in response to a change in output by one firm – is zero.

These two assumptions are alleged to mean that the slope of the individual firm's demand curve is zero: both the firm's price and the market price do not change when a single firm changes its output. However, they also mean that, if a single firm increases its output by one unit, then total industry output should also increase by one unit – since other firms won't react to the change in output by a single firm.

However, there is a problem: these two assumptions are inconsistent.

If the market demand curve is downward sloping, then an increase in total market output must mean a fall in the market price – regardless of how small a fall it might be. Since the theory assumes that other firms don't react to an increase in production by one firm, total market output must increase. Since the market demand curve is downward sloping, and supply has increased – the supply curve has shifted outwards – market price must fall.

Therefore market price *does* change because of the actions of a single firm. The only way market price could not react would be if all other firms reduced their output by as much as the single firm increased it: then the market supply curve would not shift, and the price would remain constant. But the theory assumes that firms don't react to each other's behavior.

So the market price will be affected by the actions of a single firm, in which case the demand curve facing a single firm will be downward sloping – however slight the slope may be.

Putting this critique another way, the economic argument is that if you break a large downward-sloping line (the market demand curve) into lots of very small lines (the demand curves perceived by each firm), then you will have a huge number of perfectly flat lines. Then if you add all these perfectly flat lines together again, you will get one downward-sloping line.

This is mathematically impossible. If you add up a huge number of flat lines, you will get one very long flat line. If you break one downward-

sloping line into many small lines, you will have many downward-sloping lines. The economic concept of perfect competition is based on a mathematical error of confusing a very small quantity with zero.

The market matters: marginal revenue is marginal revenue A second problem with this economic model is the nature of the marginal revenue function. Economists unconsciously reason as if the marginal revenue curve at the market level is a function of the number of firms that produce the industry's output: it exists if there is only one firm, but if there are a large number of firms, it disappears. They then show that a monopoly sets its price where marginal revenue equals marginal cost, which is consistent with their theory. However, they show a competitive industry setting price where the supply curve and the demand curve intersect, with no pesky marginal revenue curve getting in the way.

Unfortunately, marginal revenue exists independently of the number of firms in the industry. If the market demand curve is downward sloping, then so is the market marginal revenue curve, and they diverge right from the very first unit sold (as you can see in the example).

So if a competitive industry did result in output being set by the intersection of the demand curve and the supply curve, then at the collective level the competitive industry must be producing where marginal cost *exceeds* marginal revenue. Rather than maximizing profits, as economists argue firms do, the additional output – that produced past the point where marginal revenue equals marginal cost at the industry level – must be produced at a loss. This paradox means that the individual firm and the market level aspects of the model of perfect competition are inconsistent.

Creative accounting For the assertion that perfect competition results in a higher level of output at a lower price than monopoly to be correct, then in the aggregate, the individually rational profit-maximizing behavior of perfectly competitive firms must lead to a collectively irrational outcome. This would be OK if the theory actually admitted this – as do the theories of Cournot and Bertrand competition[13] – but the Marshallian model taught to undergraduates claims instead that equating marginal cost and marginal revenue maximizes profits for the competitive firm.

According to the theory, the monopoly firm produces only to the point at which marginal cost equals marginal revenue, because this output maximizes its profit. Each perfectly competitive firm likewise produces to a point at which its marginal cost equals its marginal revenue, and for the same reason – because this level of output maximizes its profit.

But at the market level, competitive firms produce to a point at which the collective marginal cost exceeds marginal revenue. The perfectly competitive *industry* produces where marginal cost equals price but exceeds marginal revenue; yet all firms in it are supposed to be producing where marginal cost equals marginal revenue.

The monopoly sets price where marginal revenue equals marginal cost, while the competitive industry sets price where the supply curve (which is the sum of all the individual firms' marginal cost curves) intersects the demand curve: this is supposed to be the result of setting marginal cost equal to marginal revenue at the firm level, which means each firm makes the maximum profit that it can. Yet at the aggregate level, while the monopoly has produced where profit is maximized, the competitive firms have produced beyond this point, so that the industry's output past the point of monopoly output has been produced at a loss – which is why the profit level for the competitive firm is lower than that for the monopoly, even though all its firms are supposed to be profit maximizers.

Where did this loss come from? It certainly can't be seen in the standard graph economists draw for perfect competition, which shows the individual competitive firm making profits all the way out to the last item produced.

Instead, this 'loss out of nowhere' is hidden in the detail that economists lose by treating infinitesimally small quantities as zeros. If perfectly competitive firms were to produce where marginal cost equals price, then they would be producing part of their output *past* the point at which marginal revenue equals marginal cost. They would therefore make a loss on these additional units of output.

As I argued above, the demand curve for a single firm cannot be horizontal – it must slope downwards, because if it doesn't, then the market demand curve has to be horizontal. Therefore, marginal revenue will be less than price for the individual firm. However, by arguing that an infinitesimal segment of the market demand is effectively horizontal, economists have treated this loss as zero. Summing zero losses over all firms means zero

losses in the aggregate. But this is not consistent with their vision of the output and price levels of the perfectly competitive industry.

The higher level of output must mean losses are incurred by the industry, relative to the profit-maximizing level chosen by the monopoly. Losses at the market level must mean losses at the individual firm level – yet these are presumed to be zero by economic analysis, because it erroneously assumes that the perfectly competitive firm faces a horizontal demand curve.

Perfect competition equals monopoly The above critique raises an interesting question: what will the price and output of a perfectly competitive industry be, if we drop the invalid assumption that the output of a single firm has no effect on the market price? The answer is: the price and output levels of a competitive industry will be exactly the same as for the monopolist (if the aggregate marginal cost curve of the competitive firms is identical to the marginal cost of the monopoly, which economic theory assumes it is).

Economic explanations of price-setting in a competitive market normally start from the level of the market, where they show that the intersection of supply and demand sets both price and quantity. They then argue that the price set by this intersection of supply and demand is taken as given by each competitive firm, so that the supply curve for the individual firm is its marginal cost curve. Then they notionally add all these marginal cost curves up, to get the supply curve for the industry – and its point of intersection with the demand curve determines the market price.

But there is a 'chicken and egg' problem here. Which comes first – price being set by the intersection of supply and demand, or individual firms equating marginal cost to price? And why should a level of output which involves making a loss on part of output (the part past where market marginal revenue equals marginal cost) determine where each individual firm perceives price as being set?

Economists have been bewitched by their own totem. They draw a downward-sloping market demand curve, and an upward-sloping supply curve, and assume that price and quantity must be set by the intersection of the two curves. But the 'supply curve' is really only the aggregate of the marginal cost curves of all the competitive firms. It isn't a supply curve

unless they can prove that, whatever the market demand curve looks like, the industry will supply the quantity given by the intersection of the demand curve and the aggregate marginal cost curve.

This isn't the case in their own model of monopoly. The intersection of marginal cost and marginal revenue determines the quantity produced, while the price charged is set by the price the demand curve gives for that quantity: price and quantity are *not* determined by the intersection of the demand curve and the marginal cost curve.

Economists claim that price and quantity are set by the intersection of the demand curve and the aggregate marginal cost curve in the case of perfect competition, but their 'proof' relies on the erroneous proposition that the demand curve perceived by each individual firm is, you guessed it, horizontal.

Once this spurious proposition is removed, the price that the competitive firm takes as given is the price determined by the intersection of the market demand curve and the aggregate marginal cost curve – which is precisely the same price as a monopoly would charge. To argue otherwise is to argue for either irrational behavior at the level of the individual firm – so that part of output is produced at a loss – or that, somehow, individually rational behavior (maximizing profit) leads to collectively irrational behavior – so that profit-maximizing behavior by each individual firm leads to the industry somehow producing part of its output at a loss. However, the essence of the neoclassical vision is that individually rational behavior leads to collectively rational behavior.

Therefore, the price that the perfectly competitive firm will take as given when it adjusts its output is not a market price set by equating price to marginal cost, but a market price set by equating marginal revenue to marginal cost. The quantity produced at this price will be equivalent, when summed, to the output of a single monopolist. On the grounds of properly amended economic theory, monopoly and perfect competition are identical.

Returns to scale and the durability of perfect competition

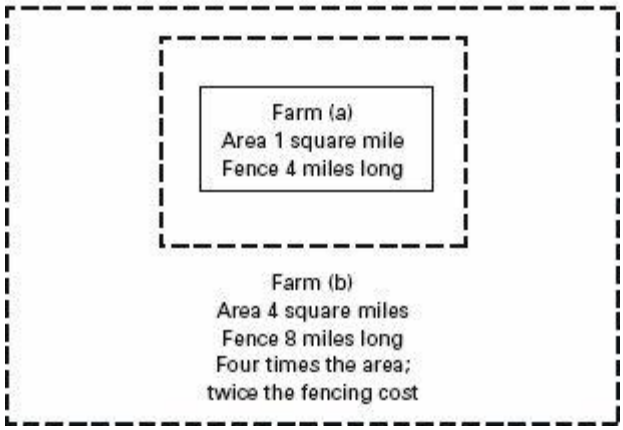
To date I have accepted the assumption that a monopoly has no scale advantages over a perfectly competitive firm, so that it is possible to sum

the cost functions of numerous small firms and come up with aggregate costs similar to those of a large firm.

In general, this assumption of scale-invariant costs will be invalid. If we are simply considering the costs of producing a homogeneous product, then it is likely that a very large firm will have scale advantages over a very small one. In the vernacular of economics, large firms benefit from returns to scale.

Returns to scale occur when the cost of production rises less rapidly than the output as the scale of production increases. A simple example is in farming, where farms need to be separated from each other by fences. The amount of fencing required depends on the perimeter of the farm. If we consider a square block of land, fencing will depend on the length of the four sides of the square. Cost is thus the cost of fencing per mile, times four times the length of a side. But the area enclosed by the fence depends on the length of a side squared. The output of a farm is related to its area, so that output is a function of the length of a side squared. Doubling the perimeter of a farm thus doubles its fencing costs, but increases its output fourfold.

As a result, large farms have a scale advantage over smaller farms. A farm with a square mile of land requires four miles of perimeter fencing to each square mile, while a farm with four square miles of land requires eight miles of perimeter fencing – or just two miles of perimeter fencing to each square mile of land.



4.8 Double the size, double the costs, but four times the output

The same concept applies in numerous ways. For a substantial range of output, a large blast furnace will be more cost effective than a smaller one, a

large ship than a smaller one, a large car factory than a smaller one.

If large firms have cost advantages over small ones, then given open competition, the large firms will drive the small ones out of business (though marketing and debt problems will limit the process, as Sraffa notes). Hence increasing returns to scale mean that the perfectly competitive market is unstable: it will, in time, break down to a situation of either oligopoly (several large firms) or monopoly (one large firm).

Economists have been well aware of this dilemma since Marshall at the end of the nineteenth century, and the fiction that has been invented to cope with it is the concept of the long run average cost curve.[14] The curve is 'u-shaped,' which asserts that there is some ideal scale of output at which the cost of production is minimized. In the long run, all inputs can be varied, so this shape is supposed to represent increasing returns to scale up to the point of minimum cost, beyond which decreasing returns to scale start to occur, so that the cost of production rises.

A competitive industry is supposed to converge to this ideal scale of output over time, in which case its many extremely big firms are safe from the predations of any much larger firm, since such a competitor would necessarily have higher costs.

This defence is specious on several counts.

First, the question of whether perfect competition can exist in a particular industry becomes an empirical one: what is the ideal scale of output, and how many firms could then occupy a particular industry at a particular time?

For some industries, the answer might well be 'very many' – the ever popular wheat farm comes to mind. However, for some other industries, the answer might well be 'very few.' It seems, for example, that the worldwide market for large intercontinental passenger airplanes can support at most three firms.

The argument that, 'in the long run,' this industry could be perfectly competitive because it could grow big enough to support hundreds or thousands of competitors is ludicrous. By the time the world was large enough to support hundreds of Boeings and Airbuses, it is highly likely that some entirely different form of transport would have superseded the airplane.

Secondly, the long run supply curve is actually constructed under the assumption of constant technology: in other words, it is not really a concept in time at all. The scale economies are supposedly there all the time, ready to be exploited.

If so, then unless an industry is already big enough to support the enormous number of firms surmised by the model of perfect competition – all operating at the ideal scale – large firms can immediately out-compete small firms. In other words, the only way competitive firms can survive is if the industry is already so large that it can support an enormous number of firms of the ideal scale.

The theoretical response of economists to this dilemma has been to presume constant returns to scale. With constant returns, 'size does not matter': a small firm will be just as cost efficient as a large one.

Unfortunately, size does matter. Economies of scale are an important part of the reason that most industries are dominated by a small number of very large firms. We do need an adequate analysis of how such an industry functions, but neoclassical economics does not provide it.

Addendum: the war over perfect competition

As noted, my plan to start work on *Finance and Economic Breakdown* (a book-length treatment of Minsky's 'Financial Instability Hypothesis') when I finished *Debunking Economics* was derailed for the next four years as I found myself embroiled in disputes with neoclassical economists – via email, in web forums, in public and in referee comments on my papers – about this argument. The end result was a substantial strengthening of the critique, the most important component of which was a proof that *equating marginal cost and marginal revenue does not maximize profits*.

I developed this proof after realizing that the key result in this chapter – that the demand curve for a competitive firm cannot be horizontal – was discovered by the neoclassical economist George Stigler over half a century ago. Why, I wondered, did he nonetheless continue to subscribe to and defend neoclassical theory?

Apart from the usual psychological explanation – that when you've committed yourself to a particular belief system and made your reputation in it, it is extraordinarily hard to accept that it might be false – there is a

technical reason in the same paper. Though he proved that the individual firm's demand curve had the same negative slope as the market demand curve, Stigler also proved that, *if firms produced where marginal cost equaled marginal revenue*, then the more firms there were in an industry, the closer industry output would be to where price equaled marginal cost.

> It is intuitively plausible that with infinite numbers all monopoly power (and indeterminacy) will vanish […] But a simple demonstration, in case of sellers of equal size, would amount only to showing that Marginal revenue = Price + Price/Number of sellers [times] Market elasticity, and that this last term goes to zero as the number of sellers increases indefinitely. (Stigler 1957: 8)

Stigler thus believed that he had neutralized his finding in the same paper. Yes, the conventional neoclassical belief that the individual competitive firm faces a horizontal demand curve is false, but if there are a large number of firms in an industry, then marginal revenue for the individual firm will be very close to the market price. Therefore the collective effect is the same: price will be set where supply equals demand. The key result of competition is restored.

From this point on, the standard failings of neoclassical research and pedagogy took over. Only a minority of economists read the paper; textbooks continued to teach the concept that Stigler had disproved; and the minority of economists who were aware of Stigler's paper defended the failure to take his result seriously because, in the end, the outcome was alleged to be the same: supply will equal demand in a competitive market.

I instead saw a logical error: Stigler's proof that marginal revenue for the individual firm would converge to market price as the number of firms increased was correct, *if* those firms all set marginal revenue equal to marginal price. But all the problems that I had identified in this chapter still remained: in particular, producing where supply equaled demand required 'profit-maximizing' firms to actually make losses on all goods sold past the point at which industry-level marginal revenue equaled marginal cost.

There was only one explanation: *equating marginal cost and marginal revenue couldn't be profit-maximizing behavior*.

I followed the logic forward and proved that the true profit-maximizing formula was quite different. If competitive firms did actually profit-maximize, they would produce an output much lower than the level where marginal cost equaled marginal revenue. The market outcome was that a competitive industry would produce the same amount as a monopoly, and market price would exceed marginal cost.

Equating marginal cost and marginal revenue does not maximize profits

The logic is fairly simple to follow if you imagine that you are running a competitive firm, and ask yourself this question: 'Is my level of output the *only* factor that can affect my profits?' The answer is, of course not: your profit depends not just on how much you produce, but also how much all the other firms in the industry produce. This is true even if you can't control what other firms do, and even if you don't try to react to what you think they might do. You work in a multi-firm industry, and the actions of all other firms impinge upon your own profits.

However, the neoclassical 'profit-maximizing' formula implies that your output is the only factor determining your profits: it uses simple calculus to advise you to produce where the change in your profits *relative to your own output* is zero. What you really need to do – if you're going to try to use calculus to work out what to do – is to work out where the change in your profits *relative to total industry output* is zero.

Intuitively, this is likely to mean that the actual amount you produce – which is something you can control – should be *less* than the amount the neoclassical formula recommends. This is because it's highly likely that the impact on your profit of changes in output by other firms – which you can't control – will be negative: if other firms increase their output, your profit is likely to fall. So when you work out the impact that changes in output by other firms has on your profits, the sign of this change is likely to be negative.

Therefore, to find the point at which your profit is at a maximum with respect to total industry output, you're likely to want the sign for the impact of *your* changes in output on *your* profit to be positive. This will mean that

your output level will be less than the level at which your marginal cost equals your marginal revenue.

The best way to solve this problem precisely is to work out when what is known as the 'total differential' of the firm's profit is equal to zero (to avoid using symbolic terms like 'n' for the number of firms in the industry, I'll work with a hypothetical industry with 1,000 firms in it, but the logic applies independently of the number of firms in the industry).

The profit of your firm will be its revenue – which will equal your firm's output times the market price – minus costs. What we have to do is work out how these two aspects of profit are influenced by the changes in output by *all* the firms in the industry, including your own.

Using a calculus procedure known as the Product Rule, the change in the revenue side of this calculation can be broken down into two bits:

- *your* output, times how much a given firm's change in output changes market price; plus
- market price, times how much a given firm's change in output causes *you* to alter your own output.

Thanks to Stigler's accurate calculus from 1957, we know that we can substitute the slope of the market demand curve for 'how much a given firm's change in output changes market price,' so the first term in the change in revenue calculation becomes your firm's output, times the slope of the market demand curve. With 1,000 firms in the industry, we get 1,000 copies of this term, which is your firm's output, multiplied by the slope of the market demand curve.

The second term in the change in revenue is the market price, times the amount that *your* output changes owing to a change in output by a given firm. Since we're working with the Marshallian model, which assumes that firms don't react strategically to what other firms do, then 999 times out of 1,000 this term will be the market price times zero. But once, it will be how much *your output* changes, given a change in *your output*. The ratio of the change in your output to the change in your output is one, so once – and only once – this calculation will return the market price.

Finally, we have to consider the cost side of the calculation: this will be how much your total costs change, given a change in output by a given firm. As with the last calculation for revenue, 999 times out of 1,000 this will be zero – because your costs don't change when the output of another firm changes. But once, and only once, it will be how much your total costs change, given a change in your output. This is your firm's marginal cost.

That gives you three terms, and when the output level you choose causes the sum of these three to be zero, you have identified the output level for your firm that will maximize your profits. These three terms are:

- the market price (a positive number);
- plus the slope of the market demand curve multiplied by *1,000 times* your output (a negative number, since the slope of the market demand curve is negative);
- minus your marginal cost.

The difference between this formula and the neoclassical formula is subtle and the size of an elephant at the same time. The neoclassical formula tells you that you maximize your profits when these three terms are equal:

- the market price;
- plus the slope of the market demand curve multiplied by your output;
- minus your marginal cost.

Whoops! The neoclassical formula has erroneously omitted 999 times your output times the slope of the market demand curve – a very large negative term (since the slope of the market demand curve is negative). It therefore takes a larger marginal cost to reduce the whole neoclassical expression to zero, which you can only achieve by producing a higher output. All of this additional output will be sold at a loss: the increase in revenue you get from selling those additional units will be less than the increase in costs the additional production causes.

The neoclassical formula is thus categorically wrong about the level of output by the individual firm that will maximize its profits – except in the one case of a monopoly, where the two formulas coincide.

If competitive firms are truly profit maximizers, then they will produce substantially less output each than neoclassical theory says they will (roughly half as much), and the sum of this output will – if they face identical costs of production – be the same as would be produced by a monopoly.

It could be argued that the accurate formula derived above requires the firm to know something that it can't possibly know – which is how many firms there are in the industry. In fact, this is less of a problem than it seems, because it's possible to reorganize this formula into a form in which the number of firms in the industry isn't that important.[15] But a more important point is that in reality firms don't 'do calculus.' They are far more likely to work out the answer to this and other questions by trial and error.

Calculus schmalculus

What firms would actually do is work out the ideal amount to produce to maximize profits by choosing some output level at random, and then vary this amount to see what happens to their profits. If a firm's profit rose, then it would continue altering its output in the same direction; but if its profit fell, then it would reverse direction.

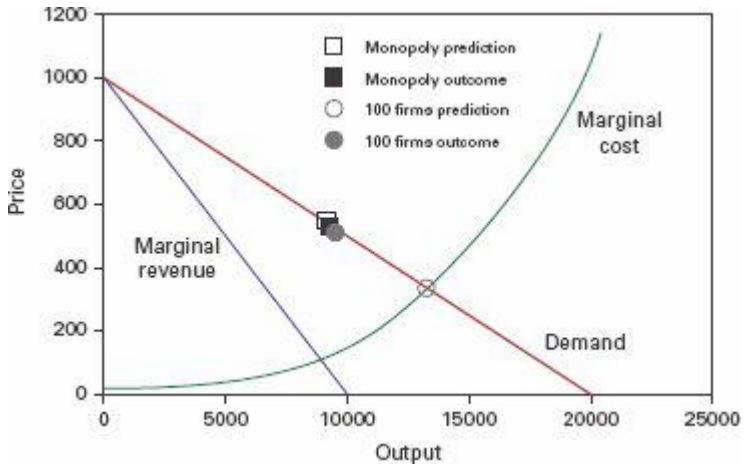
Unfortunately we can't test this using empirical data, because, as I argue later, the assumptions of the neoclassical model (a falling demand curve, a rising supply curve, and a static setting in which to maximize profits) don't even come close to describing reality. But one can today create an artificial market using a computer model that does fit the neoclassical assumptions, and then see what happens.

The next few graphs show the results of this simulation:

- firms choose an initial output level at random;
- the initial market price is determined by the sum of these randomly chosen outputs;
- each firm then chooses a random amount to vary its output by, and changes its initial output by this amount;

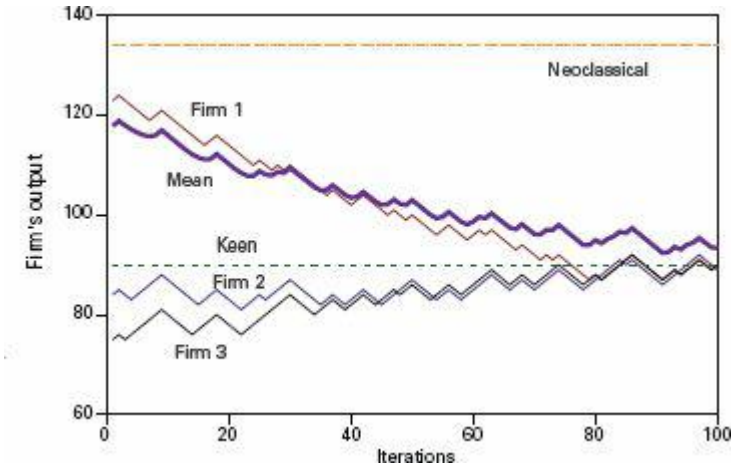
- a new market price is calculated;
- if a firm's profit has risen as a result of the change in output, it continues changing its output in the same direction;
- otherwise it reverses direction.

At the extremes considered here, of a monopoly and a 100-firm industry, neoclassical theory is correct for a monopoly, but very wrong for the 100-firm industry. It predicts that such an industry will produce effectively where marginal cost equals price – where the 'supply curve' intersects the demand curve – but in practice the 100-firm industry produces an output that is almost the same as the monopoly's.



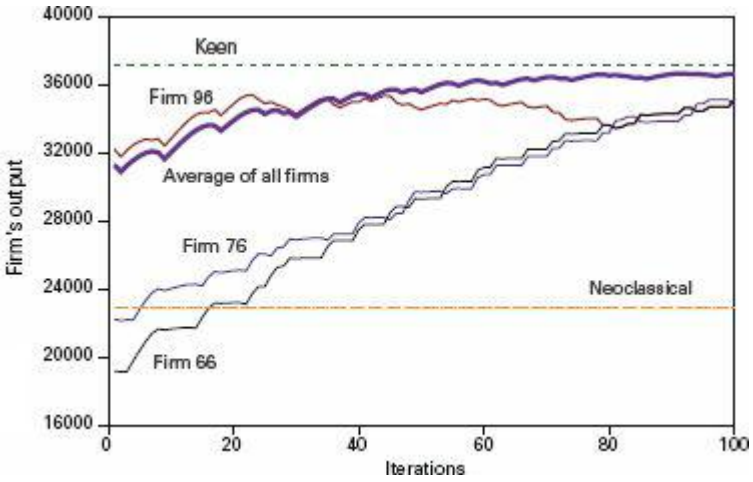
4.9 Predictions of the models and results at the market level

Neoclassical theory also predicts that industry output will converge to the competitive ideal as the number of firms in the industry rises. Simulations with between 1 and 100 firms in the industry show no pattern, though in general the output level is well below that predicted by neoclassical theory – but close to the prediction of my equation (see Figure 4.12).



4.10 Output behavior of three randomly selected firms

This market outcome is not caused by collusion, but is simply the result of profit-maximizing behavior. Firms also follow very different paths in their output, even though the basic 'strategy' is the same for each firm: vary output and try to find the output level that generates the largest profit.

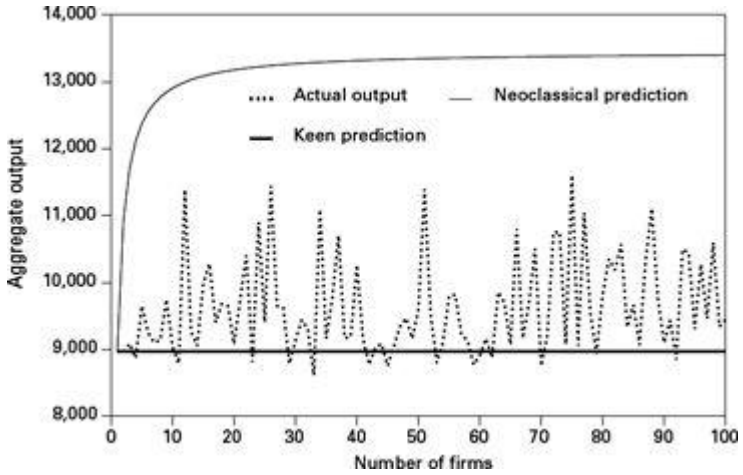


4.11 Profit outcomes for three randomly selected firms

Firms have very different outcomes with respect to profit as well, though in general most make far more profit from this 'suck it and see' algorithm than they would make if they followed the neoclassical formula.

Many different outcomes are possible with different assumptions – in particular, the introduction of some irrationality by firms (continuing to increase output when the last increase in output reduced profit, for example), or a greater dispersal in the size of the changes in output by firms

causes the aggregate result to move in the direction of the neoclassical formula (Keen and Standish 2010: 69–74). But the neoclassical proposition that strictly rational behavior leads to a competitive industry producing where individual marginal revenue equals marginal cost is strictly false.



4.12 Output levels for between 1- and 100-firm industries

Dialogue with the deaf

There are other theories of competition (the Cournot-Nash and Bertrand models of game-theoretic behavior, where firms do react strategically to what other firms might do), where a 'perfectly competitive' outcome can occur from non-profit-maximizing behavior. But the standard Marshallian model of the firm is categorically false: the demand curve for a competitive firm is not horizontal, equating marginal cost and marginal revenue does not maximize profits, and a competitive industry will produce the same amount of output as a monopoly and sell it for the same price. The 'Marshallian' model of the competitive firm is dead.

Or it should be. Instead, given the resistance of neoclassical economics to criticism, this false model is likely to live on for decades. Though this critique has been published in a range of journals – including one edited by physicists, whose mathematical capabilities far exceed those of economists – I have been unable to get it published in neoclassical economics journals. The odds that this critique will ever be recognized by economics textbooks writers are therefore effectively zero.

Every manner of excuse has been offered to avoid confronting these uncomfortable but mathematically unimpeachable results. The most remarkable excuses came from referees for the *Economic Journal* and the *Journal of Economics Education*.[16]

A referee for the former journal admitted that this result was significant, but argued that it did not matter because, he alleged, the conventional theory assumed that firms attempted to maximize their profits *while assuming that the output of other firms was fixed*. This alleged assumption cannot be found in any textbook on perfect competition, and amounts to an assumption of irrational behavior on behalf of firms: 'Needless to say, this result is worthy of publication on *Economic Journal* if it is correct. However, after reading the paper, I am not convinced that this is the case. On the contrary I think the result is due to authors' confusion about an individual firm's rationality: maximizing its profit given others' outputs fixed […]' (Referee, *Economic Journal*).

Though neoclassical economics has always insisted that it is a mathematically based theory, a referee for the *Journal of Economics Education* refused to consider that one of the most basic procedures in calculus – the Chain Rule – could be applied in microeconomics: 'Stigler's many attempts to save neoclassical theory have always caused more problems than they have solved. His version of the chain rule is contrary to the partial equilibrium method and thus is irrelevant' (Referee, *Journal of Economics Education*).

These and many other frankly irrational responses by editors and referees for other neoclassical journals emphasize a frequent refrain in this book, that neoclassical economics is far more a belief system than it is a science.

So what?

The main consequence of this critique for neoclassical economics is that it removes one of the two essential pillars of their approach to modeling the economy. Unless perfect competition rules, there is no supply curve.

This fact goes a long way to explaining why neoclassical economists cling to using a notion that is so unrealistic, and so unlike any industry in

the real world: because without it, their preferred method of modeling becomes impossible.

Economics has championed the notion that the best guarantee of social welfare is competition, and perfect competition has always been its ideal. The critiques in this chapter show that economic theory has no grounds whatsoever for preferring perfect competition over monopoly. Both fail the economist's test of welfare, that marginal cost should be equated to price.

Worse, the goal of setting marginal cost equal to price is as elusive and unattainable as the Holy Grail. For this to apply at the market level, part of the output of firms must be produced at a loss. The social welfare ideal thus requires individual irrationality. This would not be a problem for some schools of economics, but it is for the neoclassical school, which has always argued that the pursuit of individual self-interest would lead to the best, most rational outcome for all of society.

Economics can therefore no longer wave its preferred totem, but must instead only derive supply as a point determined by intersection of the marginal cost and marginal revenue curves.

Worse still, once we integrate this result with the fact that the demand curve can have any shape at all, the entire 'Totem of the Micro' has to be discarded. Instead of two simple intersecting lines, we have at least two squiggly lines for the demand side – marginal revenue and price, both of which will be curves – an aggregate marginal cost curve, and lots of lines joining the many intersections of the marginal revenue curve with the marginal cost curve to the price curve. The real Totem of the Micro is not the one shown at the beginning of this chapter, but a couple of strands of noodles wrapped around a chopstick, with lots of toothpicks thrown on top.[17]

There is thus very little left of conventional economic theory. The last two chapters leave very little of the 'Totem of the Micro' standing: in place of the simple intersecting supply and demand curve of conventional neoclassical belief, we have wavy intersecting lines. But even this is too generous to the neoclassical model, because – as Sraffa pointed out almost ninety years ago – there is no empirical justification for the one neoclassical microeconomics concept that we have not yet critiqued: the rising marginal cost curve.

5 | THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING

Why most products cost less to produce as output rises

We have already seen that both the demand and supply aspects of conventional economic analysis are unsound: first, market demand curves don't obey the 'Law' of Demand, and can have any shape at all. Secondly, a supply curve doesn't exist.

But surely, on the supply side, it makes sense that, to elicit a larger supply of a commodity, a higher price must be offered?

There is, in fact, an alternative proposition, which held sway in economics for its first century. This was the argument of the classical school of economics that price was set by the cost of production, while the level of demand determined output.[1] When this proposition is put in the same static form as economics uses to describe a commodity market, it translates as a horizontal (or even a falling) supply curve, so that the market price doesn't change as the quantity produced rises (and it can actually fall). This chapter shows that, though the modern neoclassical position is superficially more appealing and apparently more sophisticated, there are logical problems with it which mean that the classical position is a more accurate description of reality.

The kernel

One of the peculiar aspects of modern Western society is that the majority of the population has no direct experience of how the commodities it consumes are produced. Only a small and decreasing minority is directly involved in production, and only a minority of that minority has direct knowledge of how factories are designed and managed. In contrast to consumption, the conditions under which commodities are produced are a

mystery to most people, and the economic analysis of production appears to illuminate that mystery.

Neoclassical theory argues that, in the 'short run,' productivity falls as output rises, so that higher levels of output result in higher prices. The 'marginal cost curve' therefore slopes upwards, and a higher price has to be offered to entice firms to produce a higher output.

Though this sounds intuitively plausible, when this theory was put to those who do know how factories are designed and managed, they rejected it as 'the product of the itching imaginations of uninformed and inexperienced arm-chair theorizers' (Lee 1998: 73, citing Tucker).

How could something which seems so reasonable to the inexperienced be so absurd according to those 'in the know'? The answer in part lies in the assumptions economists make about production. Though these seem sound to the uninitiated, two key assumptions are in fact contradictory: if one applies for a given industry, then the other one almost certainly does not. When one applies, supply and demand become interdependent; when the other does, the marginal cost curve is likely to be horizontal.

Economic theory also doesn't apply in the 'real world' because engineers purposely design factories to avoid the problems that economists believe force production costs to rise. Factories are built with significant excess capacity, and are also designed to work at high efficiency right from low to full capacity. Only products that can't be produced in factories (such as oil) are likely to have costs of production that behave the way economists expect.

The outcome is that costs of production are normally either constant or falling for the vast majority of manufactured goods, so that average and even marginal cost curves are normally either flat or downward sloping. This causes manufacturers no difficulties, but it makes life impossible for neoclassical economists, since most of neoclassical theory depends on supply curves sloping upwards.

The roadmap

In this chapter I outline the neoclassical analysis of production, which concludes that productivity will fall as output rises, leading to rising costs of production. This in turn leads to a need for the market price to rise if

producers are to supply more, which economists represent as a 'supply curve' that slopes upward in price.

Next I detail Sraffa's argument that two crucial assumptions of this analysis – that supply and demand are independent, and that at least one input to production can't be varied in the short run – are mutually exclusive. A number of potential neoclassical rejoinders to this argument are considered and dismissed.

The outline of the neoclassical model of production below will probably convince you that the theory makes sense, but as with the corresponding section in Chapter 3, it is almost certain to bore you senseless. It is also unavoidably laden with jargon, and less accessible than Chapter 3 since few of us have any experience of production to the same depth as we have experience of consumption. So go back to the coffee pot, pour yourself a strong one, and read on.

Diminishing productivity causes rising price

The neoclassical theory of production argues that capacity constraints play the key role in determining prices, with the cost of production – and therefore prices – rising as producers try to squeeze more and more output out of a fixed number of machines, in what they call 'the short run.' The short run is a period of time long enough to change variable inputs – such as labor – but not long enough to change fixed inputs – such as machines – or for new entrants to come into the industry.

The argument has several stages: stage one puts the proposition that productivity falls as output rises; stage two takes the declining productivity argument and rephrases it as rising costs; and stage three determines the point of maximum profitability by identifying where the gap between revenue and costs is greatest.

Stage one: productivity falls as output rises Neoclassical theory asserts that the supply curve slopes upward because productivity falls as output rises. This falling productivity translates into a rising price. There is thus a direct link between what economists call 'marginal productivity' – the amount produced by the last worker – and 'marginal cost' – the cost of producing the last unit.

Table 5.1 shows an example of production as neoclassicals imagine it. This mythical firm has fixed costs of $250,000, and pays its workers a wage of $1,000.[2] It can sell as many units as it can produce at the market price of $4. To produce output at all, the firm must hire workers: with no workers, output is zero. The first worker enables the firm to produce 52 units of output. This is shown in the first row of the table: the labor input is one unit, and total output is 52 units.

The marginal product of this worker – the difference between production without him (zero) and production with – is 52 units. The marginal cost of the output is the worker's wage – $1,000 – divided by the number of units produced – 52 – which yields a marginal cost of $19.20.

The average fixed costs of output at this point are enormous – $250,000 divided by just 52, or $4,807 per unit. The average total cost is $251,000 divided by 52, or $4,827 per unit – which implies a loss of $4,823 per unit sold, if this were the chosen level of production.

At this stage, production benefits from economies of scale. Just one worker had to perform all tasks, whereas a second worker allows them to divide up the jobs between them, so that each specializes to at least that extent. With specialization, the productivity of both workers rises. The same process continues with the ninth and tenth workers, so that the marginal product of the ninth – the amount he adds to output over and above the amount produced by eight workers – is 83.6 units. Similarly, the marginal product of the tenth worker is 87.5 units.

TABLE 5.1 Input and output data for a hypothetical firm

Labor	Output	Wage bill	Total cost	Marginal product	Marginal cost	Average variable cost	Average fixed cost	Average total cost	Total revenue	Profit
1	52	1,000	251,000	52.0	19.2	19	4,808	4,827	208	−250,792
9	611	9,000	259,000	83.6	12.0	15	409	424	2,443	−256,557
10	698	10,000	260,000	87.5	11.4	14	358	372	2,793	−257,207
100	23,333	100,000	350,000	398.5	2.5	4	11	15	93,333	−256,667
276	131,111	276,000	526,000	772.5	1.3	2	2	4	524,444	−1,556
277	131,885	277,000	527,000	773.7	1.3	2	2	4	527,539	539
400	233,333	400,000	650,000	850.0	1.2	2	1	3	933,333	283,333
500	316,667	500,000	750,000	800.5	1.2	2	1	2	1,266,667	516,667
700	443,333	700,000	950,000	401.5	2.5	2	1	2	1,773,333	823,333
725	452,370	725,000	975,000	323.5	3.1	2	1	2	1,809,479	834,479
730	453,938	730,000	980,000	307.1	3.3	2	1	2	1,815,753	835,753
735	455424	735,000	985,000	290.5	3.4	2	1	2	1,821,698	836,698
740	456,827	740,000	990,000	273.7	3.7	2	1	2	1,827,307	837,307
745	458,144	745,000	995,000	256.6	3.9	2	1	2	1,832,576	837,576
746	458,397	746,000	996,000	253.1	4.0	2	1	2	1,833,588	837,588
747	458,647	747,000	997,000	249.7	4.0	2	1	2	1,834,587	837,587
748	458,893	748,000	998,000	246.2	4.1	2	1	2	1,835,572	837572
800	466,667	800,000	1,050,000	52.0	19.2	2	1	2	1,866,667	816,667

If the firm actually produced this number of units, it would lose $257,207 dollars – more than its fixed costs. However, the process of rising marginal productivity – and therefore falling marginal cost – comes to the rescue as output rises. By the 100th worker, the firm is still making a loss, but the loss is falling because its marginal cost has fallen below the sale price. The 100th worker adds 398.5 units to output, at a marginal cost of

$1,000 divided by 398.5, or just $2.50 a unit. This is less than the sale price of $4 a unit, so the firm is making a profit on the increase in output – but only enough to reduce its losses at this stage, rather than to put it into the black.

Black ink arrives with the 277th worker, who brings in $3,090 profit – the proceeds of selling the 772.5 additional units the worker produces at the sale price of $4 a unit – for the cost of his wage of $1,000.

This process of rising marginal productivity continues right up to the 400th worker hired. By this stage, marginal cost has fallen dramatically. The 400th worker adds 850 units to output, so that the marginal cost of his output is the wage of $1,000 divided by 850, or $1.18 (rounded up to $1.2 in the table). Average fixed costs, which were enormous at a tiny level of output, are relatively trivial at the output level of 233,333 units: they are down to just over a dollar.

From this point on, productivity of each new worker ceases to rise. Each new worker adds less to output than his predecessor. The rationale for this is that the ratio of workers – the 'variable factor of production' – to machinery – the 'fixed factor of production' – has exceeded some optimal level. Now each extra worker still adds output, but a diminishing rate. In economic parlance, we have reached the region where diminishing marginal productivity applies. Since marginal product is now falling, marginal cost will start to rise.

But profit continues to rise, because though each additional worker adds less output and therefore brings in less revenue, the revenue from the additional units still exceeds the cost of hiring the worker. In economic parlance, marginal revenue exceeds marginal cost.

We can see this with the 500th worker, who adds 800.5 units to output. The marginal cost of his output is the wage ($1,000) divided by 800.5, or $1.25 (rounded down in the table). This is higher than the minimum level of $1.18 reached with the 500th worker. But the additional units this worker produces can all be sold at $4, so the firm still makes a profit out of employing the 500th worker.

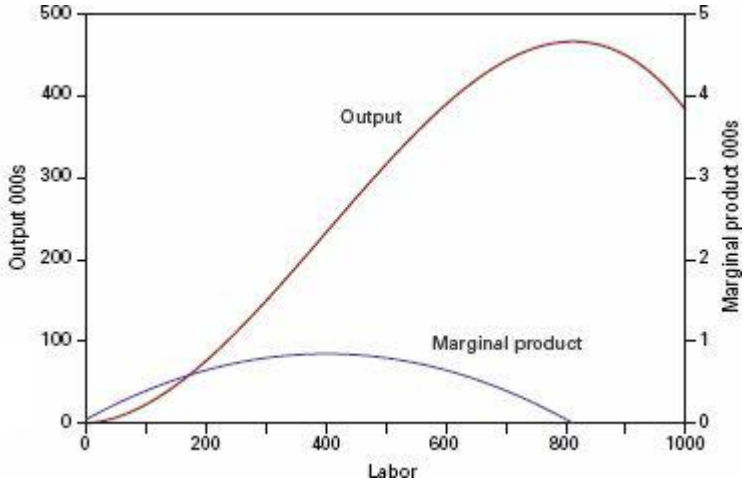
The same principle still applies for the 600th worker, and the 700th. Productivity has dropped sharply now, so that this worker adds only 401.5 units to output, for a marginal cost of $2.50. But this is still less than the

amount the additional output can be sold for, so the firm makes a profit out of this worker.

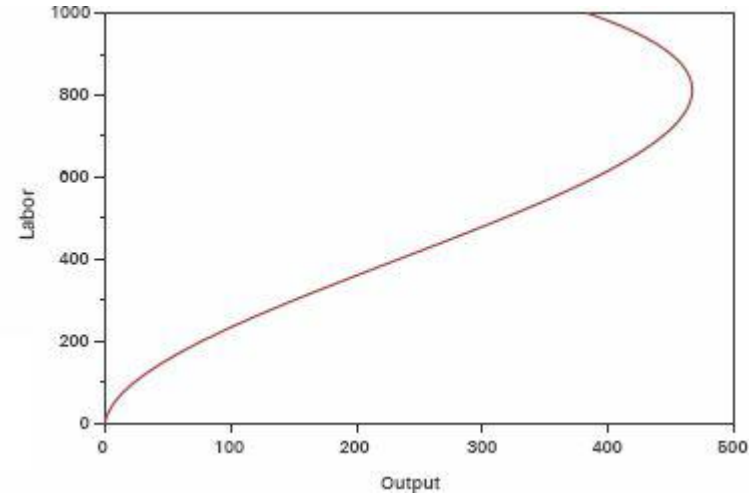
This process of rising profit comes to an end with the 747th worker, whose additional product – 249.7 units – can only be sold for $998.8, versus the cost of his wage of $1,000. From this point on, any additional workers cost more to employ than the amount of additional output they produce can be sold for.

The firm should therefore employ 746 workers, and maximize its profit at $837,588. At this point, the marginal cost of production equals the marginal revenue from sale, and profit is maximized.

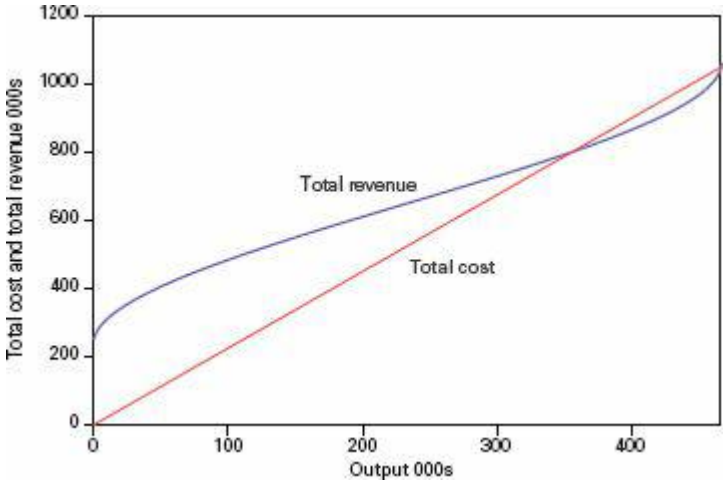
The 800th adds 52 units, for a now soaring marginal cost of $19.20. By the time we get to the 812th worker, workers are – metaphorically speaking – falling over each other on the factory floor, and this worker adds a mere 3.3 units to output, for a marginal cost of $300. The next worker actually reduces output.



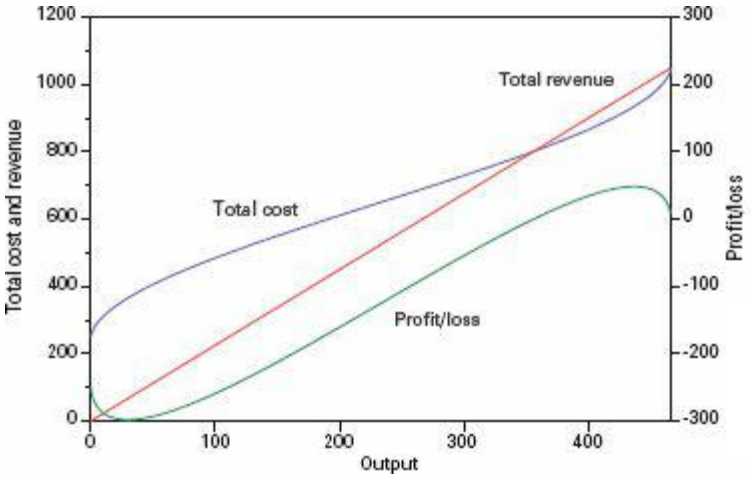
5.1 Product per additional worker falls as the number of workers hired rises



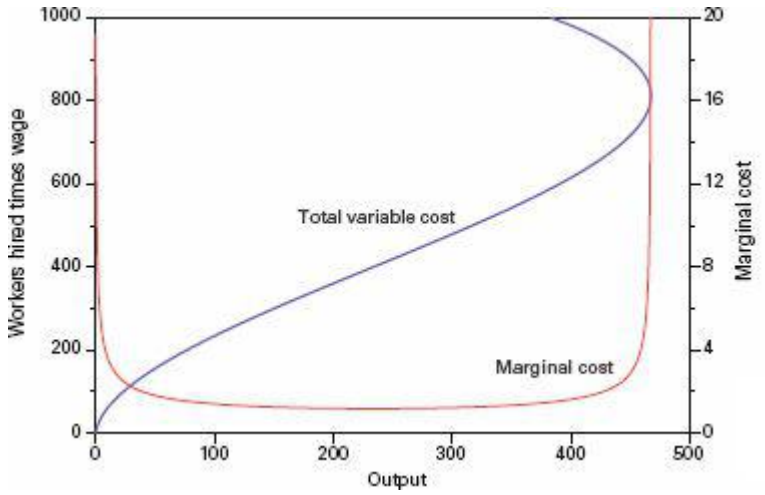
5.2 Swap the axes to graph labor input against quantity



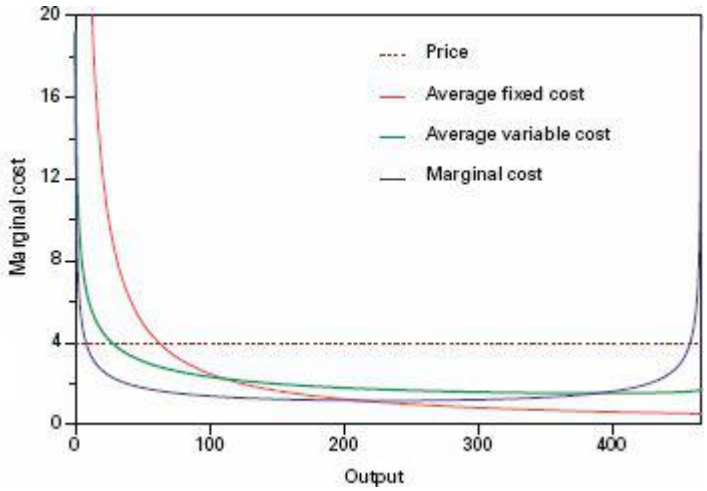
5.3 Multiply labor input by the wage to convert Y-axis into monetary terms, and add the sales revenue



5.4 Maximum profit occurs where the gap between total cost and total revenue is at a maximum

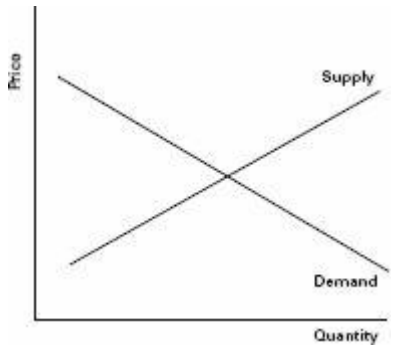


5.5 Deriving marginal cost from total cost



5.6 The whole caboodle: average and marginal costs, and marginal revenue

From minnow to market The exposition above simply describes the situation for a single firm. To derive the market supply curve, we have to aggregate the supply curves of a multitude of producers – just as to complete the derivation of the market demand curve, the demand curves of a multitude of consumers had to be added together.[3] Since each individual firm's marginal cost curve is upward sloping, the market supply curve is also upward sloping.



5.7 The upward-sloping supply curve is derived by aggregating the marginal cost curves of numerous competitive firms

Things don't add up

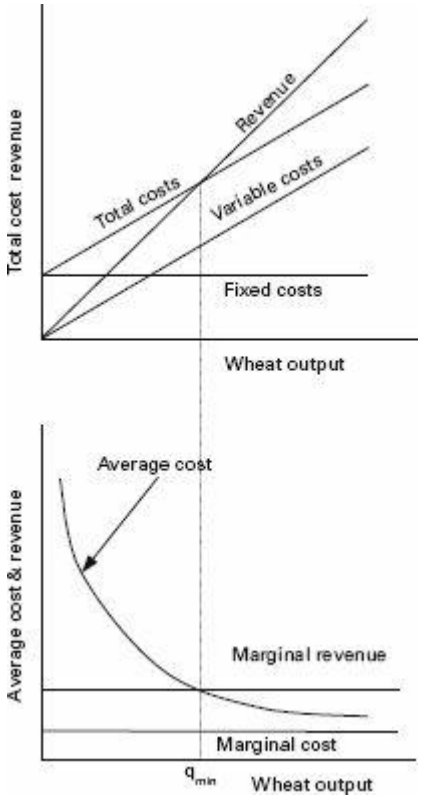
There is no doubt that the economic analysis of production has great superficial appeal – sufficient to explain much of the fealty which neoclassical economists swear to their vision of the market. But at a deeper level, the argument is fundamentally flawed – as Piero Sraffa first pointed out in 1926.

The crux of Sraffa's critique was that 'the law of diminishing marginal returns' will not apply in general in an industrial economy. Instead, Sraffa argues that the common position would be constant marginal returns, and therefore horizontal (rather than rising) marginal costs.

Sraffa's argument constitutes a fundamental critique of economic theory, since, as I've just explained, diminishing marginal returns determine everything in the economic theory of production: the output function determines marginal product, which in turn determines marginal cost. With diminishing marginal productivity, the marginal cost of production eventually rises to equal marginal revenue. Since firms seek to maximize profit, and since this equality of (rising) marginal cost to marginal revenue gives you maximum profit, this determines the level of output.

If instead constant returns are the norm, then the output function instead is a straight line through the origin, just like the total revenue line – though with a different slope. If (as a factory owner would hope) the slope of revenue is greater than the slope of the cost curve, then after a firm had met its fixed costs, it would make a profit from every unit sold: the more units it sold, the greater its profit would be.

In terms of the model of perfect competition, there would be no limit to the amount a competitive firm would wish to produce, so that neoclassical theory could not explain how firms (in a competitive industry) decided how much to produce. In fact, according to the conventional model, each firm would want to produce an infinite amount.



5.8 Economic theory doesn't work if Sraffa is right

This is so patently impossible *within the uncorrected neoclassical model* that, when told of Sraffa's critique, most neoclassicals simply dismiss it out of hand: if Sraffa was right, then why don't firms produce an infinite amount of goods? Since they don't, Sraffa must be wrong.

This knee-jerk response to Sraffa's critique brings to mind the joke that an economist is someone who, when shown that something works in practice, comments, 'Ah, but does it work in theory?' Sraffa instead put the opposite case: sure, the neoclassical model of production works in theory, if you accept its assumptions. But can the conditions that the model assumes actually apply in practice? If they can't, then regardless of how watertight the theory might be given its assumptions, it will be irrelevant in practice. It

therefore should not be used as the theory of production, because above all else, such a theory must be realistic.

Sraffa's argument focused upon the neoclassical assumptions that there were 'factors of production' which were fixed in the short run, and that supply and demand were independent of each other. He argued that these two assumptions could not be fulfilled simultaneously.

In circumstances where it was valid to say that some factor of production was fixed in the short run, supply and demand would not be independent, so that every point on the supply curve would be associated with a different demand curve. On the other hand, in circumstances where supply and demand could justifiably be treated as independent, then in general it would be impossible for any factor of production to be fixed. Hence the marginal costs of production would be constant.

Sraffa began by noting that the preceding classical school of economics also had a 'law of diminishing marginal returns.' However, for the classical school, it was not part of price theory, but part of the theory of income distribution. Its application was largely restricted to the explanation of rent.

The classical argument was that farming would first be done on the best land available, and only when this land was fully utilized would land of a lesser quality be used. Thus, as population grew, progressively poorer land would be brought into use. This poorer land would produce a lower yield per acre than the better land. Diminishing marginal returns therefore applied, but they occurred because the quality of land used fell – not because of any relationship between 'fixed' and 'variable' factors of production.

Sraffa argued that the neoclassical theory of diminishing marginal productivity was based on an inappropriate application of this concept in the context of their model of a competitive economy, where the model assumed that all firms were so small relative to the market that they could not influence the price for their commodity, and that factors of production were homogeneous. In the neoclassical model, therefore, falling quality of inputs couldn't explain diminishing marginal productivity. Instead, productivity could only fall because the ratio of 'variable factors of production' to fixed factors exceeded some optimal level.

The question then arises of when is it valid to regard a given factor of production – say, land – as fixed. Sraffa said that this was a valid

assumption when industries were defined very broadly, but this then contradicted the assumption that demand and supply are independent.

Sraffa's broad arrow If we take the broadest possible definition of an industry – say, agriculture – then it is valid to treat factors it uses heavily (such as land) as fixed. Since additional land can only be obtained by converting land from other uses (such as manufacturing or tourism), it is clearly difficult to increase that factor in the short run. The 'agriculture industry' will therefore suffer from diminishing returns, as predicted.

However, such a broadly defined industry is so big that changes in its output must affect other industries. In particular, an attempt to increase agricultural output will affect the price of the chief variable input – labor – as it takes workers away from other industries (and it will also affect the price of the 'fixed' input).

This might appear to strengthen the case for diminishing returns – since inputs are becoming more expensive as well as less productive. However, it also undermines two other crucial parts of the model: the assumption that demand for and supply of a commodity are independent, and the proposition that one market can be studied in isolation from all other markets.

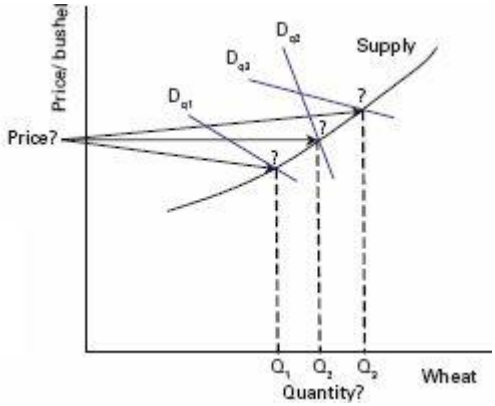
Instead, if increasing the supply of agriculture changes the relative prices of land and labor, then it will also change the distribution of income. As we saw in Chapter 2, changing the distribution of income changes the demand curve. There will therefore be a different demand curve for every different position along the supply curve for agriculture. This makes it impossible to draw independent demand and supply curves that intersect in just one place. As Sraffa expressed it:

> If in the production of a particular commodity a considerable part of a factor is employed, the total amount of which is fixed or can be increased only at a more than proportional cost, a small increase in the production of the commodity will necessitate a more intense utilization of that factor, and this will affect in the same manner the cost of the commodity in question and the cost of the other commodities into the production of which that factor enters; and since commodities into the production of which a

> common special factor enters are frequently, to a certain extent, substitutes for one another the modification in their price will not be without appreciable effects upon demand in the industry concerned. (Sraffa 1926: 539)

These non-negligible impacts upon demand mean that the demand curve for this 'industry' will shift with every movement along its supply curve. It is therefore not legitimate to draw independent demand and supply curves, since factors that alter supply will also alter demand. Supply and demand will therefore intersect in multiple locations, and it is impossible to say which price or quantity will prevail.

Thus while diminishing returns do exist when industries are broadly defined, no industry can be considered in isolation from all others, as supply and demand curve analysis requires.



5.9 Multiple demand curves with a broad definition of an industry

As you can see, Sraffa's argument here was a precursor of the Sonnenschein-Mantel-Debreu conditions that undermine the neoclassical mode of the market demand curve, analyzed from the point of view of the producer rather than the consumer. This allows an upward-sloping supply curve to be drawn, but makes it impossible to derive an independent demand curve.

Sraffa's next argument leaves the demand curve intact, but undermines the concept of a rising supply curve.

Sraffa's narrow arrow When we use a more realistic, narrow definition of an industry – say, wheat rather than agriculture – Sraffa argues that, in

general, diminishing returns are unlikely to exist. This is because the assumption that supply and demand are independent is now reasonable, but the assumption that some factor of production is fixed is not.

While neoclassical theory assumes that production occurs in a period of time during which it is impossible to vary one factor of production, Sraffa argues that in the real world, firms and industries will normally be able to vary all factors of production fairly easily.[4] This is because these additional inputs can be taken from other industries, or garnered from stocks of underutilized resources. That is, if there is an increased demand for wheat, then rather than farming a given quantity of land more intensively, farmers will instead convert some land from another crop – say, barley – to wheat. Or they will convert some of their own land which is currently lying fallow to wheat production. Or farmers who currently grow a different crop will convert to wheat. As Sraffa expressed it:

> If we next take an industry which employs only a small part of the 'constant factor' (which appears more appropriate for the study of the particular equilibrium of a single industry), we find that a (small) increase in its production is generally met much more by drawing 'marginal doses' of the constant factor from other industries than by intensifying its own utilization of it; thus the increase in cost will be practically negligible. (Ibid.: 539)

This means that, rather than the ratio of variable to 'fixed' outputs rising as the level of output rises, all inputs will be variable, the ratio of one input to another will remain constant, and productivity will remain constant as output rises. This results in constant costs as output rises, which means a constant level of productivity. Output will therefore be a linear function of the inputs: increase inputs by 20 percent, and output will rise by the same amount.

Since the shapes of the total, average and marginal cost curves are entirely a product of the shape of the output curve, a straight-line output curve results in constant marginal costs, and falling average costs.

With this cost structure, the main problem facing the firm is reaching its 'break-even point,' where the difference between the sale price and the constant variable costs of production just equal its fixed costs. From that

point on, all sales add to profit. The firm's objective is thus to get as much of the market for itself as it can. This, of course, is not compatible with the neoclassical model of perfect competition.

Irrational managers Sraffa's broad and narrow critiques accept that, if a firm's output was actually constrained by a fixed resource, then its output would at first rise at an accelerating rate, as the productivity of additional variable inputs rose; then the rate of growth of output would reach a peak as the maximum level of productivity was reached, after which output would still rise, but at a diminishing rate. Finally, when even more variable inputs were added, total output would actually start to fall. In the vernacular of economics, the firm would at first experience increasing marginal productivity, then diminishing marginal productivity, and finally negative marginal productivity.

However, Sraffa disputes even this proposition. He instead argues that a firm is likely to produce at maximum productivity right up until the point at which diminishing marginal productivity sets in. Any other pattern, he argues, shows that the firm is behaving irrationally.

His argument is probably best illustrated with an analogy. Imagine that you have a franchise to supply ice creams to a football stadium, and that the franchise lets you determine where patrons are seated. If you have a small crowd one night – say, one quarter of capacity – would you spread the patrons evenly over the whole stadium, so that each patron was surrounded by several empty seats?

Of course not! This arrangement would simply force your staff to walk farther to make a sale. Instead, you'd leave much of the ground empty, thus minimizing the work your staff had to do to sell the ice creams. There's no sense in using every last inch of your 'fixed resource' (the stadium) if demand is less than capacity.

Sraffa argued that the same logic applied to a farm, or to a factory. If a variable input displays increasing marginal returns at some scale of output, then the sensible thing for the farmer or factory owner to do is leave some of the fixed resource idle, and work the variable input to maximum efficiency on part only of the fixed resource.

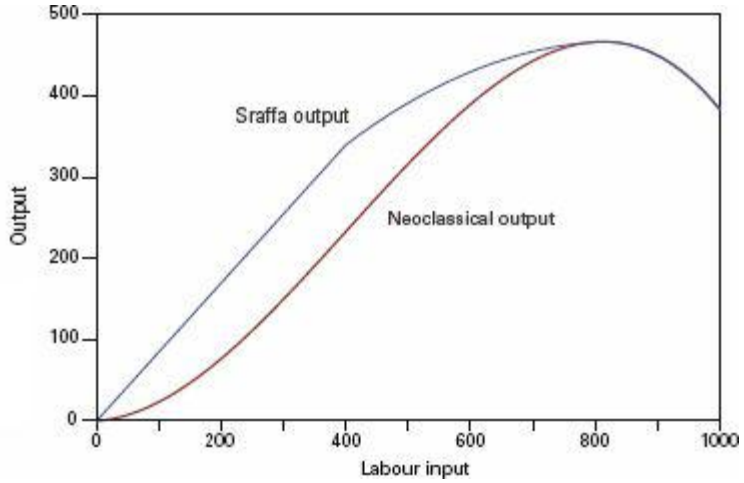
To give a numerical example, consider a wheat farm of 100 hectares, where one worker per hectare produces an output of 1 bushel per hectare, 2

workers per hectare produces 3 bushels, 3 per hectare produces 6 bushels, 4 per hectare produces 10 bushels, and 5 workers per hectare produces 12 bushels.

According to economists, if a farmer had 100 workers, he would spread them out 1 per hectare to produce 100 bushels of wheat. But, according to Sraffa, the farmer would instead leave 75 hectares of the farm idle, and work 25 hectares with the 100 workers to produce an output of 250 bushels. The farmer who behaves as Sraffa predicts comes out 150 bushels ahead of any farmer who behaves as economists expect.

Similarly, economic theory implies that a farmer with 200 workers would spread them over the farm's 100 hectares, to produce an output of 300 bushels. But Sraffa says that a sensible farmer would instead leave 50 hectares fallow, work the other 50 at 4 workers per hectare, and produce an output of 500 bushels. One again, a 'Sraffian' farmer is ahead of an 'economic' one, this time by 200 bushels.

The same pattern continues right up until the point at which 400 workers are employed, when finally diminishing marginal productivity sets in. A farm will produce more output by using less than all of the fixed input, up until this point.



5.10 A farmer who behaved as economists advise would forgo the output shown in the gap between the two curves

This might seem a minor point, but as usual with Sraffa, there is a sting in the tail. If marginal cost is constant, then average cost must be *greater* than marginal cost, so that any firm which sets price equal to marginal cost

is going to make a loss. The neoclassical theory of price-setting can therefore only apply when demand is such that all firms are producing well beyond the point of maximum efficiency. The theory therefore depends on both labor and capital normally being fully employed.

So is full employment – not just of labor, but of other resources as well – the norm in a market economy? If all you read was neoclassical economic theory, you'd believe so – right from the standard textbook definition of economics as 'the study of the allocation of limited resources to unlimited wants.'

Of course, there is recorded unemployment of labor, but neoclassical economists (at least those of the 'freshwater' variety – see pp. 255–66) attribute that to the labor–leisure choice that households make: those who are recorded as unemployed are really deciding that at the wage rates that are on offer, they'd prefer not to work. But surely firms use their capital efficiently, so that it – the 'fixed resource' – is fully employed?

5.11 Capacity utilization over time in the USA

Even a cursory look at the economic data shows that this is not so. Even during the boom years of the 1960s, at least 10 percent of the USA's industrial capacity lay idle; even during subsequent booms, capacity utilization rarely reached 85 percent; and capacity utilization has rarely exceeded 80 percent since 2000, and fell below 70 percent in the depths of the Great Recession (see Figure 5.11).

This situation may seem bizarre from a neoclassical point of view – and there is a trend towards lower utilization over time that could indicate a

secular problem – but it makes eminent sense from a very realistic perspective on both capitalism and socialism put forward by the Hungarian economist Janos Kornai.

Resource-constrained versus demand-constrained economies

Kornai's analysis was developed to try to explain why the socialist economies of eastern Europe had tended to stagnate (though with superficially full employment), while those of the capitalist West had generally been vibrant (though they were subject to periodic recessions). He noted that the defining feature of socialist economies was *shortage*: 'In understanding the problems of a socialist economy, the problem of shortage plays a role similar to the problem of unemployment in the description of capitalism' (Kornai 1979: 801).

Seeing this as an inherent problem of socialism – and one that did not appear to afflict capitalism – Kornai built an analysis of both social systems, starting from the perspective of the constraints that affect the operations of firms:

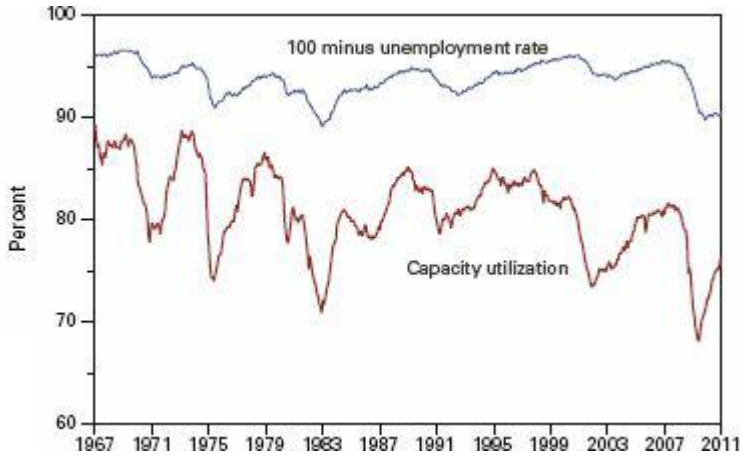
> The question is the following: what are the constraints limiting efforts at increasing production? […] Constraints are divided into three large groups:
>
> 1. Resource constraints: The use of real inputs by production activities cannot exceed the volume of available resources. These are constraints of a physical or technical nature […]
> 2. Demand constraints: Sale of the product cannot exceed the buyer's demand at given prices.
> 3. Budget constraints: Financial expenses of the firm cannot exceed the amount of its initial money stock and of its proceeds from sales. (Credit will be treated later.)[5]
>
> Which of the three constraints is effective is a defining characteristic of the social system […] (Ibid.: 803)

Kornai concluded that 'With the classical capitalist firm it is usually the demand constraint that is binding, while with the traditional socialist firm it

is the resource constraint' (Kornai 1990: 27).

This meant there were unemployed resources in a capitalist economy – of both capital and labor – but this also was a major reason for the relative dynamism of capitalist economies compared to socialist ones. Facing competition from rivals, insufficient demand to absorb the industry's potential output, and an uncertain future, the capitalist firm was under pressure to innovate to secure as much as possible of the industry's demand for itself. This innovation drove growth, and growth added yet another reason for excess capacity: a new factory had to be built with more capacity than needed for existing demand, otherwise it would already be obsolete.

Therefore most factories have plenty of 'fixed resources' lying idle – for very good reasons – and output can easily be expanded by hiring more workers and putting them to work with these idle 'fixed resources.' An increase in demand is thus met by an expansion of both employment of labor and the level of capital utilization – and this phenomenon is also clearly evident in the data.



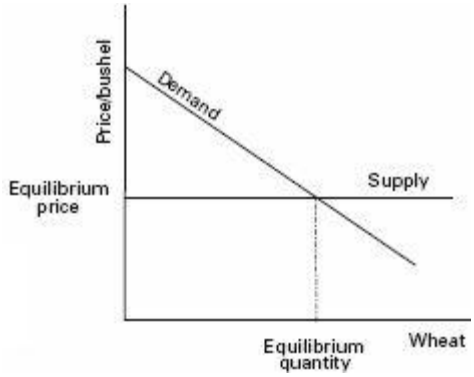
5.12 Capacity utilization and employment move together

Kornai's empirically grounded analysis thus supports Sraffa's reasoning: diminishing marginal productivity is, in general, a figment of the imaginations of neoclassical economists. For most firms, an increase in production simply means an increased use of both labor and currently available machinery: productivity remains much the same, and may even increase as full capacity is approached – and surveys of industrialists, which I discuss later in this chapter, confirm this.

Summing up Sraffa

Sraffa's critiques mean that the economic theory of production can apply in only the tiny minority of cases that fall between the two circumstances he outlines, and only when those industries are operating beyond their optimum efficiency. Then such industries will not violate the assumed independence of supply and demand, but they will still have a relatively 'fixed' factor of production and will also experience rising marginal cost. Sraffa concludes that only a tiny minority of industries are likely to fit all these limitations: those that use the greater part of some input to production, where that input itself is not important to the rest of the economy. The majority of industries are instead likely to be better represented by the classical theory, which saw prices as being determined exclusively by costs, while demand set the quantity sold. As Sraffa put it:

> Reduced within such restricted limits, the supply schedule with variable costs cannot claim to be a general conception applicable to normal industries; it can prove a useful instrument only in regard to such exceptional industries as can reasonably satisfy its conditions. In normal cases the cost of production of commodities produced competitively must be regarded as constant in respect of small variation in the quantity produced. And so, as a simple way of approaching the problem of competitive value, the old and now obsolete theory which makes it dependent on the cost of production alone appears to hold its ground as the best available. (Sraffa 1926)



5.13 Costs determine price and demand determines quantity

If not rising marginal cost, what?

Sraffa's argument dismisses the neoclassical proposition that rising costs and constant (or falling) marginal revenue determines the output from a single firm, or a single industry. This raises the question that if increasing costs don't constrain a firm's output, what does?

Sraffa's argument is simple. The output of a single firm is constrained by all those factors that are familiar to ordinary businessmen, but which are abstracted from economic theory. These are, in particular, rising marketing and financing costs, both of which are ultimately a product of the difficulty of encouraging consumers to buy your output rather than a rival's. These in turn are a product of the fact that, in reality, products are not homogeneous, and consumers do have preferences for one firm's output over another's. Sraffa mocked the economic belief that the limit to a firm's output is set by rising costs, and emphasized the importance of finance and marketing in constraining a single firm's size:

> Business men, who regard themselves as being subject to competitive conditions, would consider absurd the assertion that the limit to their production is to be found in the internal conditions of production in their firm, which do not permit of the production of a greater quantity without an increase in cost. The chief obstacle against which they have to contend when they want gradually to increase their production does not lie in the cost of production – which, indeed, generally favors them in that direction – but in the difficulty of selling the larger quantity of goods without reducing the price, or without having to face increased marketing expenses. (Ibid.)

Economics assumes this real-world answer away by assuming that products are homogeneous, consumers are indifferent between the outputs of different firms and decide their purchases solely on the basis of price, that there are no transportation costs, etc. In such a world, no one needs marketing, because consumers already know everything, and only price (which consumers already know) distinguishes one firm's output from

another. But Sraffa says that these postulates are the exception to the rule which applies in reality.

In most industries, products are heterogeneous, consumers do not know everything, and they consider other aspects of a product apart from price. Even where products are homogeneous, transportation costs can act to give a single firm an effective local monopoly. As a result, even the concept of a competitive market – in which all firms are price-takers – is itself suspect. Instead, most firms will to varying degrees act like monopolists – who, according to neoclassical theory, face a downward-sloping demand curve.

Each firm has a product that may fit within some broad category – such as, for example, passenger cars – but which is qualitatively distinguished from its rivals in a fashion that matters to a particular subset of buyers. The firm attempts to manipulate the demand for its product, but faces prohibitive costs in any attempt to completely eliminate their competitors and thus take over the entire industry. Not only must the firm persuade a different niche market to buy its product – to convince Porsche buyers to buy Volvos, for example – it must also convince investors and banks that the expense of building a factory big enough to produce for both market niches is worth the risk. Therefore, with the difficulty of marketing beyond your product's niche goes the problem of raising finance:

> The limited credit of many firms, which does not permit any one of them to obtain more than a limited amount of capital at the current rate of interest, is often a direct consequence of its being known that a given firm is unable to increase its sales outside its own particular market without incurring heavy marketing expenses. (Ibid.)

Economic theory also can't be saved by simply adding marketing costs to the cost of production, and thus generating a rising marginal cost curve. As Sraffa pointed out, there are at least three flaws with this. First, it is a distortion of the truth – marketing is not a cost of production, but a cost of distribution. Secondly, it is inconsistent with the underlying economic premise that marginal cost rises because of diminishing marginal productivity. Thirdly, it is implausible in the context of the economic theory of the firm. There is no point in 'saving' the concept of a rising marginal

cost curve by introducing marketing costs, since this requires acknowledging that one firm's product differs from another. If products differ from one firm to another, then products are no longer homogeneous, which is an essential assumption of the theory of perfect competition. It is far more legitimate to treat marketing as a cost of distribution, whose object is to alter the demand faced by an individual firm.

Sraffa's critique strengthens the case made in the preceding chapters. Rather than firms producing at the point where marginal cost equals marginal revenue, the marginal revenue of the final unit sold will normally be substantially greater than the marginal cost of producing it, and output will be constrained, not by marginal cost, but by the cost and difficulty of expanding sales at the expense of sales by competitors.[6]

Sraffa's alternative However, Sraffa was not satisfied with this revised picture, which is still dominated by intersecting marginal revenue and marginal cost curves. He instead expressed a preference for a more realistic model, which focused upon those issues that are most relevant to actual businesses.

The firm faces falling average costs as its large fixed costs are amortized over a larger volume of sales, and as its variable costs either remain constant or fall with higher output.[7] It will have a target level of output which it tries to exceed, and a target markup which it tries to maintain. The size of the firm is constrained by the size of its niche within the given market, and the difficulty of raising finance for a much larger scale of operation.

The margin between costs of production and target sale price will be set by the degree of product differentiation within the industry, competitive pressures and general market conditions. Each firm will endeavor to sell as much output as it can, but the level of output will be constrained by the size of the firm's market niche and the marketing efforts of rivals.

5.14 A graphical representation of Sraffa's (1926) preferred model of the normal firm

So what?

To the non-economist, Sraffa's conclusions might still look like fairly minor points. The supply curve should be horizontal rather than upward sloping; the output of an individual firm isn't set by the intersection of marginal revenue and marginal cost; and marketing and finance issues, rather than cost of production issues, determine the maximum scale of a firm's output. This is a big deal?

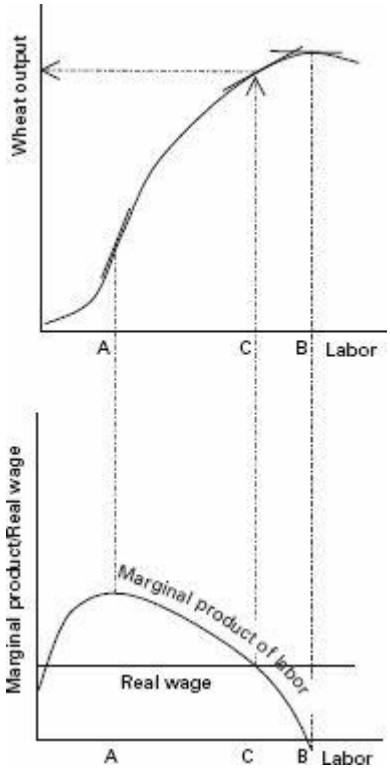
Strange as it may seem, yes, this is a very big deal. If marginal returns are constant rather than falling, then the neoclassical explanation of everything collapses. Not only can economic theory no longer explain how much a firm produces, it can explain nothing else either.

Take, for example, the economic theory of employment and wage determination (discussed in more detail in Chapter 6). The theory asserts that the real wage is equivalent to the marginal product of labor. The argument goes that each employer takes the wage level as given, since with competitive markets no employer can affect the price of his inputs. An employer will employ an additional worker if the amount the worker adds to output – the worker's marginal product – exceeds the real wage. The employer stops employing workers once the marginal product of the last one employed has fallen to the same level as the real wage.

This explains the economic predilection for blaming everything on wages being too high – neoclassical economics can be summed up, as Galbraith once remarked, in the twin propositions that the poor don't work hard enough because they're paid too much, and the rich don't work hard

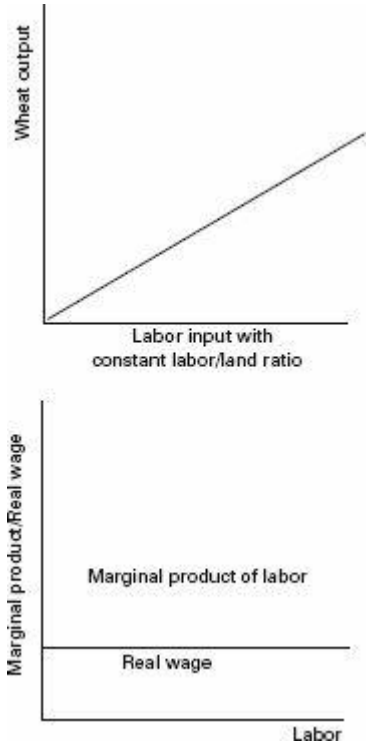
enough because they're not paid enough (Galbraith 1997). The output of the firm is subject to diminishing marginal returns, and thus marginal product falls as output increases. The real wage is unaffected by the output level of the firm. The firm will keep on hiring workers until the marginal product of the last worker equals the real wage.

Since the rational employer stops at that point, the real wage – which the employer takes as given – determines how many workers this firm employs. Since employment in turn determines output, the real wage determines the level of output. If society desires a higher level of employment and output, then the only way to get it is to reduce the real wage (and the logical limit of this argument is that output will reach its maximum when the real wage equals zero). The real wage in turn is determined by the willingness of workers to work – to forgo leisure for income – so that ultimately the level of employment is determined by workers alone.



5.15 The economic theory of income distribution argues that the wage equals the marginal product of labor

If in fact the output-to-employment relationship is relatively constant, then the neoclassical explanation for employment and output determination collapses. With a flat production function, the marginal product of labor will be constant, and it will never intersect the real wage. The output of the firm then can't be explained by the cost of employing labor, and neoclassical economics simply explains nothing: neither the level of employment, nor output, nor, ultimately, what determines the real wage.



5.16 Economics has no explanation of wage determination or anything else with constant returns

Sraffa's critique is thus a fundamental one: if his argument is accepted then the entire edifice of economics collapses.

Clearly, no such thing has happened: economics has continued on as if Sraffa's article was never even published. One might hope that this is because there is some fundamental flaw in Sraffa's argument, or because there is some deeper truth that neoclassical economics discovered to justify preserving the old model with a new explanation. Sadly, neither is the case.

The neoclassical rejoinder

Sraffa's paper evoked several responses from the economic heavyweights of the time. However, these focused upon another aspect of the paper, his critique of the notion of external economies of scale in the long run. Sraffa's primary argument, that the concept of diminishing marginal returns in the short run is invalid, was ignored – so much so that in 1927, 1928 and 1930, Pigou, Robbins and Harrod respectively set out the theory of short-run price determination by rising marginal cost in complete confidence of its validity, and without any reference to Sraffa's paper. Few, if any, conventional economists have since referred to Sraffa's paper.

There are many possible reasons for this complete neglect of a serious challenge to economic orthodoxy. The simplest explanation is that the argument was ignored because its implications, if accepted, were too destructive of conventional economics for neoclassical economists to contemplate. As Chapter 7 argues, this is a not uncommon initial response in all sciences – the key difference with economic 'science' being that this can also be the final response. However, it must be acknowledged that even Keynes – who was, like Sraffa, critical of the mainstream – failed to realize the import of Sraffa's arguments.

The situation has not improved with time. Sraffa's paper is today cited only by critics of economic orthodoxy, while the textbooks teach the theory of rising marginal cost without reference to Sraffa's counter-arguments. It is therefore difficult to put forward a neoclassical response to Sraffa. However, many economists put forward the following arguments when they are informed of Sraffa's paper.

The first is that Sraffa has completely failed to understand the concept of the short run. Neoclassical economics defines three concepts of time: the market period, during which no factor of production can be varied, so that supply is fixed and only price can vary; the short run, during which at least one factor of production cannot be varied, so that output can be varied but only at the cost of diminishing returns; and the long run, during which all inputs can be varied. Since production takes place during the short run, the remainder of the theory follows logically. Diminishing marginal returns will apply, marginal cost will rise, price and quantity will be jointly determined by supply and demand, and the entire edifice of the theory of production and distribution remains intact.

The second is that Sraffa misunderstands the nature of production in a capitalist economy. Since there is enormous pressure to be competitive, no firm can survive long with excess capacity. Therefore competition will drive all firms towards full capacity, and in this realm diminishing returns will apply.

Time and the short run

As Chapter 9 points out, time – or rather, the absence of time in its analysis – is one of the fundamental weaknesses of conventional economics. It is therefore somewhat ironic that economists defend their theory from attack by appealing to the importance of time. However, far from helping to defend economic theory from criticism, the proper analysis of time highlights a critical weakness.

A firm's revenue and costs clearly vary over time, as well as varying as the firm changes its level of output at any one point in time. The economic rule that (in the context of diminishing returns) 'profit is maximized where marginal cost equals marginal revenue' is derived by 'holding time constant' and thus describing revenue and cost as simply a function of the quantity produced. The gap between revenue and cost is widest where marginal cost equals marginal revenue.

But in fact this rule applies only 'when time stands still' – which time never does. Not even an economist can make time stand still (though some victims of economics lectures might dispute that!). Similarly, the rule tells you how to maximize profit with respect to quantity, but real businessmen are more interested in maximizing profit over both time and output.

It is possible to consider profit as a function of both time and quantity, as opposed to the economic approach of dividing time into artificial segments, by explicitly acknowledging that profit is a function of both time and quantity (which the firm can vary at any point in time, and that will also change – and hopefully grow – over time).[8] Profit therefore depends both on the amount a firm produces, and the historical time during which it produces.

Using a rule of mathematics, we can then decompose the change in profit into the contribution due to the progress of time, and the contribution

due to changes in quantity (which will also change over time). This results in the formula:

> Change in profit equals change in profit due to change in time multiplied by the change in time, plus change in profit due to change in quantity multiplied by the change in quantity.

This formula tells us how big a change in profit will be, so if a firm wants to maximize its profit, it wants this number to be as big as possible.

Change in profit due to change in quantity is the same thing as 'marginal revenue minus marginal cost.' Neoclassical theory argues that profit is maximized when marginal revenue equals marginal cost – which we already know is a fallacy – but if you followed the neoclassical profit maximization rule here, you would deliberately set this quantity to zero. Since you get zero when you multiply any number by zero, following this rule sets the second half of the formula (change in profit due to change in quantity multiplied by the change in quantity) to zero.

Therefore, economic theory tells us that the change in profit will be maximized when we eliminate the contribution that changes in quantity make to changes in profit. Change in profit is thus reduced simply to the first half of the formula, where changes due to time alone determine the change in profit. But economic theory has given us no advice about how to make change in profit due to change in time as big as possible.

What's going on? Suddenly, advice that previously seemed sensible (before we considered Sraffa's critique of the notion of a fixed factor of production) looks obviously absurd. Clearly something is wrong: but what?

An analogy might help you interpret it. Imagine you have a formula which describes how much fuel your car uses at any given speed, and you want to work out the most economical speed at which to drive. What you need to do is to work out the lowest rate at which to consume petrol per unit of distance traveled per second. If instead you first work out the most economical speed at which to travel, the answer to this first question will be zero miles per hour – because at this speed you consume the lowest possible amount of petrol per unit of time, zero. This is an accurate but useless answer, since you're not interested in staying put. If you want to work out

the speed that minimizes petrol consumed but still gets you to your destination, you have to handle both problems simultaneously.

The neoclassical theory of the firm ignores time, in the same way that the wrong answer to the 'most economical speed at which to travel' question ignores distance. But time is an essential aspect of economic behavior, in the same sense that distance is an essential aspect of travel. The neoclassical policy for profit maximization is thus false twice: first, it ignores the impact of other firms in an industry on your profit, as the previous chapter showed; then it ignores time. It is thus a kindred spirit to the advice that the cheapest way to get from point A to point B is to travel at zero miles per hour.

There is also an economic way of interpreting this apparent paradox: that advice which appears sound when you ignore (or compartmentalize) time becomes absurd when you take time into account. This is that, by ignoring time in its analysis of the firm, economic theory ignores some of the most important issues facing a firm. Its 'static' emphasis upon maximizing profit 'now' ignores the fact that, to survive, a firm must also grow over time. To grow it must invest and develop new products, and this takes energy and resources. If instead it devotes all its resources to maximizing profit now, then it will not have any energy or resources left to devote to investment and new product development.

If we try to interpret economic theory in the context of historical time, then what the theory is attempting to do is work out the ideal level of output of a product for all time. But in the real world there is no such level of output. The appropriate number of motor cars to produce in 1900 was quite different from the appropriate number to produce in 2000.

This is how something that once looked so right (before Sraffa's critique of the concept of a fixed factor of production) looks so absurd now. The formula discussed above explicitly takes time into account, and is therefore dynamic, while the economic theory of the firm is static: it ignores time, and is therefore only relevant in a world in which time does not matter. But time clearly does matter in our world, and what is right in a static setting is wrong in a dynamic one.

Let's go back to that formula, which is true by definition, and see what it tells us to do.

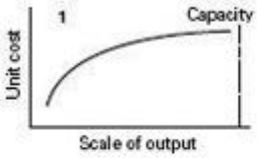
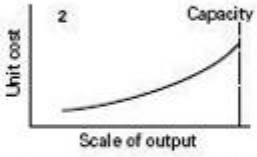
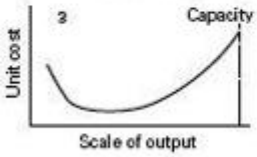
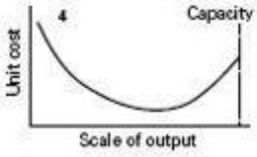
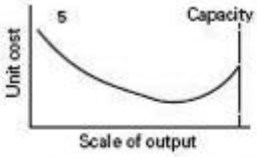
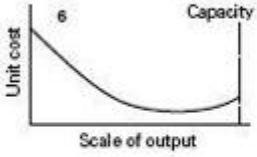
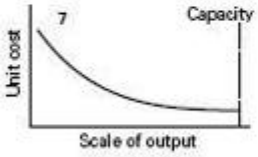
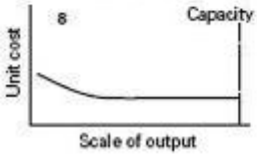
If the firm's output is growing over time, then the term change in quantity will be positive. Setting marginal revenue equal to marginal cost means multiplying this positive number by zero – which results in a smaller increase in profit than if marginal revenue exceeds marginal cost. With rising sales, you will get a higher profit if 'change in profit due to change in quantity' is also positive, which requires that marginal revenue be greater than marginal cost. Thus a careful consideration of time argues that a firm should ensure that its marginal revenue is greater than its marginal cost.

This is the position which Sraffa argued actually applies in reality, so that the proper consideration of time strengthens Sraffa's critique, rather than weakening it. It also strengthens the point I made in the previous chapter that the neoclassical 'short-run profit maximization' formula is false; it's false in the long run too.

This is one of the many instances of the phenomenon I mentioned in Chapter 1, that advice derived from static reasoning, which ignores time, is often categorically opposed to advice derived from dynamic analysis, which takes time into account. Since the economy is fundamentally dynamic, static analysis is therefore normally dangerously wrong. I explore these issues in more depth in Chapter 8.

The flaws in economic reasoning pointed out in this chapter and Chapter 4 have a very direct impact on public policy in the area of the pricing of public services. Because economists believe that competitive industries set price equal to marginal cost, economists normally pressure public utilities to price their services at 'marginal cost.' Since the marginal costs of production are normally constant and well below the average costs, this policy will normally result in public utilities making a loss. This is likely to mean that public utilities are not able to finance the investment they need in order to maintain the quality of services over time. This dilemma in turn interacts with the pressure that economists also apply to privatize public assets, and to let individuals 'opt out' of the public provision of essential services. The end result, as Galbraith so eloquently put it, is 'private affluence and public squalor.'

TABLE 5.2 Cost drawings for the survey by Eiteman and Guthrie (1952: 834–5)

Illustration	Explanation	Chosen by
	Your lowest cost point is at minimum output and unit costs increase rapidly at a decreasing rate as output expands.	0
	Your lowest cost point is at minimum output and unit costs increase slowly at an increasing rate as output expands.	0
	Unit costs are high at minimum output, […] they decline rapidly to a least-cost point, and then […] they rise until capacity output is reached.	1
	Unit costs are high at minimum output, […] they decline to a least-cost point located about midway between minimum and maximum output and then rise.	3
	Unit costs are high at minimum output, […] they decline gradually to a least-cost point near capacity, after which they rise sharply.	14
	Unit costs are high at minimum output, […] they decline gradually to a least-cost point near capacity, after which they rise slightly.	113
	Unit costs are high at minimum output, and […] they decline gradually to capacity at which point they are lowest.	203
	Unit costs are the same at all scales of operation.	0

Ironically, economic theory also makes economists essentially 'anti-capitalist,' in that they deride real businesses for pricing by a markup on cost, when theory tells them that prices should be set at the much lower level of marginal cost. Industrialists who have to cope with these attitudes in their dealings with government-employed economists are often among

the greatest closet anti-economists of all. Maybe it's time for them to come out of the closet.

Competition and capacity The argument that competition would drive all firms to use their fixed capital at full capacity does look convincing at first glance, but deeper thought reaches precisely the opposite conclusion. A firm with no spare capacity has no flexibility to take advantage of sudden, unexpected changes in the market, and it also has to consider building a new factory as soon as its output grows. Excess capacity is essential for survival in a market economy.

Wrong in fact as well as theory

Sraffa's critique was entirely based upon an appeal to logic; a defence which might appear to be open to economic theory was, of course, that the facts supported them rather than Sraffa. However, over 150 empirical studies have been conducted into what the costs of actual firms really are, and with rare unanimity, every last one of them has found that the vast majority of firms report that they have very large fixed costs, and either constant or falling marginal costs, so that average costs of production fall as output rises.

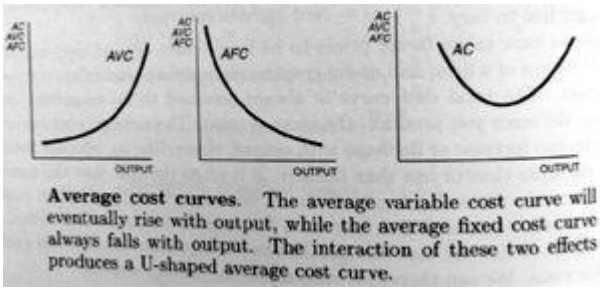
One of the most interesting such studies showed factory managers eight drawings of the shape of cost curves, only three of which bore any resemblance to the standard drawings in neoclassical textbooks (see Table 5.2).

When asked to choose which drawings most closely resembled the relationship between cost and output levels in their factories, only one of the 334 companies chose curve 3, the one that looks most like the curve drawn in virtually every neoclassical microeconomics textbook – for example, this one in Figure 5.17 from Varian's *Microeconomic Analysis* (Varian 1992: 68) – while another seventeen chose curves that looked something like it.

Ninety-five percent of the managers chose drawings that did not conform to the standard textbook model, but instead illustrated either constant or falling marginal costs (Eiteman and Guthrie 1952: 837).

Predictably, neoclassical economists ignored this empirical research – and in fact the purpose of one of the most famous papers in economics,

Milton Friedman's 'as if' paper on methodology (discussed in Chapter 8), was to encourage economists to ignore these empirical results.[9]



Average cost curves. The average variable cost curve will eventually rise with output, while the average fixed cost curve always falls with output. The interaction of these two effects produces a U-shaped average cost curve.

5.17 Varian's drawing of cost curves in his 'advanced' microeconomics textbook

This practice of ignoring empirical research continues today, even though the most recent researcher to rediscover these results was not a critic of neoclassical economics, but one-time vice-president of the American Economic Association and vice-chairman of the Federal Reserve, Alan Blinder (Blinder 1982, 1998). Blinder surveyed 200 medium-to-large US firms, which collectively accounted for 7.6 percent of America's GDP, and he put his results with beguiling honesty:

> *The overwhelmingly bad news here (for economic theory) is that, apparently, only 11 percent of GDP is produced under conditions of rising marginal cost* […]
>
> Firms report having very high fixed costs – roughly 40 percent of total costs on average. And many more companies state that they have falling, rather than rising, marginal cost curves. While there are reasons to wonder whether respondents interpreted these questions about costs correctly, *their answers paint an image of the cost structure of the typical firm that is very different from the one immortalized in textbooks*. (Blinder 1998: 102, 105; emphases added)

The neoclassical model of the u-shaped average cost curve and rising marginal cost is thus wrong in theory and wrong in fact. That it is still taught as gospel to students of economics at all levels of instruction, and believed by the vast majority of neoclassical economists, is one of the best pieces of evidence of how truly unscientific economics is.

TABLE 5.3 Empirical research on the nature of cost curves (summarizing Table 4 in Eiteman and Guthrie 1952: 838)

	By firms	By products
Supports MC=MR	18	62
Contradicts MC=MR	316	1,020
Percent supporting MC=MR	5.4	5.7

A totem in tatters

While the neoclassical model of why production costs rise with output is thus fallacious, it is still feasible that, in some instances, price will rise as output rises. Feasible 'real-world' reasons for this include the inflexibility of supply in some markets across some timeframes (something economics attempts to deal with by its concept of the market period, as opposed to the short-run theory debunked in this chapter), firms exploiting high demand to set higher margins, and, in some circumstances, wage demands rising during periods of high employment.[10] But the neoclassical attempt to link higher prices directly to declining productivity is a failure.

This of itself would not be catastrophic, were it not for the extent to which diminishing marginal productivity permeates neoclassical economics. It is a foundation stone which, when it is withdrawn, brings down virtually everything else with it. Sraffa's critique thus provides one more illustration of the remarkable fragility of this outwardly confident social theory we call economics. Economics is not the emperor of the social sciences, but the Humpty Dumpty.

Just as with Humpty Dumpty after his fall, it is impossible to reconstruct the totemic supply and demand diagram after the criticisms outlined in this and the preceding chapters. First, the Sonnenschein-Mantel-Debreu conditions show that 'diminishing marginal utility,' which in theory applies at the individual level and means that individual demand curves slope downwards, doesn't survive aggregation to the market level – so that a market demand curve can have any shape at all (apart from doubling back on itself, or intersecting itself). Secondly, the marginal revenue curve derived from this demand curve will be even more unstable. Thirdly, equating marginal revenue and marginal cost isn't profit-maximizing.

Finally, diminishing marginal productivity is a theoretical and empirical fallacy, so that for most factories, marginal cost is either constant or falling.

Taken together, these critiques eliminate the 'Totem of the Micro' completely. Virtually every concept in neoclassical microeconomics depends on diminishing marginal productivity for firms on the one hand, and diminishing marginal utility for the community on the other. If both these foundations are unsound, then almost nothing else remains standing. Without diminishing marginal productivity, neoclassical economists cannot explain how a firm decides how much to produce. This alone invalidates their analysis of market structures and income distribution. Without a community utility map, everything from the analysis of optimum output levels to the theory of international trade collapses.

Yet still they teach the standard mantra to their students, and still they apply the same simplistic logic to many other areas of economics.

In the chapters to come, we will temporarily 'forget' the criticisms of these fundamental building blocks, and examine the validity of neoclassical theory as it is applied to specific issues. As you will see, even if we allow, for the sake of argument, that demand falls smoothly as price rises, that production is subject to diminishing marginal returns, and that demand and supply set prices, the neoclassical theories of the distribution of income, the behavior of the macroeconomy, and the role of finance are all intellectually unsound.

6 | TO EACH ACCORDING TO HIS CONTRIBUTION

Why productivity doesn't determine wages

One of the most striking aspects of the late twentieth century was the increase in the gap between the poorest worker and the richest. While many bemoaned this increase in inequality, economists counseled that the growing gap merely reflected the rising productivity of the highly paid.

The basis for this advice is the proposition that a person's income is determined by his contribution to production – or more precisely, by the marginal productivity of the 'factor of production' to which he contributes. Wages and profits – or 'factor incomes,' as economists prefer to call them – reflect respectively the marginal product of labor and of capital. The argument that highly paid workers – merchant bankers, managers of major corporations, stock and money market traders, financial commentators, etc. – deserve the high wages they receive compared to the less highly paid – nuclear physicists, rocket scientists, university professors, schoolteachers, social workers, nurses, factory workers, etc. – is simply an extension of this argument to cover subgroups of workers. Members of the former group, we are told, are simply more productive than members of the latter, hence their higher salaries.

I'll defer discussion of the proposition that profits reflect the marginal productivity of capital until the next chapter. Here we'll consider the argument that wages equal the marginal product of labor.

Once again, the argument relies heavily on concepts we have already dismissed: that productivity per worker falls as more workers are hired; that demand curves are necessarily downward sloping; that price measures marginal benefit to society; and that individual supply curves slope upwards and can easily be aggregated. Even allowing these invalid assumptions, the economic analysis of the labor market is still flawed.

The kernel

Economists prefer to treat everything, including labor, as a simple commodity, subject to the same 'laws of supply and demand' as the simple apple. Yet their own analysis of labor shows that it is fundamentally different. In all other markets, demand decisions are made by consumers and supply decisions by producers. But in the labor market, supply decisions are made by *consumers* (households supplying labor), whereas labor demand decisions are made by *producers* (firms hiring labor). Thus the conventional economic analysis of markets, which is suspect enough on its own terms, is highly unlikely to apply in this most crucial of markets. As a result, wages are highly unlikely to reflect workers' contributions to production, as economists argue.

The roadmap

In this chapter, I outline the economic analysis of labor supply, and the normal economic argument in favor of letting the market decide both wages and the level of employment.

I show that irregularities in the supply of labor – when compared to a normal commodity – are easily derived from this analysis, yet economists unjustifiably assume that labor supply will be an upward-sloping function of the wage. However, these labor market irregularities can make the supply of labor 'backward-bending,' so that reducing wages could actually cause the supply of labor to rise rather than fall.

Though economists normally oppose unions, there are economic arguments in favor of a cartel when sellers (such as workers selling their labor) face a buyer with market power. The opposition economists normally present to unions, to interventionist labor market policies, and to attempts to reduce income inequality are thus shown to be unjustified, even on the grounds of standard economic logic.

Labor demand and supply: an inverted commodity

The economic theory that a person's income reflects his contribution to society relies on being able to treat labor as no different from other commodities, so that a higher wage is needed to elicit a higher supply of labor, and reducing the wage will reduce supply. In fact, economic theory

supports no such conclusion. Even economists can't escape the fact that, as commodities go, labor is something out of the ordinary.

The demand for ordinary commodities is determined by consumer incomes and tastes, while supply is determined by the costs of production. However, unlike other commodities, no one actually 'consumes' labor: instead, firms hire workers so that they can produce other commodities for sale. Secondly, unlike all other commodities, labor is not produced for profit – there are no 'labor factories' turning out workers according to demand, and labor supply certainly can't be said to be subject to the law of diminishing returns (whatever parents might think!).

These two peculiarities mean that, in an inversion of the usual situation, the demand for labor is determined by producers, while the supply of labor is determined by consumers. Demand reflects firms' decisions to hire workers to produce output for sale; supply reflects workers' decisions about how long to work, on the basis of their preferences for income on the one hand and leisure time on the other.

If economists are to argue that the labor market is to behave like all other markets, then these peculiarities must not complicate the usual totemic duet of a downward-sloping demand curve and an upward-sloping supply curve. Unfortunately for economists, they do.

Marginal workers

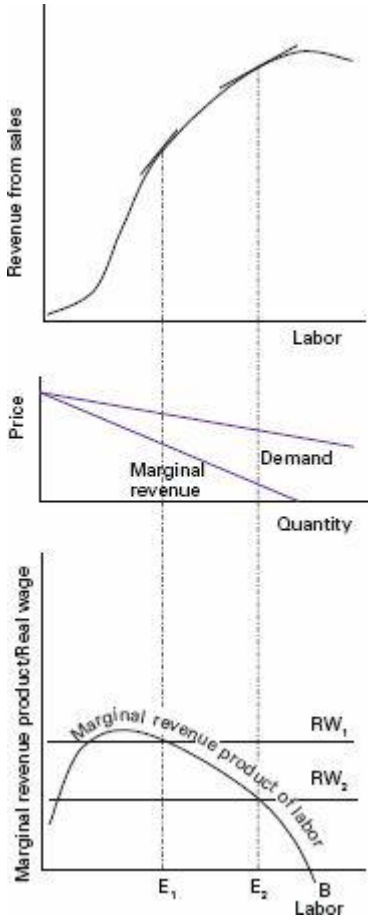
According to economic theory, a firm's labor-hiring decision is determined simply by the impact that each additional worker has on the firm's bottom line. If hiring an additional worker will add to the firm's profit, he is hired; if not, the firm stops hiring.

With a perfectly competitive labor market, the firm can hire as many workers as it wishes to at the going wage. However, since one input (capital) is fixed in the short run, output is subject to diminishing returns: each additional worker hired adds a lesser amount to output than his predecessor. Diminishing marginal productivity therefore rules the hiring roost.

For each firm, the wage is a constant (set by the labor market in which each firm is an infinitesimally small actor). The amount each worker adds to profits, however, is variable. The firm keeps hiring workers up until the

point at which the wage equals the amount for which the last worker's additional output can be sold.

If the industry itself is perfectly competitive, the additional units can be sold without the firm having to reduce its price (yes, I know that's been debunked already; but let's pretend otherwise). In general, the revenue the firm gains by hiring its last employee is equal to the price for which it sells its output, multiplied by the marginal product of the last worker. The firm's demand for labor is therefore the marginal physical product of labor multiplied by the price of the output.



6.1 The demand for labor curve is the marginal revenue product of labor

A disaggregated picture of this is used to explain why some workers get much higher wages than others. They – or rather the class of workers to which they belong – have a higher marginal revenue product than more poorly paid workers. Income disparities are the product of differential

contributions to society, and though sociologists may bemoan it, both the rich and the poor deserve what they get.

Aggregate demand

The demand curves for individual firms are aggregated to form this industry's demand curve for labor, which itself will be a small part of the economy-wide demand curve for labor (since workers can generate many different kinds of output). The real wage is set by the point of intersection of this aggregate demand for labor curve – labor's aggregate marginal revenue product curve – with the aggregate supply curve.

Aggregate supply, in turn, is simply the sum of the supply decisions of individual workers. According to economists, a worker's decision about how much labor to supply is made the same way he decides how much to consume.

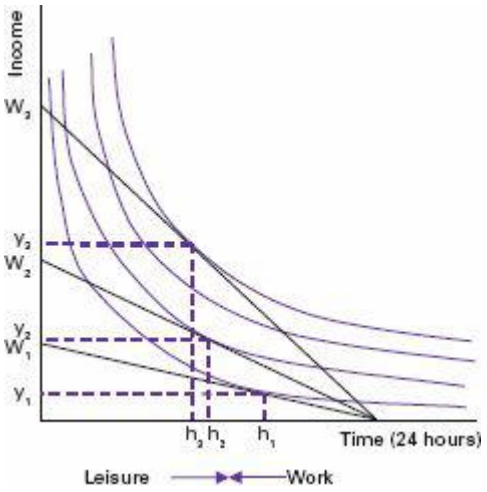
Indifferent workers

Individual labor supply is determined by the individual's choice between work and leisure. Work is a 'bad' in Bentham's calculus: work is a 'pain' while leisure is a 'pleasure.' Therefore the pain of work must be compensated for by the pleasure of the wage, to make up for the sacrifice of leisure required to earn the wage.

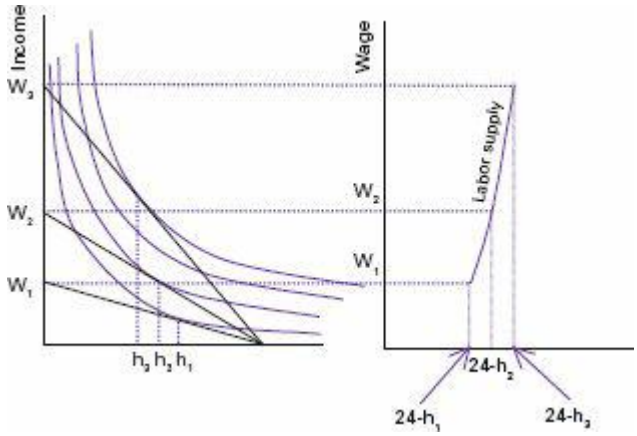
This choice is represented, as always, by indifference curves where potential income is one of the goods, and potential leisure time is the other. The indifference map represents a consumer's preferences between leisure and income, while the budget line represents the hourly wage rate: the higher the wage, the steeper the budget line.

This model has one peculiarity when compared to that applied to normal commodities. With standard commodities, the budget line can be drawn anywhere, so long as it reflects the relative price of the commodities in its slope, and the consumer's income. But with labor, one end of the budget line is fixed at twenty-four hours, since that's the maximum amount of leisure anyone can have in a day. For this reason, all that the budget line can do in this model is pivot about the twenty-four-hour mark, with the slope representing the hourly wage. The distance from zero to the twenty-

four-hour mark represents the maximum possible leisure of twenty-four hours a day.



6.2 The individual's income–leisure trade-off determines how many hours of labor he supplies



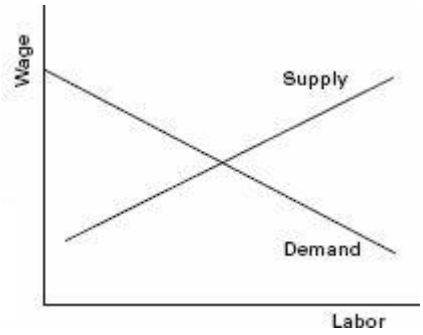
6.3 An upward-sloping individual labor supply curve

As with the consumption of bananas and biscuits, the amount of leisure and income that a consumer will 'consume' is worked out by varying the wage, and seeing what combination of leisure and work the consumer chooses. This generates an individual labor *supply* curve – not a demand curve – from this worker.

The individual supply curve is then summed with that of all other workers to produce the market supply curve. We are back in the familiar economic territory of a downward-sloping demand curve and an upward-

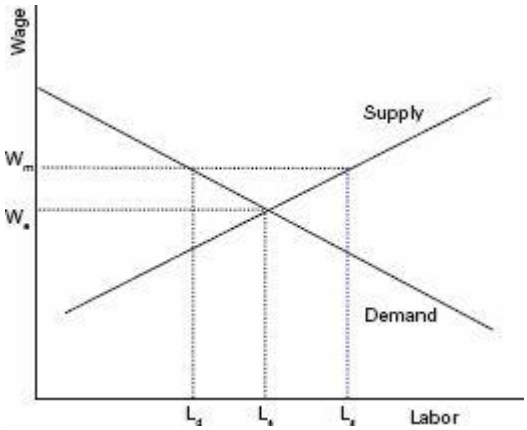
sloping supply curve intersecting to determine an equilibrium price: the average wage. The 'Totem of the Micro' is once again held aloft.

This argument, which strictly speaking applies to labor in the aggregate, is extended by analogy to a disaggregated level in order to explain why some workers get much higher wages than others.



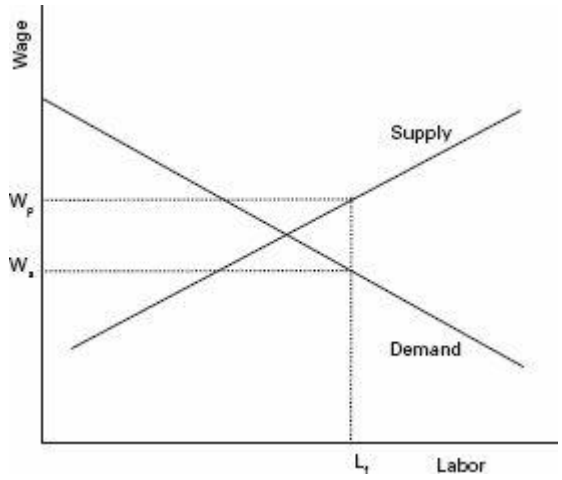
6.4 Supply and demand determine the equilibrium wage in the labor market

At a policy level, this model is used to emphasize the futility of minimum wage legislation, demand management policies, and any other attempts to interfere with the free working of the market mechanism in this most political of markets. If a government attempts to improve workers' incomes by legislating a minimum wage, then this will result in unemployment, because it will increase the number of hours workers are willing to work, while reducing the demand from employers because the wage will now exceed the marginal product of labor. The gap between the increased hours offered and the reduced hours demand represents involuntary unemployment at this artificially high wage level.



6.5 Minimum wage laws cause unemployment

Demand management measures – trying to boost aggregate demand to increase employment – will also fail, because they can't alter the marginal physical product of labor, which can only be done by raising the productivity of labor on the supply side. Attempts to increase aggregate demand will thus merely cause inflation, without increasing the real returns to firms.



6.6 Demand management policies can't shift the supply of or demand for labor

The essential message is that 'you can't beat the market.' Whatever society may think is a fair wage level, or a socially desirable level of unemployment, ultimately the market will decide both income distribution and the rate of unemployment. Moreover, both these market outcomes will be fair: they will reflect individual productivity on the one hand, and the labor–leisure preferences of individuals on the other.

Problems

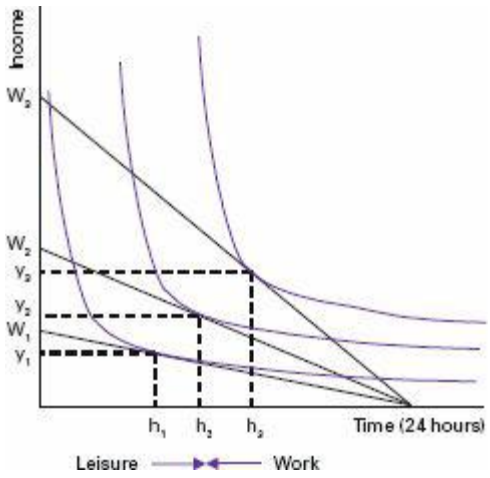
There are at least six serious problems with this meritocratic view of income distribution and employment determination:

- the supply curve for labor can 'slope backwards' – so that a fall in wages can cause an increase in the supply of labor;
- when workers face organized or very powerful employers, neoclassical theory shows that workers won't get fair wages unless they also organize;

- Sraffa's observations about aggregation, noted in Chapter 3, indicate that it is inappropriate to apply standard supply and demand analysis to the labor market;
- the basic vision of workers freely choosing between work and leisure is flawed;
- this analysis excludes one important class from consideration – bankers – and unnecessarily shows the income distribution game between workers and capitalists as a zero-sum game. In reality, there are (at least) three players in the social class game, and it's possible for capitalists and workers to be on the same side in it – as they are now during the Great Recession; and
- most ironically, to maintain the pretense that market demand curves obey the Law of Demand, neoclassical theory had to assume that income was redistributed by 'a benevolent central authority' (Mas-Colell et al. 1995: 117) prior to exchange taking place.

Backward-bending supply curves

Neoclassical economists blithely draw upward-sloping individual and aggregate labor supply curves, but in fact it is quite easy to derive individual labor supply curves that slope downwards – meaning that workers supply *less* labor as the wage rises.

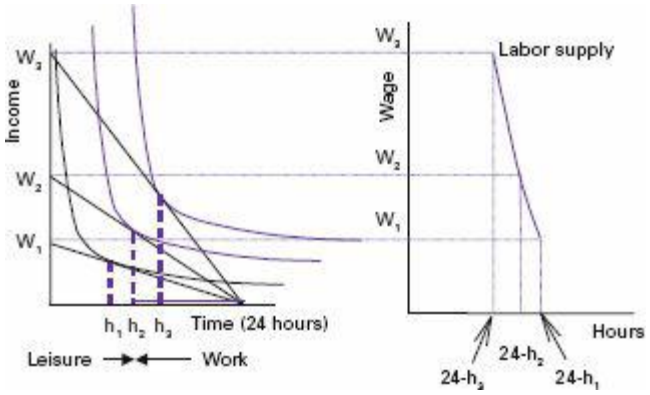


6.7 Indifference curves that result in less work as the wage rises

The logic is easy to follow: a higher wage rate means that the same total wage income can be earned by working *fewer* hours. This can result in an individual labor supply curve that has a 'perverse' shape: less labor is supplied as the wage rises. Economists normally get around anomalies like this by dividing the impact of a higher price into its income and substitution effects – where this time the price of labor is the hourly wage. The substitution effect necessarily means that you'll provide more labor, since each hour of leisure that you forgo gives you a higher return. It's the income effect which stuffs things up – the fact that with a higher wage you can manage to get both a higher income and work fewer hours.

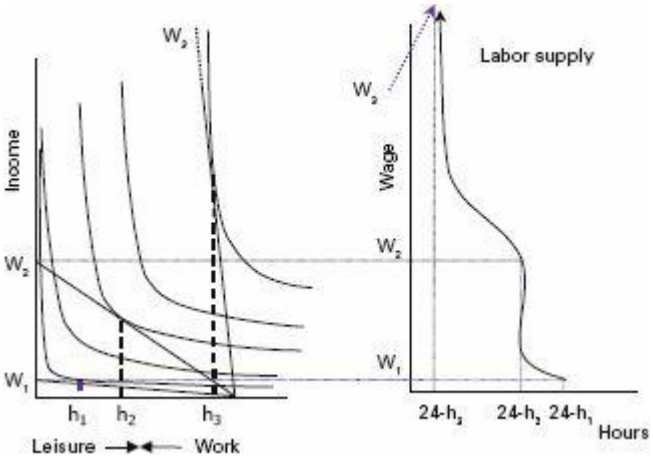
This ruse works when you're considering normal commodities: you simply notionally alter a consumer's income – this was the basis of the 'Hicksian compensated demand curve' that played a role in the proof of the Law of Demand for an individual consumer in Chapter 2. However, this is no use when considering labor supply, because while it's quite easy to notionally add or subtract income from a consumer – thus varying uniformly the amount of both biscuits and bananas that he can consume – it's not possible to add or subtract hours from a day: you can't magically give a worker twenty-eight hours in a day, or take away four.

As a result, it makes no sense to separate the impact of an increase in the wage rate into its substitution effect and income effect: the fact that the substitution effect will always result in an increase in hours worked is irrelevant, since everyone will always have twenty-four hours to allocate between work and leisure.



6.8 Labor supply falls as the wage rises

Since an increase in wages will make workers better off, individual workers are just as likely to work fewer hours as more when the wage rate increases. Individual labor supply curves are just as likely then to slope backwards – showing falling supply as wages rise – as they are to slope forwards.

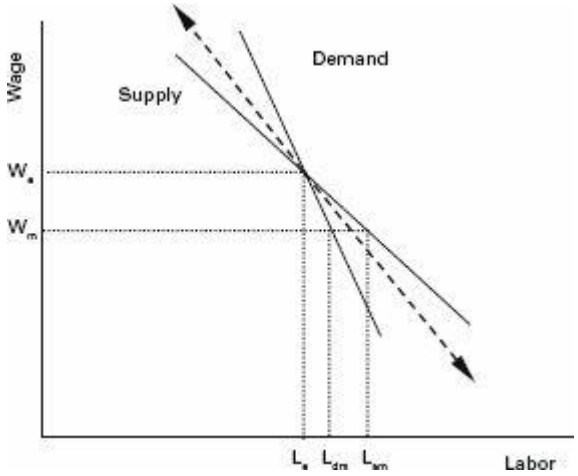


6.9 An individual labor supply curve derived from extreme and midrange wage levels

At the aggregate level, a labor supply curve derived by summing many such individual supply curves could have any shape at all. There could be multiple intersections of the supply curve with the demand curve (accepting, for the moment, that a downward-sloping demand curve is valid). There may be more than one equilibrium wage rate, and who is to say which one is valid? There is therefore no basis on which the aggregate amount of labor that workers wish to supply can be unambiguously related to the wage offered. Economic theory thus fails to prove that employment is determined by supply and demand, and reinforces the real-world observation that involuntary unemployment can exist: that the employment offered by firms can be less than the labor offered by workers, and that reducing the wage won't necessarily reduce the gap.

This imperfection in the theory – the possibility of backward-bending labor supply curves – is sometimes pointed out to students of economics, but then glossed over with the assumption that, in general, labor supply curves will be upward sloping. But there is no theoretical – or empirical – justification for this assumption.

This strong assumption would be of little consequence if economists didn't derive such strong conclusions from their model of the labor market. Declarations that minimum wage legislation is ineffective and causes unemployment, or that demand management policies can't alter the rate of unemployment, are hardly insignificant pronouncements. Their truth is dependent in part on the supply curve for labor being upward sloping.



6.10 An unstable labor market stabilized by minimum wage legislation

For example, if the aggregate demand and supply curves for labor both slope downwards, then the 'equilibrium' of the two could be unstable: falling supply could be met by falling demand, resulting in runaway unemployment. Putting a floor to this process via a minimum wage could actually make the labor market stable and decrease unemployment.

Didactic policy positions should be based upon robust intellectual or empirical foundations, rather than the flimsy substrate of mere fancy. Neoclassical economists are quite prone to dismissing alternative perspectives on labor market policy on this very basis – that they lack any theoretical or empirical foundations. Yet their own policy positions on the labor market are based as much on wishful thinking as on wisdom.

Monopoly and monopsony

The conclusion that workers receive the value of their marginal contribution to output depends upon the assumption that both the product market and the labor market are perfectly competitive. The notion of perfect

competition has already been debunked, but even if it were intellectually sound, it is clearly a dubious thing to assume for an overall economy.

If we instead accept that in practice both product and labor markets will not be perfectly competitive, then economic theory predicts that workers will not, in general, receive the value of their marginal contribution to production. In this more general case, economic theory concedes that workers' incomes are determined not only by their contribution to production, but also by the relative bargaining power of workers and employers.

Let's first consider the case in which the product market is not perfectly competitive: workers are being hired by firms that have to reduce their average selling price to increase output. In this case, the price received per unit falls as output increases. Marginal revenue is thus less than price, and the worker's marginal revenue product is the product of marginal revenue and marginal productivity.

One ironic consequence of this analysis – given how vehemently anti-union most neoclassical economists are – is that neoclassical theory can be shown to favor the existence of trade unions. Without trade unions, the labor supply will be competitive and will therefore be 'exploited,' because the wage will be less than the price for which the marginal worker's output can be sold. With a trade union acting as a single seller of labor, however, the price charged for each additional worker will rise as more workers are hired. This situation – known as a monopsony or single seller – means that the marginal cost of supply lies above the supply curve.

With a monopoly seller of labor confronting non-competitive purchasers of labor, the wage is indeterminate. It will lie between the minimum set by the marginal revenue product of labor (which means that firms are exploiting workers), and the maximum set by the rising marginal cost of workers (which means that workers are exploiting firms). The final position will be determined by the relative bargaining power of the two groups, which cannot be determined by the market.

Thus while economists normally portray unions as bad because they restrict competition in the labor market, this may be a preferable situation to leaving competitive workers to be exploited by less than perfectly competitive hirers of labor.

Sraffa's observations on aggregation

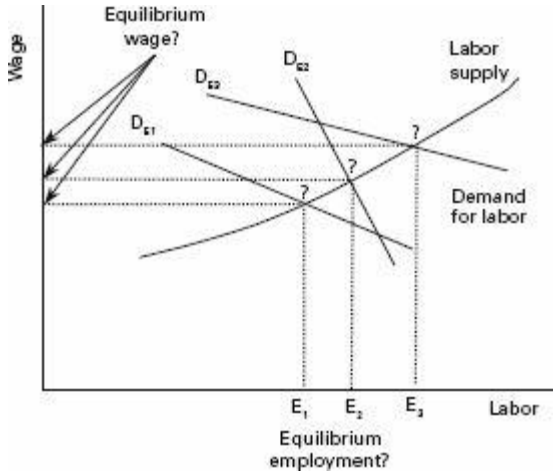
You will remember from Chapter 5 that Sraffa had two criticisms of economic demand and supply analysis: one for a broad definition of an industry, the other for a narrow definition. The labor market is clearly a broadly defined industry, and Sraffa's first critique is therefore relevant to it.

The critique was that, with a broad definition of an industry, it is not feasible to draw independent demand and supply curves, since any change in supply will have income distributional effects which will in turn alter demand.

This is clearly the case when the supply curve refers to the entire labor force. Remember that the aggregate demand curve, in this market, is supposed to represent the aggregate marginal revenue product for labor. This in turn is a product of physical labor productivity on the one hand, and the price for which output produced by that labor is sold.

If an increase in supply requires an increase in the price of labor – if, in other words, the supply curve for labor is upward sloping – then this is clearly going to alter income distribution, the demand for commodities, and hence their prices. This means that a different 'demand curve' for labor will apply at every different point along a labor supply curve.

This means that multiple equilibria will exist, none of which can be said to be more fundamental than any other. It is also quite feasible that 'perverse' outcomes will apply: that, for example, a higher wage could be associated with a higher level of employment rather than a lower one (this dilemma is explored in detail in Chapter 7, in the context of the demand for capital).



The economist's ubiquitous tool of supply and demand analysis is therefore particularly unsuited to analyzing this crucial market.

Freedom and labor

The vision of a worker deciding how many hours to work on the basis of his preferences between income and leisure, and offering more labor as the wage rises, is, like so much else of economic theory, superficially appealing. But, again like so much else in economics, it implicitly raises a question which undermines the superficial appeal. In this case, the question is 'how can one enjoy leisure time without income?'

If there is a positive relationship between the wage rate and hours worked, then as the wage rate falls, so too will the number of hours worked. As a result, income – the product of the hourly wage times the number of hours worked – falls even faster. So according to economists, a fall in the wage rate should mean that workers will substantially reduce their incomes, and simultaneously devote more time to 'leisure activities.'

In reality, the only 'leisure activity' which one can devote more time to with less income is sleeping (just ask a homeless person). Most leisure activities are just that – active – and cost money. The only way that workers could behave as economics fantasizes is if they have alternative sources of income.

This in effect is the economic vision of a worker: someone who has alternative means to generate income at his disposal, and has to be enticed by the wage to undertake wage labor for an employer over the alternative of working for himself.

For that choice to be a reality, workers need something else: capital, his own means of production.

Some workers are so endowed. Some farmers can be enticed into working as farm laborers if the wage is high enough, and if it's not, then they can work their own land. Some office workers have the alternative of working for a wage, or operating as independent consultants out of their home offices. Some 'wage slaves' can make the transition from employee

to employer by an innovative idea, hard work, good luck, skill or good timing – or fraud.

But the majority do not have that choice – or rather don't have it to the degree that they could avoid bankruptcy or starvation by turning to self-employment. For this majority, work is not an option but – in the absence of a very generous social security system – a necessity. Rather than smoothly choosing between work and leisure, in a completely free market system they face the choice of either working or starving. In a market economy attenuated by the welfare state, this choice is less stark, but still present.

A three-horse race

This point will become clearer in later chapters, when I outline the monetary approach to economics that I take, in which bankers are treated as a separate social class to capitalists. The précis for now is that bankers' incomes depend on the level of debt, and if a Ponzi scheme develops, then the level of debt can escalate dramatically. This then transfers income from *both workers and capitalists* to bankers, and to the detriment of society in general since it also normally results in a lower level of real investment.

This issue might seem arcane now, but it has serious implications during a financial crisis, such as the one we are currently in. Neoclassical efforts to get out of such a crisis – once they've gotten over the shock of one actually happening, and revert to form after behaving like 'born-again Keynesians' when the crisis begins – invariably argue that wages have to fall to end the crisis, because high employment clearly indicates that wages are too high.

In fact, policies based on this notion actually make a debt deflation worse, because they drive down the general price level and actually increase the debt burden on society. What is really needed is not lower wages, but lower debt levels – and paradoxically that can be achieved by *increasing* wages. A boost to money wages during a depression can cause inflation far more effectively than 'printing money,' and this inflation can reduce the real debt burden.

If such a policy is ever proposed, you can bet your bottom dollar that the main opposition to it will come from neoclassical economists – and their advice, as always, will be wrong.

'A benevolent central authority'

I've saved the unkindest cut of all for last: even though neoclassical economists are normally vehement opponents of the redistribution of income by the state – everything, they normally argue, should be decided by the market – their own theory of demand and supply only works if, and only if, a 'benevolent central authority' (Mas-Colell et al. 1995: 117) redistributes income in order to 'keep the ethical worth of each person's marginal dollar equal' (Samuelson 1956: 21).

This nonsensical condition is yet another 'proof by contradiction' that neoclassical economics is unsound. Starting from the assumption that the market economy maximizes social welfare, it concludes that this is possible only if, prior to the market operating, a dictatorship redistributes wealth so that everyone in society is happy with the resulting distribution.

This is, of course, absurd. Rather than using neoclassical economics to justify dictatorships, that neoclassical theory literally needs a dictatorship to make its model work is a reason to abandon neoclassical theory. The fact that neoclassical economists not only cling to their theory but argue against income redistribution in policy debates also shows how little they understand their own theory.

Normally this happens because the analysis that establishes bizarre results like this is only in the journal literature that most neoclassical economists don't read – in this case, Samuelson's 1956 paper 'Social indifference curves.' However, here I have to thank Andreu Mas-Colell and colleagues for putting this nonsense in their market-dominating PhD textbook *Microeconomic Theory*, which makes it impossible for neoclassical economists to hide behind their ignorance of their own literature. This section is worth reiterating here, even though I previously cited some of it in Chapter 3:

> For it to be correct to treat aggregate demand as we did individual demand […] there must be a positive representative consumer. However, although this is a necessary condition for the property of the aggregate demand that we seek, it is not sufficient. We also need to be able to assign welfare significance to this fictional individual's demand function. This will lead to the definition of a normative representative consumer. To do so, however, we first have to be more specific about what we mean by the term social

> welfare. We accomplish this by introducing the concept of a social welfare function […]
>
> The idea behind a social welfare function is that it accurately expresses society's judgments on how individual utilities have to be compared to produce an ordering of possible social outcomes […] *Let us now hypothesize that there is a process, a benevolent central authority perhaps, that* […] *redistributes wealth in order to maximize social welfare* […] this indirect utility function provides a positive representative consumer for the aggregate demand function […]
>
> If there is a normative representative consumer, the preferences of this consumer have welfare significance and the aggregate demand function can be used to make welfare judgments […] In doing so however, *it should never be forgotten that a given wealth distribution rule is being adhered to and that the 'level of wealth' should always be understood as the 'optimally distributed level of wealth.'* (Mas-Colell et al. 1995: 116–18; emphases added)

Ahem; please, stop snoring – that was important! In the turgid and boring prose of a neoclassical textbook – and one which has been used in the training of virtually every American PhD student since the late 1990s – you've just been told that neoclassical economics has to assume the existence of a dictator (benevolent of course!).

Most neoclassical economists don't realize this – if they did, they would, I hope, abandon the neoclassical approach as a waste of time. But instead it's likely they don't even read this section of their 1,000-page instruction manual, let alone realize the import of what it says at this point.

I hope you do, however. Certainly, this conundrum makes anything neoclassical economists have to say about the distribution of income irrelevant.

So what?

Few issues provide better examples of the negative impact of economic theory on society than the distribution of income. Economists are forever

opposing ‘market interventions’ which might raise the wages of the poor, while defending astronomical salary levels for top executives on the basis that if the market is willing to pay them that much, they must be worth it. In fact, the inequality which is so much a characteristic of modern society reflects power rather than justice. This is one of the many instances where unsound economic theory makes economists the champions of policies which, if anything, undermine the economic foundations of modern society.

Economics should accept that labor is unlike any other commodity, and develop an analysis suited to its peculiarities, rather than attempt to warp this most personal of markets to fit the conventional cloth of supply and demand.

Keynes did just that in the General Theory. But mainstream economics after Keynes pulled away from this innovation on the basis that Keynes’s argument ‘did not have good microeconomic foundations.’ As this and the preceding three chapters have shown, conventional microeconomic theory itself has unsound foundations. And things get even worse when we turn our attention to problems with the other ‘factor of production,’ capital.

PART 2 | COMPLEXITIES

ISSUES OMITTED FROM STANDARD COURSES THAT SHOULD BE PART OF AN EDUCATION IN ECONOMICS

7 | THE HOLY WAR OVER CAPITAL

Why the productivity of capital doesn't determine profits

The economist Dharma Kumar is said to have once remarked that 'Time is a device to stop everything from happening at once, and space is a device to stop everything from happening in Cambridge.'

Nevertheless, a lot did happen at Cambridge during the 1960s and 1970s, where 'Cambridge' refers to both Cambridge, Massachusetts, USA, and Cambridge, England. The former is home to the Massachusetts Institute of Technology (better known by its initials MIT); the latter is the home of the famous University of Cambridge. MIT was the bastion for the leading true believers in economics, while the University of Cambridge housed an important group of heretics.

For twenty years, these two Cambridges waged a theoretical 'Holy War' over the foundations of neoclassical economics. The first shot was fired by the heretics, and after initial surprise the true believers responded strongly and confidently. Yet after several exchanges, the leading bishop of the true believers had conceded that the heretics were substantially correct. Summing up the conflict in 1966, Paul Samuelson observed that the heretics 'merit our gratitude' for pointing out that the simple homilies of economic theory are not in general true. He concluded that 'If all this causes headaches for those nostalgic for the old time parables of neoclassical writing, we must remind ourselves that scholars are not born to live an easy existence. We must respect, and appraise, the facts of life' (Samuelson 1966: 583).

One might hope that such a definitive capitulation by as significant an economist as Paul Samuelson would have signaled a major change in the evolution of economics. Unfortunately, this was not to be. While many of the bishops have conceded that economics needs drastic revision, its priests

preach on in a new millennium, largely unaware that they lost the holy war thirty years earlier.

The kernel

The term 'capital' has two quite different meanings in economics: a sum of money, and a collection of machinery. Economists assume that they can use the two terms interchangeably, and use the money value of machines as a proxy for the amount of machinery used in production. They prefer to abstract from the complexity that there are many different types of machines, many of which (such as, for example, blast furnaces) are solely suited to producing one particular commodity, and instead work with the generic term 'capital' – as if there is some ubiquitous productive substance which is just as suited to turning out sheep as it is to producing steel. For the economic theories of production and distribution to work, the behavior of this hypothetical generic substance must be little different from the behavior of the actual real world of many different machines.

However, a careful analysis of production as a system by which commodities are produced by combining other commodities and labor shows that the money value of machinery cannot be used as a proxy for the amount of machinery used in production. As a result, the economic theory of how commodities are produced is wrong, and the theory's argument that profit is a reward for capital's contribution to production is also wrong. This reinforces the observations made in Chapter 6, that the distribution of income is not the result of impersonal market forces, but instead reflects the relative power of different social classes.

The roadmap

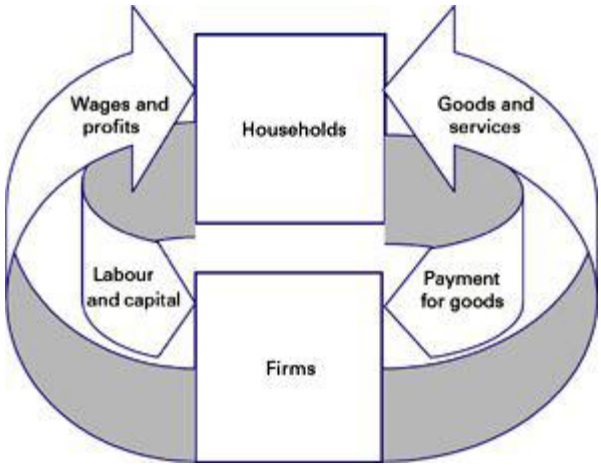
This quite difficult chapter begins with an outline of the economic theory of the production of commodities by 'factors of production,' with its assumption that all machinery can be lumped into the aggregate called 'capital' and measured by the money value placed upon those machines. Then Sraffa's 'abstraction-free' analysis of production is outlined. It is shown that, rather than the rate of profit depending upon the amount of capital, as neoclassical economists argue, the measured amount of capital in fact depends upon the rate of profit.

Measuring capital

Though the war began in earnest only in 1960, the possibility of conflict was first flagged by Piero Sraffa in his 1926 paper 'The law of returns under competitive conditions' (discussed in Chapter 5). In passing, Sraffa observed that an essential aspect of the economic theory of production was the assumption that the interdependence of industries could be ignored. The problem was that this assumption was invalid when changes in one industry's output affected the costs of many other industries, which in turn determined the costs facing the first industry. As Sraffa put it,

> the assumption becomes illegitimate, when a variation in the quantity produced by the industry under consideration sets up a force which acts directly, not merely upon its own costs, but also upon the costs of other industries; in such a case the conditions of the 'particular equilibrium' which it was intended to isolate are upset, and it is no longer possible, without contradiction, to neglect collateral effects. (Sraffa 1926)

Sraffa spent the better part of the next thirty-five years turning this observation into a rigorous theoretical argument. The product was a book with the bland but descriptive title of *The Production of Commodities by Means of Commodities* (Sraffa 1960), and the rather more revealing but still oracular subtitle of 'Prelude to a critique of economic theory.' Essentially, Sraffa provided the techniques needed to highlight fundamental internal inconsistencies in the economic theory of production.



This theory argues that commodities – everything from cornflakes to steel mills – are produced by 'factors of production.' These are normally reduced to just labor on the one hand, and capital on the other. This concept is normally embodied in a 'circular flow diagram' like that of Figure 7.1, which shows factors of production 'flowing' from households to the factory sector, and goods flowing from the factory sector to households.

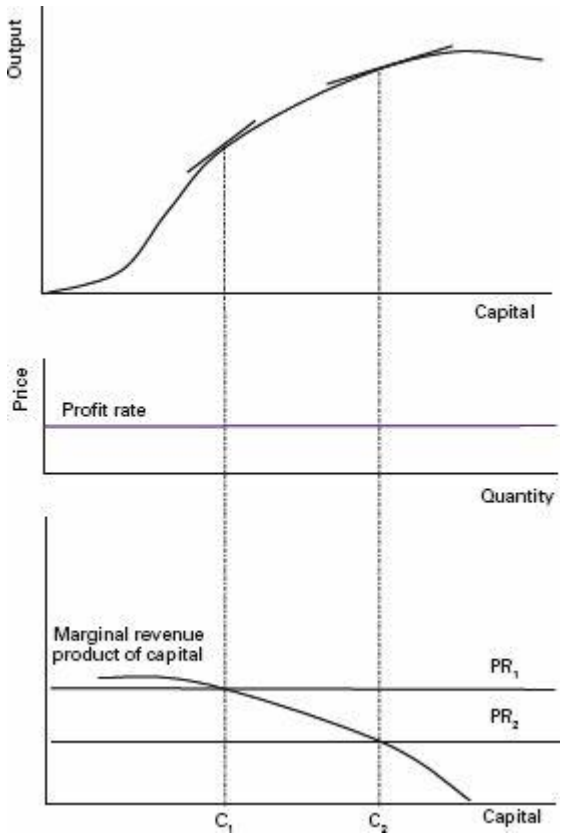
For this flow to be truly circular, households must transform goods into factors of production, while factories must transform factors of production into goods. The factories-to-households half of the circle is reasonable: factories can transform capital and labor inputs into goods. To complete the circle, households must transform the goods they receive from factories into factors of production – labor and capital.

The proposition that households convert goods into labor is unproblematic. However, the questionable proposition is that households also convert goods into capital. This raises a vital question: what is capital, in the context of this diagram? Is it machinery, etc., or is it financial instruments? If it is the former, then this raises the question of where these machines are produced. The model implies that households take goods produced by firms and internally convert them into machines, which are then sold to firms by households. Clearly this is nonsense, since in this case 'households' must also be factories. Therefore, the flow of capital from households to firms must be a financial flow.

However, economic theory treats this financial flow as directly contributing to production: the 'capital' from households to firms generates a profit flow back from firms to households, where that profit reflects the marginal productivity of capital.

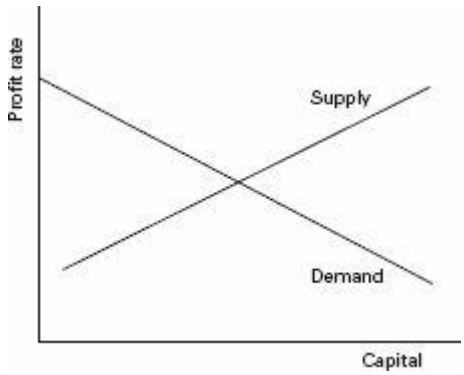
One way this would be possible is if financial instruments directly produced output (in combination with labor) – which clearly they don't.

There is only one other solution, which is to acknowledge that the model is not complete. Factories actually produce capital machines, and this is left out of the diagram. The flow of capital from households to firms is therefore a financial flow, but hopefully there is a direct and unequivocal relationship between the measurement of capital in financial terms and its physical productivity.



7.2 The rate of profit equals the marginal product of capital

A standard 'education' in economics simply ignores these complexities, and explains profit just as it explains wages: the payment to capital represents its marginal productivity. The argument goes that a profit-maximizing firm will hire capital up to the point at which its marginal contribution to output just equals the cost of hiring it. The cost of hiring it is the rate of interest, while its marginal contribution is the rate of profit. The two are equal in equilibrium, so the demand curve for capital slopes downwards – just like all other demand curves – reflecting rising demand for capital as the cost of capital falls.



7.3 Supply and demand determine the rate of profit

The sum of all the individual demand for capital curves gives the market demand curve for capital, while the supply curve – the willingness of households to supply capital – rises as the rate of interest increases. The point of intersection of this downward-sloping demand curve with the upward-sloping supply curve yields the equilibrium rate of profit.

This argument should already be looking somewhat suspect to you, after the previous chapters. For instance, production is supposed to occur in the short run, when at least one factor of production can't be varied. That notion appears at least arguably OK when capital is the fixed factor – though we've shown it to be invalid even there. But it makes no apparent sense to imagine that machinery is now variable while labor is fixed. Surely machinery should be the least flexible factor of production – so that if it can be varied, then everything else can be varied too?

The arguments put by Sraffa against the concept of diminishing marginal productivity can also be applied here in a simple and devastating critique, which was first put formally by Bhaduri in 1969. As with the labor market, the 'capital market' is a broadly defined 'industry': there would be thousands of products being lumped together into the general rubric of 'capital,' and there is no industry which does not use some 'capital' as an input. This raises Sraffa's argument in Chapter 5, that a change in the price of such an input would affect numerous industries, and therefore alter the distribution of income. This is a similar point to that made earlier for the labor market, but it can now be put in a more explicit form.[1]

If we notionally divide all people into either workers or capitalists, then total income will be the sum of wages and profits. Profits in turn are the

product of the rate of profit, times the amount of capital hired. Applying this at the level of the single firm, this gives us the relationship that:

> Income equals
>
> (a) the wage rate multiplied by the number of employees plus
>
> (b) the rate of profit multiplied by the stock of capital

If we now consider changes in output (which we have to do to derive the marginal product of capital), then a rule of mathematics tells us that the changes in output have to equal the changes in wages and profits. Another rule of mathematics lets us decompose the change in profits into two bits: the rate of profit times the change in capital, and capital times the change in the rate of profit.[2] This yields the relationship that:

> Change in income equals
>
> a) change in the wages bill (which we leave aggregated), plus
>
> b) change in profit (which we disaggregate)

Disaggregating changes in profit leads to the statement that:

> Change in income equals
>
> a) change in the wages bill, plus
>
> b) the rate of profit multiplied by the change in capital, plus
>
> c) the amount of capital multiplied by the change in the rate of profit

At the level of the individual firm, economists assume that (a) and (c) are zero: a change in the firm's level of output caused solely by hiring more capital has no impact on either the real wage or the rate of profit. Thus the relationship can be reduced to:

> Change in income equals
>
> a) change in wages [zero], plus

b) the rate of profit multiplied by the change in capital [one[3]], plus

c) capital multiplied by the change in the rate of profit [zero]

Canceling out the terms we know are zero or one yields the desired relationship:

Change in output due to a change in capital (marginal product) equals the rate of profit

However, while this is a reasonable approximation at the level of the individual firm, it is not true at the level of the overall economy. There, any change in capital will definitely have implications for the wage rate, and for the rate of profit. Therefore the aggregate relationship is

Change in output due to a change in capital (marginal product) equals

a) change in wages due to change in capital [non-zero], plus

b) the rate of profit, plus

c) the amount of capital multiplied by the change in the rate of profit due to the change in capital [non-zero]

The rate of profit will therefore not equal the marginal product of capital unless (a) and (c) exactly cancel each other out.[4] Thus at the aggregate level, the desired relationship – the rate of profit equals the marginal product of capital – will not hold true. This proves Sraffa's assertion that, when a broadly defined industry is considered, changes in its conditions of supply and demand will affect the distribution of income.

A change in the capital input will change output, but it also changes the wage, and the rate of profit. These changes alter the distribution of income between workers and capitalists, and will therefore alter the pattern of demand. Exactly the same argument applies to wages, so that in general a person's income will not be equal to their marginal contribution to output. As a result, the distribution of income is neither meritocratic nor determined by the market. The distribution of income is to some significant degree

determined independently of marginal productivity and the impartial blades of supply and demand.

This adds what mathematicians call an additional 'degree of freedom' to the model of the economy. To be able to work out prices, it is first necessary to know the distribution of income; and there will be a different pattern of prices for every different division of the economic cake between workers and capitalists. There is therefore nothing sacrosanct about the prices that apply in the economy, and equally nothing sacrosanct about the distribution of income. It reflects the relative power of different groups in society – though it is also constrained by limits set by the productive system, as we will soon discuss.

This contradicts economic theory, which says that the distribution of income is uniquely determined by the market (via the mechanisms discussed in these two chapters), and therefore there's nothing that policy-makers can or should do to alter it.[5] Instead, rather than prices determining the distribution of income as economists allege, the distribution of income determines prices. Within limits, the distribution of income is something which is determined, not by market mechanisms, but by relative political power.

Bhaduri's critique still accepts the assumption that it is possible to define a factor of production called capital. However, as I intimated above, the machinery aspect of the term 'capital' covers too great a multitude of things to be easily reduced to one homogeneous substance. It includes machines and the buildings that house them; trucks, ships and planes; oil wells, steel works and power stations. Each of these items itself consists of numerous other sub-assemblies which are themselves commodities. A truck contains an engine, which contains valves, springs and cables, the manufacture of which requires inputs from other types of capital, and so on.

The only thing that such disparate commodities obviously have in common is a price, and this is how economists would prefer to aggregate capital. But the price of a piece of capital should depend on the rate of profit, and the rate of profit will vary as prices change: there is an impossible circularity in this method of aggregation.

This problem was explicitly considered by Sraffa in his 1960 magnum opus. His purpose was to provide a firm foundation upon which a critique

of the economic theory of production and income distribution could be built. He built his argument up stage by stage, with great care taken at each stage to make sure that the analysis was sound.

This meticulous method uncovered a number of paradoxes that invalidated the simplistic beliefs economists held about the relationship between productivity and income. Just as the peculiar conditions of 'production' of labor complicate the argument that the wage equals the marginal product of labor, so do the more conventional conditions of the production of capital disturb the argument that profit represents the marginal productivity of capital.

Note: the next section is possibly the most difficult part of this entire book. If you're satisfied with the debunking above, then you can skip this section for now and move to the next chapter. But I do recommend reading this section at some stage.

The whole box and dice

Sraffa's technique was to eschew the initial aggregation of capital, and to say, in place of 'factors of production produce goods,' that 'goods produce goods' – in concert with labor. Sraffa then used this 'assumption-free' model of production to show that the economic theories of price and of income distribution were invalid.

The essential point in his analysis was that capital does not exist as an easily definable entity, yet such an existence is necessary for the simple parable that profit represents the marginal productivity of capital to be true. He made this point by constructing a series of models that directly confronted the true complexity of a system of commodity production.

Sraffa built his models up very carefully, from a simple model with very little real-world realism to a more complex model which, with one exception, was a fairly realistic rendition of a market system of production.

The exception was that Sraffa considered an economy in equilibrium, when a real-world economy is certain not to be in equilibrium. However, Sraffa's purpose was to critique economics on its own terms, and since economics assumes equilibrium, Sraffa made the same assumption. He took it to its logical conclusion, considering an economy which was not only in equilibrium now, but had been in equilibrium for the indefinite past.

Model one: production with no surplus His first model was one in which the economy was just able to reproduce itself, and in which there was no 'fixed capital' – instead, all inputs were 'circulating capital' which are used up in each round of production.

In this economy, the output of each industry was just sufficient to supply the demand for its output by itself and the other industries. Labor was not explicitly treated, but it was feasible to envisage that part of the inputs to an industry represented workers receiving a subsistence wage. Sraffa's example is shown in Table 7.1.

In this hypothetical economy, combining 240 quarters of wheat, 12 tons of iron and 18 pigs in a production process results in an output of 450 quarters of wheat. Similarly, 90 quarters of wheat, 6 tons of iron and 12 pigs are used to produce 21 tons of iron; and 120 quarters of wheat, 3 tons of iron and 30 pigs are used to produce 60 pigs.

The total output of each sector just equals the amount of its output used to produce both its own output and that of all other sectors. Thus the total demand for wheat as an input is 450 quarters: 240 in wheat production, 90 in iron and 120 in pig production.

Sraffa posed the question of what would determine prices in this hypothetical economy, and the answer was not 'demand and supply,' but 'the conditions of production': each sector's price had to enable it to just purchase its inputs. Specifying this for the wheat industry, this meant that 240 times the price of wheat, plus 12 times the price of iron, plus 18 times the price of pigs, had to just equal 450 times the price of wheat.

Similar equations applied for iron and pigs, and with three equations (the price equations for each sector) and three unknowns (the prices), there was one unique set of prices which made it possible for the economy to reproduce.[6]

TABLE 7.1 Sraffa's hypothetical subsistence economy

Industries	Wheat input (qrs)	Iron input (tons)	Pig input (pigs)	Total outputs
Wheat	240	12	18	450 qrs
Iron	90	6	12	21 tons
Pigs	120	3	30	60 pigs
Total inputs	450	21	60	

Neoclassical economists might have endeavored to find this set of prices by considering the demand curves for wheat, pigs and iron, and the supply curves for wheat, pigs and iron, and solving these to find the set of relative prices that equated supply and demand in each industry. However, in this context this would have been overkill: the only prices that work for this economy are those that enable each sector to buy its inputs.

Model two: production with a surplus The next step towards realism was to consider an economy which produced a surplus: where at least one sector produced more of its output than was used up to produce itself and all other commodities. This step closer to a real market economy raises the issue of profits – which weren't an issue in the first model. For this economy to be in equilibrium, the rate of profit has to be the same across all sectors – even if only one sector produced a physical surplus. Otherwise, capitalists in sectors with a low rate of profit would be tempted to move to sectors with a high rate of profit, and the economy would not be in equilibrium. Sraffa used a two-sector example, as shown in Table 7.2.

TABLE 7.2 Production with a surplus

Industries	Wheat input	Iron input	Total output
Wheat	280	12	575
Iron	120	8	20
Total inputs	400	20	

This economy uses 280 quarters of wheat and 12 tons of iron to produce 575 quarters of wheat; another 120 quarters of wheat and 8 tons of iron are used to produce 20 tons of iron. 175 bushels of wheat are produced over and above the 400 used in production, whereas the entire 20 tons of iron are used up in producing wheat and iron.

For a uniform rate of profit r to apply, the prices in this economy must be such that the 'money' value of inputs, multiplied by $(1+r)$, must equal the money value of its outputs. For this example economy, the price ratio is 15 bushels of wheat for 1 ton of iron, and the uniform rate of profit is 25 percent.

Model three: production with a surplus and explicit labor The economy above had to have labor in it, since nothing can be produced without labor.[7] However, this was not explicitly shown. The next model added further realism by showing that output was produced by combining both commodities and labor in a production process.

This introduces the wage as an additional unknown, and establishes the first element in Sraffa's critique of the economic theory of income distribution: rather than prices determining the distribution of income, the distribution of income between wages and profits must be known before prices can be calculated.[8]

Sraffa then shows that there is an appropriate measuring stick (the 'standard commodity') which reveals a simple, linear relationship between the wage w, the actual rate of profit r, and the maximum feasible rate of profit for a given economy, R.[9] The wage w falls linearly as the rate of profit r rises towards its maximum value R.

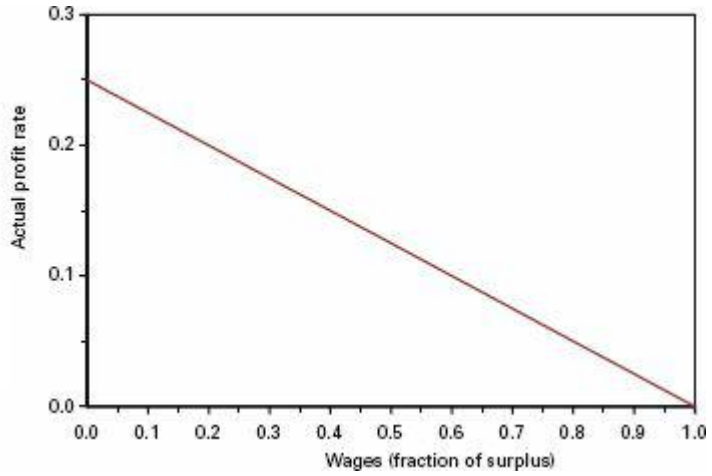
The example economy in Table 7.2 has a maximum rate of profit of 25 percent, and results in the wage/profit function shown in Figure 3.1. If the wage w is .8 – which means that workers' wages represent 80 percent of the surplus output of this economy – then the corresponding rate of profit r is 5 percent. This is shown numerically in Table 7.3.

TABLE 7.3 Relationship between maximum and actual rate of profit and the wage share of surplus

Maximum R Wage (% of surplus)	25% Profit rate
0	25
10	23
20	20
30	18
40	15
50	13
60	10
70	8
80	5
90	3
100	0

What this table says is that if workers, for example, get a zero wage, then all of the surplus goes to the capitalists, who then make a profit of 25

percent. If, however, workers get 10 percent of the surplus as their wage, then the rate of profit falls to 23 percent (rounded up). The same linear process continues right out to the point at which workers get 100 percent of the surplus, at which point capitalists get nothing and therefore have a rate of profit of zero.



7.4 The wage/profit frontier measured using the standard commodity

Clearly, this analysis is reasonably realistic, and therefore, one might think, rather innocuous. However, this apparently innocuous step sets up the *coup de grâce* for the economic theory of income distribution.

The punchline: capital behaving badly

The key concept in the neoclassical theory of income distribution is that factors get paid in accordance with their marginal contribution to output in the context of diminishing marginal returns. This means that as the supply of a factor increases, its return should fall.

The difficulty is, as alluded to earlier, that it is not easy to see how one can add units of capital together. Workers can be aggregated by adding up the number of hours they work – after notionally standardizing for different levels of productivity by multiplying the hours of skilled labor by some amount to reflect higher productivity. Land can be aggregated by adding up acres – and again by adjusting numerically for varying degrees of fertility.

But machines have no apparent common property apart from price. This is in fact how economic theory aggregates capital, but this involves an

obvious circularity, because the price of a machine reflects the profit expected from it, yet the rate of profit is the ratio of profit to price.

Sraffa proposed an ingenious and logically sound method of aggregation: to reduce capital to dated inputs of labor. The previous linear relationship between the wage and the rate of profit was an essential element in this analysis.

All items of capital are produced by other items of capital and labor. When an economy has been in equilibrium for the indefinite past, it is thus possible to regard the value of a machine as being equal to the value of the machines used to produce it, plus the value of the labor involved, times a rate of profit to reflect the passage of time. If we notionally treat the period of production as a year, then if the equilibrium rate of profit is 5 percent, 1.05 times the value of the inputs last year should equal the value of the machine this year.

The same argument applies to all the machines and labor inputs used to produce the inputs, and to all the machines and labor that produced them, and so on.

If we repeat this process, and each time reduce machinery inputs to the machinery and labor used to produce them, then we get a set of labor terms and a declining – *but never zero* – residual of machinery inputs. Each labor input is multiplied both by the wage, and by one plus the rate of profit raised to a power which reflects how many years ago the input was made.

If, for example, we are considering a machine manufactured eleven production periods ago, then this term will be the amount of direct labor bestowed in producing all the relevant components in the twelfth year, times the wage, plus the capital input, all raised to the twelfth power. It is therefore possible to substitute an expression in terms of labor for the capital inputs used up in producing a given commodity.[10]

TABLE 7.4 The impact of the rate of profit on the measurement of capital

Profit rate (%)	Years								
	0	1	2	3	4	5	10	20	25
0	1	1	1	1	1	1	1	1	1
1	0.96	0.97	0.98	0.99	1.00	1.01	1.06	1.17	1.23
2	0.92	0.94	0.96	0.98	1.00	1.02	1.12	1.37	1.51
3	0.88	0.91	0.93	0.96	0.99	1.02	1.18	1.59	1.84
4	0.84	0.87	0.91	0.94	0.98	1.02	1.24	1.84	2.24
5	0.80	0.84	0.88	0.93	0.97	1.02	1.30	2.12	2.71
10	0.60	0.66	0.73	0.80	0.88	0.97	1.56	4.04	6.50
20	0.20	0.24	0.29	0.35	0.41	0.50	1.24	7.67	19.08
21	0.16	0.19	0.23	0.28	0.34	0.41	1.08	7.24	18.78
22	0.12	0.15	0.18	0.22	0.27	0.32	0.88	6.40	17.31
23	0.08	0.10	0.12	0.15	0.18	0.23	0.63	5.03	14.15
24	0.04	0.05	0.06	0.08	0.09	0.12	0.34	2.95	8.66
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

We can now approximately[11] express the value of a machine in terms of the sum of the value of the labor inputs used to produce it. Each element in this sum consists of a physical quantity of labor, multiplied by two terms: one representing the wage, and another representing the impact of accumulated profit over time.

The former term is a negative function of the rate of profit (as in Table 7.3); the latter is as a positive function of the rate of profit, raised to a power. The former will fall in size as the rate of profit rises; the latter will rise, and it will also rise more for inputs made a long time ago.

This combination of opposing effects – one term that falls as r falls, the other that rises as r falls – evokes the possibility that one effect can prevail for a time, only to be overwhelmed by the opposite effect at a higher rate of profit. Therefore, the individual terms that interact to determine the value of an item of capital can rise for a while as the rate of profit rises, only to fall as the rate of profit rises still further.

This can be illustrated using Sraffa's example economy where the maximum rate of profit was 25 percent, and considering a machine which was made using one unit of labor as an input at some time in the past.

If the rate of profit was zero, then no matter how many years ago that machine was made, if a machine cost one (standard commodity) unit to make, its measured value would still be 1, as shown by the first row of Table 7.4. If the rate of profit was instead 1 percent, then the measured

value of that machine falls to 0.96, if it is used today – reflecting the lower value of labor in terms of Sraffa's measuring stick.

The value of the machine rises a bit if it was made two years ago, because its value is calculated to be 0.96 times 1 plus the rate of profit. This is 0.96 times 1.01, or roughly 0.97. This larger amount, though, is still less than 1, which would have been its value if the rate of profit had been zero. The same applies if the machine was used two periods ago, in which case its calculated value would be 0.98 – or 0.96, multiplied by 1.01 squared.

However, if the machine was produced five years ago, then its value in terms of the standard commodity rises to 1.01. This is because, while one part of the overall term has fallen to 0.96, the other has risen to 1.01 multiplied by itself five times – which roughly equals 1.05 – and 1.05 times 0.96 gives us 1.01.

The same effect applies across the row of the table, showing that as the rate of profit rises, the measured value of this capital input rises. The second term, 1.06, is 0.96 times 1.05 raised to the 10th; the third, 0.96 times 1.05 raised to the 15th; and so on.

The measured value of the machine therefore falls because of a higher rate of profit, but then rises if it was used many years ago. And the table has even more complications.

Notice that as we go down the table – so that the rate of profit increases – the value of a machine input today falls smoothly. However, the value of a machine applied five years ago rises for a while, but then falls. This accurate picture is a lot more complicated than economists assumed it to be, and these complications rule out the simple correspondence economists believed existed between the 'amount' of capital and the rate of profit.

The complications arise because the two different effects in Sraffa's accurate measure of capital don't cancel each other out. The first is the value of a wage unit, given the rate of profit r. On the first row, that is 1 (reflecting a zero rate of profit); on the second, 0.96 (at a 1 percent rate of profit); the third, 0.92 (at a 2 percent profit rate); and so on. But the second effect is $1+r$, raised to a power of 5, reflecting how many years ago the input was made. On the first row, that term is 1 – because the rate of profit is zero – and 1 times 1 is 1. On the second row, it is 0.96 times 1.05, which is 1.01 raised to the fifth power. This is roughly 1.01, so the measured value of the machine has risen. On the third row, it has risen further to 1.02 –

which is 0.92 times 1.1, which is 1.02 raised to the 5th. On the fourth, it is roughly the same – 0.88 times 1.16, which is 1.03 raised to the 5th.

But by the time we get to a 10 percent rate of profit, the value goes down to 0.97: here we have 0.6 times 1.61, which is 1.10 raised to the 10th. The impact of the falling value of the first term now outweighs the impact of the rising value of the second. By the time we get to a rate of profit of 20 percent, the value of this machine (in terms of the standard commodity) has fallen to just 0.5, having been as high as 1.02 at lower rates of profit.

So the measured value of a machine rises and then falls as the rate of profit rises, and also rises and then falls as the time at which the machine was used to produce a commodity becomes farther in the past.

This is not exactly how economists think about capital as a factor of production. They had hoped that the rate of profit would fall smoothly as the amount of capital used in production rose, so that capital, like labor, would manifest diminishing marginal productivity. But Sraffa instead showed that not only was there no uniform relationship between the rate of profit and the amount of capital, but also the direction of causation was the opposite of what economists wanted. Rather than the rate of profit depending on the 'amount' of capital, the measured amount of capital actually depended on the rate of profit. This makes it impossible to argue that the rate of profit is determined by the marginal productivity of capital, and so this second leg of the economic theory of income distribution collapses.

Not only that, but the perverse relationship that exists between the measurement of capital and the rate of profit is going to cause perverse effects in production. A rising rate of profit might for a while make one method of producing a commodity cheaper than alternatives, but then at a still higher rate of profit, it might make it more expensive.

Sraffa provides one illustration of this by comparing the price of two commodities which start out equal when the rate of profit is zero, and where one becomes more expensive than the other as the rate of profit rises, only to have the other become more expensive as the rate of profit rises farther still. One product has relatively more 'direct labor' applied to its production in the recent past, while the other has more direct labor applied in the far distant past. Sraffa likens the latter to wine produced by being aged in a barrel; the former could be regarded as producing wine of identical quality

using advanced chemical processes.[12] The latter process would be regarded as 'capital intensive,' since so much machinery is used directly in its production, while the former would be called perhaps 'time intensive' (or labor intensive if you imagine the barrels being tended over the years by cellar masters).

At a zero rate of profit, the cost of each barrel of wine equals simply the sum of the wages paid to produce the wine – and for both methods of production to exist in equilibrium, the cost of the two techniques must be identical.

As the rate of profit rises from zero to a moderate uniform rate, the far distant application of labor needed to produce the barrel has comparatively little impact, so that the wine produced using modern technology is more expensive. In this range of the rate of profit, production using modern technology would cease, since it would be uncompetitive with wine produced using the aging process.

However, as the rate of profit becomes higher still, the effect of compounding the rate of profit on the making of the cask becomes enormous, so that the aged wine becomes more expensive than its mass-produced cousin. Mass production would take over again – we would switch back to the apparently more 'capital intensive' means of production.

Finally, when the rate of profit reaches its maximum value and wages fall to zero, the cost of wine falls to simply the cost of the irreducible commodity components (the original grapes, etc.), and the price of the two types of wine could again coincide.

Subsequent economists used Sraffa's building blocks to illustrate that a method of production could start out superior to all others at a zero profit rate, become less profitable than some other methods at a higher rate, only to once again become the most profitable at a higher rate still.

This phenomenon of 'reswitching' destroyed the simple proposition that the rate of return on capital represented the marginal product of capital. If a particular production technique had lost primacy to others at one rate of profit, then it could not regain that primacy at a higher rate of profit still, unless for a period it benefited from increasing marginal product. But if marginal product could alternately rise and fall, then there was no necessity that the market for capital should be well behaved. Demand curves could

slope up as well as down, supply curves down as well as up, and no unique equilibrium position could be defined.

The causes of this apparent paradox are that the concept of capital as a homogeneous substance is an illusion, and that what is capital intensive depends on the rate of profit. If the rate of profit is low, then the labor embodied in an ancient wine barrel is of little consequence, and the process of aging wine may well appear to be labor intensive. But if the rate of profit is high, then compounding of this high rate of profit makes that ancient wine barrel of great value – and the process could be described as capital intensive. Rather than the rate of profit depending on the quantity of capital, the quantity of capital (in terms of its value measured in embodied labor value) depends upon the rate of profit.

The intricate and interdependent processes of production thus generate many opportunities for factor returns to move one way and then the other as factor intensities rise. There is therefore no consistent relationship between factor productivity and factor incomes. Instead, the distribution of income between wages and profits is largely independent of the system of production. The distribution of income is a social phenomenon.

Economists fought against this conclusion, but every apparent victory was shown to be invalid. Ironically, the rebuttals to economic rejoinders often showed that the only conditions under which the economic position could hold would be if the ratio of capital to output was the same in all industries. This is the same condition needed to make Marx's labor theory of value hold, yet the neoclassical revolution which gave us modern economic theory was supposedly free of the nonsense conditions needed by its Marxian rival.

So what?

Just as Chapter 6 showed that the wage can't be explained as the marginal product of labor, this chapter has established that economic theory cannot justify the existing rate of profit as somehow reflecting the marginal productivity of capital. Instead, the rate of profit reflects relative power in our society, as well as the technical capabilities of factories and the success or otherwise of recent waves of investment. It is clearly possible for the rate

of profit to be ‘too high’ or ‘too low,’ but conventional economics is of no use in establishing either level.

Ignorance is bliss

Of course, the average economist would never tell you that economic theory had suffered such a devastating blow. This is because the average young economist doesn’t even know that this intellectual bout took place – the concepts in this debate don’t make it onto the curriculum for either undergraduate or postgraduate students. Older economists cannot avoid some knowledge of the war, but they either erroneously believe that their camp won, or they dismiss the issue completely.

Today, economic theory continues to use exactly the same concepts which Sraffa’s critique showed to be completely invalid – capital as an amorphous mass that can be costlessly moved from producing any commodity to any other, whose return reflects its marginal productivity, and which can be aggregated by adding up its price times quantity.

There are few better signs of the intellectual bankruptcy of economics than this.

However, this madness is often justified by an appeal to a methodological precept that the absurdity of a theory’s assumptions is irrelevant – all that matters is that the theory’s predictions accord with reality. We now turn to consider this popular but false defense of economics.

8 | THERE IS MADNESS IN THEIR METHOD

Why assumptions do matter, and why economics is so different from the true sciences

Economics would have us believe that it is a science, fully able to stand tall beside the more conventional physical sciences and mathematics.

After the preceding chapters, you should be inclined to reject that belief. Surely, whatever 'science' is, one might hope that it is undertaken with more impartiality, regard for the facts and logical consistency than economics has displayed.

However, the critiques of conventional economics which form the substance of this book were devised by critical economists (and sometimes, inadvertently, by conventional economists themselves) and some of these critiques have been acknowledged as valid by some conventional economists. There is also a small but robust minority working on other approaches to economic analysis, as you'll find in Chapter 18. There are thus some systematic and logical aspects to what economists in general do, which could qualify as scientific behavior.

The position I now favor is that economics is a pre-science, rather like astronomy before Copernicus, Brahe and Galileo. I still hold out hope of better behavior in the future, but given the travesties of logic and anti-empiricism that have been committed in its name, it would be an insult to the other sciences to give economics even a tentative membership of that field.[1]

Before better behavior can take widespread root, economics will have to wean itself from a methodological myth. This is the proposition, first put by Milton Friedman, that a theory cannot be judged by its assumptions, but only by the accuracy of its predictions.

Leaving aside the question of whether economics has ever accurately predicted anything, the argument that 'the more significant the theory, the more unrealistic [are] the assumptions' is simply bad philosophy.

The kernel

Have you heard the joke about the chemist, the physicist and the economist who get wrecked on a desert isle, with a huge supply of canned baked beans as their only food? The chemist says that he can start a fire using the neighboring palm trees, and calculate the temperature at which a can will explode. The physicist says that she can work out the trajectory of each of the baked beans, so that they can be collected and eaten. The economist says, 'Hang on, guys, you're doing it the hard way. Let's assume we have a can opener.'[2]

That assumption is not too different from the type of assumption that economists routinely make, and yet they defend themselves on the apparently convincing grounds that the assumptions don't matter – a theory can be evaluated only on the basis of the accuracy of its predictions.

This methodological defense is invalid, because it confuses 'negligibility' assumptions, which argue that some minor details can be ignored, with 'domain' assumptions, which determine the range of applicability of a given theory. Assumptions also do matter to economists, in that they genuinely believe that their theories describe reality, and they reject economic argument that is not based upon their preferred set of assumptions.

The roadmap

In this chapter I outline the paper in which Friedman introduced the notion that 'assumptions don't matter.' Following Musgrave, I classify assumptions under three headings: negligibility assumptions, domain assumptions, and heuristic assumptions. Friedman's paradoxical statement that 'the more significant the theory, the more unrealistic the assumptions' is only partially true of the first class of assumptions, and manifestly untrue of the latter two classes. Finally, I detail the many ways in which assumptions do matter to economists.

A paradoxical proposition

There would be few if any academic economists who have not had a lecture disturbed by some recalcitrant student, interjecting that the assumptions of the model being discussed are unrealistic. Fortunately, there is a simple weapon at hand: an appeal to the authority of Milton Friedman that a theory can't be judged by its assumptions, but only by how well its predictions accord with reality.

In fact, Friedman's case went farther: he argued that unrealistic assumptions were the hallmark of good theory. In what Paul Samuelson later dubbed 'the F-twist,' Friedman argued that

> Truly important and significant hypotheses will be found to have 'assumptions' that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions (in this sense). The reason is simple. A hypothesis is important if it 'explains' much by little, that is, if it abstracts the common and crucial elements from the mass of complex and detailed circumstances surrounding the phenomena to be explained and permits valid predictions on the basis of them alone. To be important, therefore, a hypothesis must be descriptively false in its assumptions; it takes account of, and accounts for, none of the many other attendant circumstances, since its very success shows them to be irrelevant for the phenomena to be explained.
>
> To put this point less paradoxically, the relevant question to ask about the 'assumptions' of a theory is not whether they are descriptively 'realistic,' for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions. (Friedman 1953)

The proposition that a theory is not regarded as a description of reality, but merely as a way of predicting the future, is known as 'instrumentalism.' This position is superficially appealing, and sufficiently persuasive to

quieten the average interjector. It appears scientific, in that most scientists would admit that their theories can never exactly describe reality. It also implies a healthy dose of theoretical agnosticism, in that the economist is purportedly detached from his theory, and is only really interested in 'the facts.'

However, despite its superficial appeal, instrumentalism suffers from several flaws, which were clearly set out by the philosopher Alan Musgrave in 1981. Musgrave argued that there were three classes of assumptions, and that Friedman's dictum was only partially true in the least important of them.

Negligibility assumptions Negligibility assumptions state that some aspect of reality has little or no effect on the phenomenon under investigation. Friedman's paper made heavy use of the example of a ball being dropped near the earth, which fell very nearly 'as if' it had been dropped in a vacuum. In this instance it was valid to assume that the ball was falling in a vacuum, since air resistance has negligible impact on the ball's fall. However, the same was obviously not true of a feather dropped under the same circumstances.

Friedman argued that though it was unrealistic to say 'assume the ball was dropped in a vacuum,' the theory of gravity had great explanatory power: it explained much (the acceleration of bodies in free fall close to the earth) with very little (a gravitational constant and simple calculus). This theory should be dropped in favor of another only if a rival is at least as accurate and equally acceptable on other grounds, or 'when there exists a theory that is known to yield better predictions but only at a greater cost' (Friedman 1953).

Musgrave argued that many of Friedman's musings were reasonable in this domain, but that even here his 'dialectical' proposition that 'the more significant the theory, the more unrealistic the assumptions' is overblown. In fact, it is possible to rephrase these 'unrealistic' statements as 'realistic' ones: for example, it is realistic to say that air resistance is negligible for dense bodies falling from rest over short distances. As Musgrave put it, these assumptions:

> are not necessarily 'descriptively false,' for they do not assert that present factors are absent but rather that they are 'irrelevant for the phenomena to be explained' […] Galileo's assumption that air-resistance was negligible for the phenomena he investigated was a true statement about reality, and an important part of the explanation Galileo gave of those phenomena. (Musgrave 1981)

However, negligibility assumptions are the minnows of the assumptions family. Far more important are domain assumptions, and it is these to which rightly troubled students often object.

Domain assumptions A domain assumption specifies the conditions under which a particular theory will apply. If those conditions do not apply, then neither does the theory.

An economic example of this is the assumption that risk can be used as a proxy for uncertainty – an assumption that permeates the conventional theories of macroeconomics and finance, which we will investigate in Chapters 10 and 11.

Risk applies to situations in which the regularity of past events is a reliable guide to the course of future events. Gambling gives us many such examples: if a tossed coin is seen to land showing heads roughly half the time, then you can reliably bet that there will be a 50:50 chance of heads in the future. If anyone bet you that heads would in future come up only 40 percent of the time, it would be sensible to take the bet. A risky event will have a probability associated with it, and a variance of outcomes around those probabilities, which can be reliably estimated using the techniques of statistics.

Uncertainty applies when the past provides no reliable guide to future events. Though the fact that we cannot predict the future is the essence of the human condition, the very nebulousness of uncertainty means that many people – and certainly the vast majority of economists – have difficulty grasping the concept. As a result, they act as if the quantifiable concept of risk can be safely substituted for unquantifiable uncertainty.

A somewhat intimate example might illustrate the fallacy of identifying uncertainty with risk.[3] Imagine that you are very attracted to a particular individual, and that you know this person has gone out with 20 percent of

those who have asked him or her out in the past. Does this mean that you have a 20 percent chance of being lucky if you 'pop the question'?

Of course not. Each instance of attraction between two people is a unique event, and the past behavior of the object of your desires provides no guide as to how your advances will be received. How he or she will react cannot be reduced to some statistical prediction based on past apparent regularities. From your perspective, their reaction is truly uncertain – and this uncertainty is at the root of much of the angst that romantic attraction generates.

A similar observation can be made about each new business investment. Even if similar investments have been made in the past, the economic environment of a new investment differs from those which have gone before. Past trends therefore cannot be confidently extrapolated to predict future performance – but this procedure is the essential assumption behind using statistics to calculate risk.

The assumption that risk can be used as a proxy for uncertainty when evaluating investments is therefore unrealistic. A theory that makes such an assumption is quite clearly not better than an alternative one which does not – quite the opposite in fact. This assumption says that the domain of relevance of the theory is a world in which the future is simply subject to chance.

Since there is no such world, the domain of applicability of theories which make such an unrealistic assumption is 'nowhere.' Yet assumptions of this type abound in economic theory (especially, it must be said, in the work of Milton Friedman).

Such an assumption should be made only if it fits into Musgrave's third class, the heuristic assumption.

Heuristic assumptions A heuristic assumption is one which is known to be false, but which is made as a first step towards a more general theory. Musgrave gives the example of Newton's assumption that the solar system consisted only of the sun and the earth. This gave rise to the theory that planets would follow elliptical orbits (which is a reasonable medium-term guide to actual planetary orbits in our solar system).

The next major step came with Poincaré in 1899, when he tried to develop a formula describing planetary motion in a system with more than

one planet. His proof that there was no such formula – and that the actual orbits would interact in wildly unpredictable ways – ushered in what is now known as 'chaos theory' or 'complexity theory' (though it lay dormant for sixty-eight years until modern computers allowed its accidental rediscovery).

The modern theory of planetary behavior now recognizes that the stable orbits of our solar system can only have evolved – over an enormous period of time – from far less stable orbits, which must have led to collisions between proto-planets. It is now accepted that the moon, for example, was the product of a collision between another proto-planet and the early earth.

Collisions are not possible in a single-planet solar system – the kind of system that Newton assumed to derive his initial theory. Though that heuristic assumption was a major step in the development of the scientific mode of thinking about astronomy, dropping it led to a better theory, not a worse one.

When heuristic assumptions are made consciously by a theorist in the course of developing a theory, they are normally explicitly described as such. For instance, when developing the theory of relativity, Einstein at one point stated that the distance covered by a person walking from one side to the other of a moving train is equal to the sum of the distance covered by the train, and the width of the carriage. However, he continued that 'We shall see later that this result cannot be maintained; in other words, the law that we have just written down does not hold in reality. For the time being, however, we shall assume its correctness' (Einstein 1961 [1916]). When Einstein dropped this heuristic assumption, the theory of relativity was the result.

The greater realism at the heart of Einstein's theory transformed our understanding of reality, and dramatically expanded the physical and intellectual capabilities of our species. Yet if we accept Friedman's methodology, then we would have to argue that Einstein's theory was poorer than Newton's because it was more realistic.

In general, then, and contrary to Friedman, abandoning a factually false heuristic assumption will normally lead to a better theory – not a worse one.

Judging the assumptions Theories can therefore be evaluated by their assumptions to some extent, if one has an intelligent taxonomy of

assumptions. A theory may well draw power from 'unrealistic' assumptions if those assumptions assert, rightly, that some factors are unimportant in determining the phenomena under investigation. But it will be hobbled if those assumptions specify the domain of the theory, and real-world phenomena are outside that domain.

These assumptions may be justified if they are merely heuristic devices used to simplify the process of deriving a more general theory – but only if that more general theory is in fact derived. Economists often imply, when they fob off some critical student, that the unrealistic assumptions in introductory economics courses are dropped in more advanced theory – which portrays these assumptions as heuristic tools. In fact, as preceding chapters have illustrated, the assumptions used in more advanced theory are often more unrealistic than those presented in introductory lectures.

Scientific realism versus instrumentalism Musgrave also points out that most scientists reject an instrumental view of science in favor of 'scientific realism' – the belief that scientific theories should not merely predict reality but should, in some sense, represent it.

Ironically, this is actually the belief that most economists have about economic theory. Friedman's instrumentalism is little more than a smokescreen behind which to hide when one wishes to quell a budding class rebellion. It is often evident to the student objector that, though professing that the assumptions don't matter, his teachers continue to use the same small class of assumptions over and over again: rational utility-maximizing individuals, profit-maximizing firms, and a plethora of ancillary assumptions built on these foundations.

These assumptions are used because economists believe that these assumptions do capture essential elements of reality, and regard any theory which does not use these building blocks as 'unrealistic.' This belief is most clearly seen in the manner in which the 'bibles' of economics, its academic journals, filter out papers that do not make this core set of assumptions.

Assumptions do matter – to economists The proposition that assumptions don't matter implies that economists would be quite willing to accept a theory which assumed irrational behavior if the model generated results which accorded with observation. It also implies that the development of

economic theory would be driven primarily by the desire to produce theories that provide a closer fit to observed data.

Both these implications are strongly at variance with reality.

As any non-orthodox economist knows, it is almost impossible to have an article accepted into one of the mainstream academic economic journals unless it has the full panoply of economic assumptions: rational behavior (according to the economic definition of rational!), markets that are always in equilibrium, risk as an acceptable proxy for uncertainty, and so on. When it comes to safeguarding the channels of academic advancement, little else matters apart from preserving the set of assumptions that defines economic orthodoxy.

Similarly, the development of economic theory over time has been propelled by the desire to make every aspect of it conform to the preferred economic model. Macroeconomics, when it first began, bore little resemblance to microeconomics. Fifty years later, macroeconomics is effectively a branch of microeconomics. As I outline in Chapter 10, a major factor behind this tribal coup was the belief that, regardless of its predictive validity, macroeconomics was unsound because its assumptions did not accord with those of microeconomics. It was therefore extensively revised, especially during the 1970s and 1980s, so that macroeconomic theory was more consistent with microeconomic assumptions. Far from assumptions not mattering to economists, assumptions in fact drove the development of economic theory.

Assumptions and logic Assumptions matter in a more profound sense because, as this book shows, assumptions can be logically incoherent. For example, as discussed in Chapter 4, the economic model of the firm is internally contradictory. A theory that contains logically inconsistent assumptions will be a bad theory – and, as this book shows, economics is replete with logical inconsistencies.

This is a science? The behavior of economists hardly fits the stereotype of scientists as dispassionate seekers of truth. But their behavior does fit modern, sociological theories of how scientists behave.[4]

Briefly, these theories argue that each 'science' is as much a society as it is an intellectual discipline. A collection of scholars in a science will share a

perspective on what defines their discipline, and what constitutes scientific behavior. This shared mindset includes core beliefs, which cannot be challenged without threatening your membership of the group (and hence your status as a scientist), ancillary beliefs which are somewhat malleable, a set of analytic techniques, and as yet unsolved problems to which these techniques should be applied. The core beliefs are known as the 'hard core' – since they cannot be altered without rejecting, in some crucial sense, the very foundations of the science. The ancillary beliefs are known as the 'protective belt,' since their function is to protect the core beliefs from attack.

The scholars expect that their beliefs and techniques will be able to solve the outstanding problems, thus increasing the explanatory power of their science. If they fail, then the first response is to adjust the ancillary beliefs rather than the core propositions. Only when the problem proves both intractable and crucial is there any possibility that core beliefs will be abandoned, leading to the formation of a new school of thought – or the ascendancy of an existing rival school. While a school of thought is expanding the range of phenomena it can explain using its core beliefs – by experiments that confirm its predictions, or extensions of its theories to novel areas – then it is said to be a 'progressive' scientific research program which manifests a 'positive heuristic.' If, instead, experimental results contradict its predictions, and its theories are adjusted to rationalize these failures, then it is said to be 'degenerative' with a 'negative heuristic.'

It is possible for more than one such collection of scholars to exist in a science at any one time, so it makes sense to speak of schools of thought within a science. Each school of thought will compete with the others, emphasizing their weaknesses and its own strengths.

Clearly this sociological description of a science fits the historical record of economics. At the beginning of the third millennium, there are at least five schools of thought. The neoclassical school is clearly dominant, but there are several other competing schools – in particular, the post-Keynesian, Austrian, and evolutionary schools of economics. Each is developing its own approach to explaining similar phenomena, and there is clearly a rivalry between the minority schools and neoclassical economics – the other schools criticize neoclassical economics while it largely ignores its rivals.

However, it might be thought that this provides a fairly demeaning perspective on science itself. Surely this behavior is aberrant, and true sciences are beyond this petty bickering? No, strange as it may seem, a similar picture can be painted even of the queen of sciences, physics.

Quantum uncertainty? In order to comprehend some of the bizarre results of experimental particle physics, most physicists argue that matter is in some sense 'probabilistic,' and that the observer fundamentally affects reality. If an observer tries to 'tie down' one aspect of a particle – say, its location – then some other aspect becomes fundamentally unknowable. Physicists say that an elementary particle is always in a 'superposition' of both states, and testing for one leads to the other state resolving itself in a completely random way. The act of observing a particle thus directly – but unpredictably – alters its state. This is not because of any statistical properties of large numbers of electrons, but because randomness is an inherent feature of fundamental particles.

Two crucial aspects of this 'Copenhagen school' interpretation of quantum reality are (a) that particles can be treated as 'wave functions' in what is known as the wave–particle duality, so that a fundamental particle can be completely represented by its wave function; and (b) that there are two sets of physical laws, one which applies when there is no observer ('superposition') and one which exists when there is an observer.

The most famous popular representation of what this means, when put in terms of everyday objects, is 'Schrodinger's cat.' This is a thought experiment in which a box contains a cat, a radioactive element, and a vial of poison. If the radioactive element emits a particle, the vial opens and the cat dies. If it doesn't, the cat lives.

What state is the cat in before an experimenter opens the lid to see whether it is alive or dead? In the Copenhagen school interpretation, the cat is in a superposition of being both alive and dead. The act of the observer opening the box resolves the cat into one or other state.

But this is not the only way to make sense of the experimental data. A rival interpretation, established by David Bohm, provides a completely deterministic interpretation, with none of the 'quantum uncertainty' of the Copenhagen school. It can explain the same experimental results as can the Copenhagen school – and some which it can't explain – without resorting to

the apparently metaphysical position that the observer somehow affects reality at the quantum level. In Bohm's theory, Schrodinger's cat is either alive and well if the radioactive element hasn't emitted a particle, or dead if it has, independent of the human observer who eventually opens the box to check.

How have physicists reacted to this coexistence of two rival explanations of reality? As the physicist David Albert sees it, in much the same way that economists have reacted to alternative schools of thought – by refusing to take them seriously. It is worth citing Albert at some length to show that, quite possibly, scientists in other disciplines are no different from economists when it comes to their reaction to intellectual challenges to accepted dogma:

> Despite all the rather spectacular advantages of Bohm's theory, an almost universal refusal even to consider it, and an almost universal allegiance to the standard formulation of quantum mechanics, has persisted in physics, astonishingly, throughout most of the past 40 years. Many researchers have perennially dismissed Bohm's theory on the grounds that it granted a privileged mathematical role to particles. The complaint was that this assignment would ruin the symmetry between position and momentum, as if ruining that symmetry amounted to a more serious affront to scientific reason than the radical undermining, in the Copenhagen formulation, of the very idea of an objective reality. Others dismissed Bohm's theory because it made no empirical predictions (no obvious ones, that is) that differed from those of the standard interpretation – as if the fact that those two formulations had much in common on that score somehow transparently favored one of them over the other. Still others cited 'proofs' in the literature that no deterministic replacement for quantum mechanics of the kind that Bohm had already accomplished was even possible. (Albert 1994)

After the above was published in the first edition, several physicists contacted me and put forward criticisms of Bohm's theory. However, the relevance of his theory in the context of this chapter was the alleged

behavior of physicists in rejecting this alternative perspective in the manner described by Albert.

At this sociological level, therefore, economics appears to have some similarities to the conventional sciences – though the extent to which alternative perspectives are suppressed in economics is far greater than in physics.

A degenerate scientific research program There was a time when the neoclassical school of economics was clearly progressive, while its main rival was clearly degenerate. When the neoclassical school coalesced in the 1870s in the works of Jevons, Menger and Walras, the preceding classical school was in crisis. The classical school always had a difficulty in explaining the relationship between what it called value and prices; yet it insisted that value was in some way fundamental to the determination of price. This problem was accentuated by the work of the final member of the classical school, Karl Marx (the subject of Chapter 17).

At the same time, the neoclassical school was expanding its core belief that human behavior was driven by the desire to maximize utility. This had developed from a guiding principle, in Bentham's hands, to a coherent theory of consumer and producer behavior in the hands of Jevons, and to an explanation for the overall coordination of a market economy in Walras. At the turn of the nineteenth century, neoclassical economists were confident that their science could continue expanding its explanation of the economy. It was clearly then a progressive scientific research program.

Though the majority of economists still believe that this is the case today, there are manifest signs that this is no longer true. Instead, the theory today is degenerate: rather than expanding the range of phenomena it can explain, the leading edge of the theory is dominated by adjusting the protective belt of ancillary beliefs to defend the hard-core beliefs from attack. For example, the Sonnenschein-Mantel-Debreu conditions (discussed in Chapter 3) are a way of maintaining the hard-core belief that individual behavior is driven by utility maximization, despite the proof that individual preferences cannot be aggregated. A similar interpretation could be given of responses of neoclassical economics to the many logical problems documented in this book.

But the problems with economics go beyond just this, since if economics were as fully a science as astronomy, eventually its litany of failures would lead to at least a general acknowledgment of crisis.

The incredible inertness of economics What makes economics different from and inferior to other sciences is the irrational tenacity with which it holds to its core beliefs in the face of either contrary factual evidence or theoretical critiques that establish fundamental inconsistencies in its intellectual apparatus.

The discovery, for example, that firms believe they experience constant or falling marginal costs (Eiteman and Guthrie 1952), and generally set prices by placing a markup on average cost, led not to the abandonment of the economic theory of price-setting, but to a welter of papers arguing that in a competitive market, the effect of markup pricing was the same as if firms did consciously equate marginal cost to marginal revenue (Langlois 1989). On the same note, Sraffa's theoretical argument that diminishing marginal returns were unlikely to occur in practice was ignored.

As a result, students at the beginning of the twenty-first century are receiving much the same instruction about how firms set prices as did their counterparts at the end of the nineteenth century.

Physical sciences hold on to their core beliefs with some tenacity, but nowhere near this much – even Albert's paper goes on to observe that 'serious students of the foundations of quantum mechanics rarely defend the standard formulation anymore' (Albert 1994). As a result, revolutions in physical sciences – where one dominant paradigm is replaced by another – occur much more frequently than they do in economics. Often, these revolutions outpace the popular understanding of a science.

Astronomy provides an example of this. I expect that most lay people think that the dominant theory of how the universe came into being is the 'Big Bang.' In this theory, the universe originated in a 'quantum singularity' some 12–15 billion years ago. This explosion kick-started matter and time, leading to the immense universe we observe today. Back in the 1950s, this theory won out against its rival, that the universe had always been in a 'steady state' of expansion.

The Big Bang was indeed the dominant theory for some time – until it was pointed out that, according to calculations from quantum mechanics,

the Big Bang would have resulted in a universe consisting of a mere handful of elementary particles.

A rival theory then developed which argued that, for a substantial period of time, the laws of physics of the current universe did not apply. Matter, for example, could move much faster than the speed of light. This 'inflationary universe' theory has subsequently been embellished to predict that there are many universes – as opposed to the one universe postulated by the Big Bang.

The shifts from the Big Bang paradigm to the inflationary universe, to 'multiverses,' are big ones conceptually. The first envisages a single finite universe, while the last muses that ours may be only one of many universes, each with different 'fundamental' physical laws. But the science of astronomy made this move over a period of about twenty years, and it continues to undergo development today. Now even the inflationary/multiverse theory is under challenge, as measurements imply that the rate of expansion of the universe is actually increasing with time.[5]

Economics, in contrast, has had only one acknowledged revolutionary episode in the last century – the Keynesian revolution during the 1930s. Yet at the end of the twentieth century, the dominant school of thought in economics retains nothing from that revolution, and is in fact a direct descendant of pre-Keynesian neoclassical economics.

Think of the many revolutions in our understanding of the physical world which have occurred in the twentieth century: from Newtonian to Einsteinian physics; from Mendelian genetics to DNA and the human genome; from determinism to chaos theory. Any scientist from the nineteenth century would be bewildered by what is commonplace today in his discipline – save an economist.

Why is economics so resistant to change? Is it because everything economists believed at the end of the nineteenth century was correct? Hardly, as this book shows. Instead, to understand the incredible inertness of economics, we have to consider an essential difference between social sciences in general and the physical sciences, and the thorny topic of ideology.

My kingdom for an experiment In the nineteenth century, scientists and philosophers of science generally believed that what distinguished the

social sciences from the physical sciences was that the latter could undertake experiments to test their theories, whereas the former could not. In the twentieth century, Popper instead argued that the distinction between a science – like physics – and a non-science – like astrology – was not that one could undertake experiments and the other could not, but that one made falsifiable statements, while the other did not. Popper's distinction between science and non-science wasn't completely relevant to the 'experiments versus no experiments' distinction, but it did tend to play down the importance of experimentation in deciding what was and what was not a science.

The history of economics implies that Popper's distinction does not give sufficient attention to whether or not a falsifiable statement can in fact be experimentally falsified. For example, Milton Friedman is famous as the father of the now defunct sub-branch of economics known as monetarism. One falsifiable statement he made was that inflation is caused by the government increasing the money supply more rapidly than the economy is going.

This implied that, to reduce inflation, all the government had to do was to increase the money supply more slowly than the economy was growing. This was the basis of the economic policies of Margaret Thatcher, yet eventually this approach was abandoned. One reason why was that the government was never able to meet its targets for the rate of growth of the money supply – it might aim to increase it by, say, 6 percent, only to see it grow by 11 percent. Also, the relationship between the three crucial variables in Friedman's theory – the rate of inflation, the rate of growth of the economy, and the rate of growth of the money supply – was never as watertight in practice as it appeared to be in his theory.

You could thus argue that Friedman's statement – that inflation is caused by the government expanding the money supply faster than the rate of growth of the economy – had been falsified. Did this lead Milton and his supporters to abandon his theory? Of course not: monetarists instead argued that all sorts of attenuating features disturbed the results.

In other words, because the monetarist experiment in Great Britain wasn't a controlled experiment, monetarist economists could refuse to accept that their theory had been falsified.

The same observation can be made about Marxist economists, and their attitude toward the data on Marx's theory that the rate of profit would tend to fall, or the inevitability of socialism, and so on. In other words, this isn't just a disease of the political right, but an endemic problem in economics: without the ability to undertake controlled experiments, statements which could be falsified will be unfalsifiable in practice. Economists of all persuasions are therefore liable to hang on to beliefs that they argue are scientific, but which in the end are ideological.

The experience of another social science, psychology, provides some support for the argument that the ability to undertake experiments is crucial to scientific progress. For much of the twentieth century, psychology was dominated by the 'behaviorist' school. This school argued that an organism's behavior had to be understood as a response to an external stimulus: it was 'unscientific' to postulate any unobservable mental processes of the organism which mediated between the stimulus and the response. To this school, complex behavior – such as playing a piano – had to be understood as a chain of stimuli and responses. However, experiments showed that

> even average pianists move their hands too quickly for the tactile information to pass along the sensory nerves to the central nervous system and for the command to move the hands to be sent down the motor nerves […] Therefore, the behaviorist hypothesis that each new action is a response to an external stimulus is implausible. (Bond 2000)

This and several other experimental falsifications of behaviorism led to its demise, and replacement by cognitive psychology, which accepts that 'there are cognitive processes that determine our behavior which we, as psychologists, must explain, even if they are not directly observable' (ibid.). Thus psychology, with the help of experiments, was able to undergo a revolution from one dominant school to another – while economics continues to be dominated by the same school (which, ironically, has a very behaviorist view of human behavior). Unless it develops a means to undertake experiments to test rival theories, economics may be unable to break from the grip of ideology.

Equilibrium and an invisible ideology Economics as a discipline arose at a time when English society was in the final stages of removing the controls of the feudal system from its mercantile/capitalist economy. In this climate, economic theory had a definite (and beneficial) political role: it provided a counter to the religious ideology that once supported the feudal order, and which still influenced how people thought about society. In the feudal system the preordained hierarchy of king, lord, servant and serf was justified on the basis of the 'divine right of kings.' The king was God's representative on earth, and the social structure which flowed down from him was a reflection of God's wishes.

This structure was nothing if not ordered, but this order imposed severe restrictions on the now dominant classes of merchants and industrialists. At virtually every step, merchants were met with government controls and tariffs. When they railed against these imposts, the reply came back that they were needed to ensure social order.

Economic theory – then rightly called political economy – provided the merchants with a crucial ideological rejoinder. A system of government was not needed to ensure order: instead, social order would arise naturally in a market system in which each individual followed his own self-interest. Smith's phrase 'the invisible hand' came along rather late in the process, but the notion played a key role in the political and social transformations of the late eighteenth and early nineteenth centuries.

An essential aspect of this market social order was equilibrium.

From the outset, economists presumed that the market system would achieve equilibrium. Indeed, the achievement of equilibrium was often touted as an advantage of the free market over any system where prices were set by fiat. Equilibrium was therefore an essential notion of the economic defense of capitalism: the equilibrium of the capitalist market would replace the legislative order of the now defunct feudal hierarchy.

More importantly, whereas the feudal order endowed only the well born with welfare, the equilibrium of the market would guarantee the best possible welfare for all members of society. The level of individual welfare would reflect the individual's contribution to society: people would enjoy the lifestyle they deserved, rather than the lifestyle into which they had been born.

If, instead of equilibrium, economists had promised that capitalism would deliver chaos; if, instead of meritocracy, economists had said that the market would concentrate inequality, then economists could have hindered rather than helped the transition to capitalism (though they more likely would have been ignored).

By the middle of the nineteenth century, the transition to capitalism was complete: what was left of feudalism was a mere vestige. But rather than the promised equilibrium, nineteenth-century capitalism was racked by cycles and enormous disparities of wealth. A major depression occurred roughly every twenty years, workers' conditions would improve and then rapidly deteriorate, prices rise and then fall, banks expand and then collapse. New 'robber barons' replaced the barons of old. It appeared that, while promising a meritocratic equilibrium, capitalism had instead delivered unbalanced chaos. A new political challenge arose: that of socialism.

Once again, economics rose to the challenge, and once again equilibrium was a central tenet. This time the defense was mounted by what we today call neoclassical economics, since classical economics had been turned into a weapon against capitalism by the last great classical economist, Karl Marx.

In contrast to the hand-waving of Smith, the neoclassical economists of the late nineteenth century provided a substantive mathematical analysis of how equilibrium could be achieved by an idealized market economy, and how this equilibrium could be fair to all. However, unlike the earlier classical championing of capitalism, this technical edifice provided very little in the way of libertarian slogans for the battle against the ideology of socialism. Instead of arming capitalism's defenders with rhetoric to deploy against socialists, it gave birth to the academic discipline of economics.

Capitalism eventually transcended the challenge of socialism, with little real assistance from economic theory. But while the economics had little impact upon capitalism, the need to defend capitalism had a profound impact upon the nature of economic theory. The defensive imperative, and the role of equilibrium in that defense, cemented equilibrium's role as a core belief of economic theory.

At the beginning of the third millennium, there is no competing social system against which capitalism must prove its superiority. Feudalism is

long dead, and those socialist societies which remain are either socialist in name only, or bit players on the world stage.

Today, most economists imperiously dismiss the notion that ideology plays any part in their thinking. The profession has in fact devised the term 'positive economics' to signify economic theory without any value judgments, while describing economics with value judgments as 'normative economics' – and the positive is exalted far above the normative.

Yet ideology innately lurks within 'positive economics' in the form of the core belief in equilibrium.[6] As previous chapters have shown, economic theory has contorted itself to ensure that it reaches the conclusion that a market economy will achieve equilibrium.[7] The defense of this core belief is what has made economics so resistant to change, since virtually every challenge to economic theory has called upon it to abandon the concept of equilibrium. It has refused to do so, and thus each challenge – Sraffa's critique, the calamity of the Great Depression, Keynes's challenge, the modern science of complexity – has been repulsed, ignored, or belittled.

This core belief explains why economists tend to be extreme conservatives on major policy debates, while simultaneously believing that they are non-ideological, and motivated by knowledge rather than bias.

If you believe that a free market system will naturally tend towards equilibrium – and also that equilibrium embodies the highest possible welfare for the highest number – then, *ipso facto*, any system other than a complete free market will produce disequilibrium and reduce welfare. You will therefore oppose minimum wage legislation and social security payments – because they will lead to disequilibrium in the labor market. You will oppose price controls – because they will cause disequilibrium in product markets. You will argue for private provision of services – such as education, health, welfare, perhaps even police – because governments, untrammeled by the discipline of supply and demand, will either under- or oversupply the market (and charge too much or too little for the service).

In fact, the only policies you will support are ones that make the real world conform more closely to your economic model. Thus you may support anti-monopoly laws – because your theory tells you that monopolies are bad. You may support anti-union laws, because your theory asserts that collective bargaining will distort labor market outcomes.

And you will do all this without being ideological.

Really?

Yes, really – in that most economists genuinely believe that their policy positions are informed by scientific knowledge, rather than by personal bias or religious-style dogma. Economists are truly sincere in their belief that their policy recommendations will make the world a better place for everyone in it – so sincere, in fact, that they often act against their own self-interest.

For example, there is little doubt that an effective academic union could increase the wages paid to academic economists. If economists were truly self-motivated – if they behaved like the entirely self-interested rational economic man of their models – they would do well to support academic unions, since the negative impacts they predict unions to have would fall on other individuals (fee-paying students and unemployed academics). But instead, one often finds that economists are the least unionized of academics, and they frequently argue against actions that, according to their theories, could conceivably benefit the minority of academics at the expense of the greater community. However ideological economists may appear to their critics, in their hearts they are sincerely non-partisan – and, ironically, altruistic.

But non-partisan in self-belief does not mean non-partisan in reality. With equilibrium both encapsulating and obscuring so many ideological issues in economics, the slavish devotion to the concept forces economists into politically reactionary and intellectually contradictory positions.

Of course, if economists were right that equilibrium embodies the best possible outcome for the greatest number, then their apparently ideological policy positions would be justified – if the economy always headed back to equilibrium when disturbed from its nirvana. In the next chapter, we'll put aside the critiques which establish that the building blocks of equilibrium are invalid, and instead ask whether economic equilibrium, as defined by economic theory, is in fact stable.

9 | LET'S DO THE TIME WARP AGAIN

Why economics must finally treat time seriously

Forget everything you know about riding a bicycle, and imagine that someone who purports to be a 'bicycle guru' has convinced you that there are two steps to learning how to ride a bike. In step 1, you master balancing on a stationary bike. In step 2, you master riding a moving bike, applying the skills acquired at step 1.

After several difficult months at step 1, you would know that to remain upright, you must keep your center of gravity directly above the point of contact between your wheels and the road.

Step 2 arrives. Applying the lessons in stage 1, you keep your bike at a perfect 90 degrees to the ground, balance against the uneven pressure of your legs, get up some speed and you're away.

So far, so good. But what if you want to change direction? The handlebars appear to provide the only means of turning, so you rotate them in the direction you wish to go – and fall flat on your face.

What went wrong with this apparently logical advice? 'Elementary, my dear Watson': the gyroscopic force which keeps you upright when a bike is moving simply doesn't exist when it is stationary. Manipulating this force is what enables you to turn a moving bike, and the lessons learnt in the static art of balancing a stationary bike are irrelevant to the dynamic art of actually riding one.[1]

Replace the bicycle with the economy, and the point still stands: the procedures which apply in a static economy are irrelevant to a dynamic, changing one; the forces which apply in a static economy simply don't exist in a dynamic one.[2] Lessons learnt from managing an economy in which processes of change either don't occur, or in which changes occur instantly,

are irrelevant to an economy in which change does occur, and takes time to occur.

The kernel

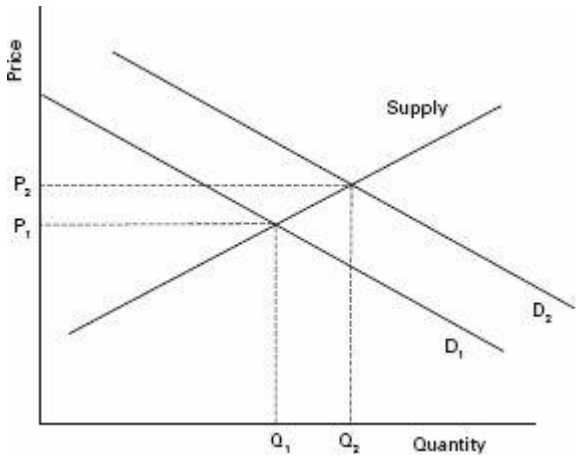
Neoclassical economic models in general ignore processes which take time to occur, and instead assume that everything occurs in equilibrium. For this to be allowable, the equilibrium of the dynamic processes of a market economy must be stable, yet it has been known for over forty years now that those processes are unstable: that a small divergence from equilibrium will *not* set up forces which return the system to equilibrium. The dynamic path of the economy therefore cannot be ignored, and yet most economists remain almost criminally unaware of the issues involved in analyzing dynamic, time-varying systems.

The roadmap

In this chapter I detail the roots of the economic propensity to ignore time, and to instead focus on what happens in equilibrium. Then I point out that economic research in the 1950s and 1960s established that the equilibrium of a market economy was unstable, so that the economy could never be in equilibrium. A brief discussion of chaos theory outlines the type of analysis which economists should undertake, but do not.

Cobwebs of the mind

Economic processes clearly take time, and yet economists don't consider time in analyzing demand, supply, or any of their other key variables. For example, the quantity demanded of a commodity and the quantity supplied are both treated as functions of price, and the outcome is an equilibrium quantity. To illustrate what they believe will happen if the demand for a commodity rises, neoclassical economists compare one equilibrium with another, using what they call comparative statics. The time path from one equilibrium to another is ignored.



9.1 Standard neoclassical comparative statics

But what if the initial market price happens not to be the equilibrium price? Then demand and supply will be out of balance: if price exceeds the equilibrium, demand will be too low and supply too high. For equilibrium to be restored, this disequilibrium must set off dynamic processes in supply and demand which cause them both to converge on the equilibrium price. This dynamic process of adjustment will obviously take time. However, in general, economists simply assume that, after a disturbance, the market will settle down to equilibrium. They ignore the short-term disequilibrium jostling, in the belief that it is just a short-term sideshow to the long-run main game of achieving equilibrium.

A similar belief permeates even some of the alternative schools of economics. The dynamic process is ignored because it is believed to be a short-term, transitory phenomenon, and attention is focused on the long-term, allegedly enduring phenomenon of equilibrium. As a result, time itself, the change in variables over time, and disequilibrium situations are all ignored. Even econometric programs which attempt to forecast the future value of macroeconomic variables such as output and employment assume that the current levels are equilibrium values, and they predict what the future equilibrium values will be.

Economics has invented numerous intellectual devices to enable itself to ignore time, and focus upon the equilibrium situations rather than consider the processes of change over time in an economy. One of these devices is one to which many budding students of economics initially object: the 'all other things being equal,' or '*ceteris paribus*' assumption

that nothing changes outside the single market being considered. This assumption lies behind the analysis of supply and demand in a single market, which we've already debunked in Chapters 3 to 5.

Such troubled students are reassured that at higher levels of analysis, this 'partial equilibrium' assumption is dropped for the more realistic proposition that all things are interrelated. However, rather than this more general analysis being more realistic, dynamic, and allowing for disequilibrium as well as equilibrium, it is in fact 'general *equilibrium*': a model of how all aspects of an economy can be in equilibrium simultaneously.

Budding economists who object to the assumption of *ceteris paribus* would walk away in disgust if they were immediately told of the assumptions needed to sustain the concept of general equilibrium. However, their fears assuaged by the promise of more realistic notions to come, they continue up the path of economic inculcation. By the time they confront general equilibrium in graduate education, they treat these assumptions and the analysis which goes with them as challenging intellectual puzzles, rather than as the asinine propositions they truly are. Normally, these students work at less rarefied levels of economic theory, and confidently presume that the leading lights of the profession will generalize the assumptions and solve the remaining puzzles.

As is so often the case with neoclassical economics, the leading lights have done their job very well, but they have not delivered the goods expected of them by the troops. Instead, they have proved that, in general, general equilibrium is unattainable. Even economic models will not achieve general equilibrium, let alone the real economies that general equilibrium once purported to model. General equilibrium is at one and the same time the crowning achievement of economic theory and its greatest failure.

General equilibrium

In the late nineteenth century, three economists in different countries independently gave birth to the neoclassical school of thought: Jevons in England, Menger in Austria, and Walras in France. Today, Walras is the most exalted of these, because his model of general equilibrium set the mold by which economics has since been crafted.

Groping towards equilibrium According to neoclassical theory, equilibrium occurs in a particular market when demand at a given price equals supply at the same price. For equilibrium to occur in all markets simultaneously, the price in every market has to be such that demand and supply are equal in all markets. However, a change of price in one market will affect consumer demand in all other markets. This implies that a move towards equilibrium by one market could cause some or all others to move away from equilibrium. Clearly it is possible that this 'dance of many markets' might never settle down to equilibrium.

This will be especially so if trades actually occur at disequilibrium prices – as in practice they must, since who could ever know when one real-world market was in equilibrium, let alone all of them simultaneously? A disequilibrium trade will mean that the people on the winning side of the bargain – sellers if the price is higher than equilibrium – will gain real income at the expense of the losers, compared to the alleged standard of equilibrium. This shift in income distribution will then affect all other markets, making the dance of many markets even more chaotic.

Walras provided a simple initial abstraction to sidestep this dilemma: he assumed that no trades take place until equilibrium is achieved in all markets. Having satisfied himself that, in the absence of trade, the jiggling of prices up and down would eventually converge to equilibrium, he extended the same faith to a system with production and exchange at disequilibrium prices.

Walras envisaged the market as being a huge, and very unusual, auction. The audience for this auction includes all the owners of the goods for sale, who are simultaneously the buyers for all the goods on sale. At a normal auction, the quantity of each commodity offered for sale is fixed. In Walras's auction, the total amount of each commodity is fixed, but sellers will offer anywhere from none to all of this for sale, depending on the price offered. The quantity offered rises as the price rises, and vice versa, with any amount not sold being taken back home by the seller for his/her own consumption (there are no stocks; everything is either sold or consumed by the producer).

The most peculiar features of Walras's auction market are that, rather than selling each commodity one at a time, the 'auctioneer' attempts to sell all goods at once; and rather than treating each commodity independently,

this auctioneer refuses to accept any price for a commodity until supply equals demand for all commodities. In Walras's words:

First, let us imagine a market in which only consumer goods and services are bought and sold […] Once the prices or the ratios of exchange of all these goods and services have been cried at random in terms of one of them selected as numeraire, each party to the exchange will offer at these prices those goods or services of which he thinks he has relatively too much, and he will demand those articles of which he thinks he has relatively too little for his consumption during a certain period of time. The quantities of each thing effectively demanded and offered having been determined in this way, the prices of those things for which the demand exceeds the offer will rise, and the prices of those things of which the offer exceeds the demand will fall. New prices now having been cried, each party to the exchange will offer and demand new quantities. And again prices will rise or fall until the demand and the offer of each good and each service are equal. Then the prices will be current equilibrium prices and exchange will effectively take place. (Walras 1954 [1874])

This is clearly not the way markets work in the real world.[3] Nonetheless, this mythical construct became the way in which economics attempted to model the behavior of real-world markets.

Walras's auctioneer starts the market process by taking an initial stab at prices. These arbitrarily chosen prices are almost certainly not going to equate demand and supply for each and every commodity – instead, for some commodities, demand will exceed supply, while for others supply will exceed demand. The auctioneer then refuses to allow any sale to take place, and instead adjusts prices – increasing the price of those commodities where demand exceeded supply, and decreasing the price where demand was less than supply. This then results in a second set of prices, which are also highly unlikely to balance demand and supply for all commodities; so another round of price adjustments will take place, and another, and another.

Walras called this iterative process of trying to find a set of prices which equates supply to demand for all commodities '*tatonnement*' – which literally translates as 'groping.' He believed that this process would eventually converge to an equilibrium set of prices, where supply and

demand are balanced in all markets (so long as trade at disequilibrium prices can be prevented).

This was not necessarily the case, since adjusting one price so that supply and demand are balanced for one commodity could well push demand and supply farther apart for all other commodities. However, Walras thought that convergence would win out because the direct effects on demand – of increasing the price of a commodity where demand exceeds supply, which directly reduces demand – would outweigh the indirect effects of changes in demand for other commodities. In his words:

> This will appear probable if we remember that the change from p'b to p''b, which reduced the above inequality to an equality, exerted a direct influence that was invariably in the direction of equality at least so far as the demand for (B) was concerned; while the [consequent] changes from p'c to p''c, p'd to p''d, which moved the foregoing inequality farther away from equality, exerted indirect influences, some in the direction of equality and some in the opposite direction, at least so far as the demand for (B) was concerned, so that up to a certain point they cancelled each other out. Hence, the new system of prices (p''b, p''c, p''d) is closer to equilibrium than the old system of prices (p'b, p'c, p'd); and it is only necessary to continue this process along the same lines for the system to move closer and closer to equilibrium. (Ibid.)

'Generalizing' Walras Walras's ruse, of an auctioneer who stopped any trades taking place until such time as demand equaled supply in all markets, was clearly artificial. However, it enabled economists to make use of the well-known and relatively simple techniques for solving simultaneous linear equations.

The alternative was to describe the dynamics of a multi-commodity economy, in which trades could occur at non-equilibrium prices in anywhere from a minimum of two to potentially all markets. At a technical level, modeling non-equilibrium phenomena would have involved nonlinear difference or differential equations. In the nineteenth century, the methodology for them was much less developed than it is now, and they are inherently more difficult to work with than simultaneous linear equations.

Walras's auctioneer was therefore arguably a justifiable abstraction at a time when, as Jevons put it, it would have been 'absurd to attempt the more difficult question when the more easy one is yet so imperfectly within our power' (Jevons 1888: ch. 4, para. 25).

But it suggests an obvious, dynamic, research agenda: why not see what happens when the artifact of no non-equilibrium trades is dispensed with? Why not generalize Walras's general equilibrium by removing the reliance upon the concept of equilibrium itself? Why not generalize Walras by dropping the fiction that everything happens at equilibrium?

This potential path was, for economics, the path not chosen.

Instead, the neoclassical 'Holy Grail' became to formalize Walras's concept of equilibrium: to prove that general equilibrium existed, and that it was the optimum position for society.

Unfortunately, reality had to be seriously distorted to 'prove' that general equilibrium could be attained. But, for the reasons given in Chapter 8, economists would rather sacrifice generality than sacrifice the concept of equilibrium.

The pinnacle of this warping of reality came with the publication in 1959 of Gerard Debreu's *Theory of Value*, which the respected historian of economic thought Mark Blaug has described as 'probably the most arid and pointless book in the entire literature of economics' (Blaug 1998). Yet this 'arid and pointless' tome set the mold for economics for the next forty years – and won for its author the Nobel Prize for economics.

'The formal identity of uncertainty with certainty'

Walras's vision of the market, though highly abstract, had some concept of process to it. Buyers and sellers would haggle, under the guidance of the auctioneer, until an equilibrium set of prices was devised. Exchange would then take place, and those prices would also determine production plans for the next period. There is at least some primitive notion of time in this series of sequential equilibria.

No such claim can be made for Debreu's vision of general equilibrium. In this model, there is only one market – if indeed there is a market at all – at which all commodities are exchanged, for all times from now to eternity. Everyone in this 'market' makes all their sales and purchases for all of time

in one instant. Initially everything from now till eternity is known with certainty, and when uncertainty is introduced, it is swiftly made formally equivalent to certainty. A few choice extracts give a clearer picture of Debreu's total divorce from reality:

> For any economic agent a complete action plan (*made now for the whole future*), or more briefly an action, is a specification for each commodity of the quantity that he will make available or that will be made available to him, i.e., a complete listing of the quantities of his inputs and of his outputs.
>
> For a producer, say the *jth* one, a production plan (made now for the whole future) is a specification of the quantities of all his inputs and all his outputs. *The certainty assumption implies that he knows now what input-output combinations will be possible in the future* (although he may not know the details of technical processes which will make them possible).
>
> As in the case of a producer, the role of a consumer is to choose a complete consumption plan. His role is to choose (and carry out) a consumption plan made now for the whole future, i.e., a specification of the quantities of all his inputs and all his outputs.
>
> The analysis is extended in this chapter to the case where uncertain events determine the consumption sets, the production sets, and the resources of the economy. A contract for the transfer of a commodity now specifies, in addition to its physical properties, its location and its date, an event on the occurrence of which the transfer is conditional. *This new definition of a commodity allows one to obtain a theory of uncertainty free from any probability concept and formally identical with the theory of certainty developed in the preceding chapters*. (Debreu 1959; emphases added)

I can provide no better judgment of the impact this brazenly irrelevant theory had on economics than that given by Blaug:

> Unfortunately this paper soon became a model of what economists ought to aim for as modern scientists. In the process, few readers realized that Arrow and Debreu had in fact abandoned the vision that had originally motivated Walras. For Walras, general equilibrium theory was an abstract but nevertheless realistic description of the functioning of a capitalist economy. He was therefore more concerned to show that markets will clear automatically via price adjustments in response to positive or negative excess demand – a property that he labeled 'tatonnement' – than to prove that a unique set of prices and quantities is capable of clearing all markets simultaneously.
>
> By the time we got to Arrow and Debreu, however, general equilibrium theory had ceased to make any descriptive claim about actual economic systems and had become a purely formal apparatus about a quasi economy. It had become a perfect example of what Ronald Coase has called 'blackboard economics,' a model that can be written down on blackboards using economic terms like 'prices,' 'quantities,' 'factors of production,' and so on, but that nevertheless is clearly and even scandalously unrepresentative of any recognizable economic system. (Blaug 1998)

A hobbled general It is almost superfluous to describe the core assumptions of Debreu's model as unrealistic: a single point in time at which all production and exchange for all time is determined; a set of commodities – including those which will be invented and produced in the distant future – which is known to all consumers; producers who know all the inputs that will ever be needed to produce their commodities; even a vision of 'uncertainty' in which the possible states of the future are already known, so that certainty and uncertainty are formally identical. Yet even with these breathtaking dismissals of essential elements of the real world, Debreu's model was rapidly shown to need additional restrictive assumptions – the Sonnenschein-Mantel-Debreu conditions discussed in Chapter 3. Rather than consumers being able to have any utility function consistent with what

economists decreed as rational, additional restrictions had to be imposed which, as one economist observed, came 'very close to simply assuming that the consumers in the aggregate have identical tastes and income' (Diewert 1977: 361).

This was not the end of the restrictions. As Blaug observes above, Walras hoped to show that the process of *tatonnement* would lead, eventually, to equilibrium being achieved, and that the same outcome would follow even if disequilibrium trading occurred. In mathematical terms, he hoped to show that general equilibrium was stable: that if the system diverged from equilibrium, it would return to it, and that if the process of *tatonnement* began with disequilibrium prices, it would eventually converge on the equilibrium prices. Debreu abandoned this aspect of Walras's endeavor, and focused solely on proving the existence of general equilibrium, rather than its stability. But stability cannot be ignored, and mathematicians have shown that, under fairly general conditions, general equilibrium is unstable.

Positive prices and negative stability Walras's assumption that the direct effects of the price change would outweigh the indirect effects – so that the process of *tatonnement* would converge on the set of equilibrium prices – was reasonable, given the state of mathematics at the time. However, mathematical theorems worked out in the twentieth century established that, in general, this assumption is wrong.

These theorems established that the conditions which ensure that an economy can experience stable growth simultaneously guarantee that Walras's *tatonnement* process is unstable (Blatt 1983). Therefore if the auctioneer's first stab at prices is only a tiny bit different from the set of prices which would put all markets in equilibrium, his next stab – derived by increasing prices for goods where demand exceeded supply, and vice versa – will be farther away from the equilibrium set of prices. The process of *tatonnement* will never converge to the equilibrium set of prices, so if equilibrium is a prerequisite for trade, trade will never take place.

These theorems[4] are too complex to be conveyed accurately by either words or figures, but in keeping with the objectives of this book, I'll attempt an explanation. If you don't want to twist your mind around the

mathematical concepts involved, then please skip to the following heading ('A transitional methodology?').

The 'general equilibrium problem' is to find a set of prices which result in the amount consumers demand of each and every product equaling the amount supplied. Prices obviously have to be positive, as do the quantities demanded and the quantities produced.[5]

Before commodities can be demanded, they must be produced, and the means of production are simply other commodities. If the economy is going to last indefinitely, the system of production must be able to generate growth.

This can be described by a set of equations in which the prices are the variables, and the quantities required to produce each commodity are the coefficients. A single equation adds up the cost of inputs needed to produce a given commodity at a given price. There will be as many equations as there are commodities to produce.

It is then possible to separate the prices into a column of numbers called a vector, and the quantities into a square of numbers called a matrix – where, as noted earlier, every element is either a positive number or zero. The properties of this matrix can then be analyzed mathematically, and its mathematical properties can be used to answer economic questions.

This matrix is known as a Leontief input-output matrix, after the Russian economist who first developed this method of analysis. The first row of such a matrix effectively says that 'a units of commodity a combined with b units of commodity b and z units of commodity z will produce 1 unit of commodity a.' It is the simplest method of describing a system of production, in that it implies that there is one and only one best way to make each commodity: no substitution of one technology for another is allowed.

While this is a much simpler model of production than economists like to work with, it turns out that the properties of this very simple system determine whether the equilibrium of any more general model is stable. If this simple system can't guarantee stability, then no more complex system is going to either (this is a general property of dynamic models: the stability of the system very close to its equilibrium is determined by its 'linear'

parts, and Leontief's matrix is the linear component of any more complex model of production).

There are two stability conditions in the simple Leontief system: the quantities produced each year have to enable the system to reproduce itself (this won't happen if, for example, the required inputs of iron for year 10 exceed the output of iron in year 9); and the prices must be feasible (the iron-producing sector can't depend on the price of some required input to producing iron being negative, for example).

It turns out that the first stability condition is governed by a characteristic of the input-output matrix, whereas the second stability condition is governed by the same characteristic of the inverse of that matrix. As with simple constants, a matrix and its inverse have, to some extent, opposite properties. Thus if you have a constant a which is less than 1, then a squared will be much less than 1, a cubed even more so, and higher powers of a will eventually converge to zero. However, the inverse of a, $1/a$, will be greater than 1, and powers of $1/a$ will blow out to infinity. If the stability of some system depends upon both a and the inverse of a being less than 1, then no number can fulfill both requirements, and the system is going to be unstable.

Since economic models are supposed to concern themselves with real economies, which can and do change in size, the general conclusion is that a real economy will never be in a state of general equilibrium. If economics is to have any relevance to the real world – if economics is even to be internally consistent – then it must be formulated in a way which does not assume equilibrium. Time, and dynamic analysis, must finally make an appearance in economic analysis.

A transitional methodology?

The founding fathers of economics had no problem accepting such a conclusion. In fact, to them, static analysis was merely a stop-gap measure, a transitional methodology which would be superseded by dynamic analysis as economics reached maturity. Jevons, for example, argued that 'If we wished to have a complete solution we should have to treat it as a problem of dynamics.' But he instead pioneered static analysis because 'it would

surely be absurd to attempt the more difficult question when the more easy one is yet so imperfectly within our power' (Jevons 1888).

Similarly, and at more length, Marshall noted that

> The Mecca of the economist lies in economic biology rather than in economic dynamics. But biological conceptions are more complex than those of mechanics; a volume on Foundations must therefore give a relatively large place to mechanical analogies; and frequent use is made of the term 'equilibrium,' which suggests something of statical analogy. This fact, combined with the predominant attention paid in the present volume to the normal conditions of life in the modern age, has suggested the notion that its central idea is 'statical,' rather than 'dynamical.' But in fact it is concerned throughout with the forces that cause movement: and its key-note is that of dynamics, rather than statics. (Marshall 1920 [1890]: Preface, para. 19)

At the end of the nineteenth century, J. B. Clark, the economist who developed the marginal productivity theory of income distribution (critiqued in Chapter 5), looked forward to the twentieth century as the period during which economic dynamics would supplant economic statics:

> A point on which opinions differ is the capacity of the pure theory of Political Economy for progress. There seems to be a growing impression that, as a mere statement of principles, this science will fairly soon be complete. It is with this view that I take issue. The great coming development of economic theory is to take place, I venture to assert, through the statement and solution of dynamic problems. (Clark 1898)

In this paper, Clark gave many good reasons why economics should be analyzed using dynamics rather than statics. Foremost among these was that 'A static state is imaginary. All actual societies are dynamic; and those that we have principally to study are highly so. Heroically theoretical is the study that creates, in the imagination, a static society' (ibid.).

One century later, economic dynamics has indeed been developed – but not by the school to which J. B. Clark belonged. Instead, neoclassical economics still by and large ignores the issue of time. Students are often

told that dynamics is important, but they are taught nothing but statics. A typical undergraduate macroeconomics textbook, for example, states that 'the examination of the process of moving from one equilibrium to another is important and is known as dynamic analysis.' However, it then continues that 'Throughout this book we will assume that the economic system is stable and most of the analysis will be conducted in the comparative static mode' (Taslim and Chowdhury 1995).

The leading textbook used today to teach graduate students makes a similar claim – that while other disciplines use dynamics, economists model processes as if they occur in equilibrium because economists are good at identifying equilibrium! Two-thirds through his voluminous 1,000-page tome, Mas-Colell, the current doyen of neoclassical instruction, writes:

> We have, so far, carried out an extensive analysis of equilibrium equations. A characteristic feature that distinguishes economics from other scientific fields is that, for us, the equations of equilibrium constitute the center of our discipline. Other sciences, such as physics or even ecology, put comparatively more emphasis on the determination of dynamic laws of change. In contrast, up to now, we have hardly mentioned dynamics.
>
> The reason, informally speaking, is that economists are good (or so we hope) at recognizing a state of equilibrium but poor at predicting how an economy in disequilibrium will evolve.
>
> Certainly there are intuitive dynamic principles: if demand is larger than supply, then the price will increase, if price is larger than marginal cost then production will expand, if industry profits are positive and there are no barriers to entry, then new firms will enter and so on. The difficulty is in translating these informal principles into precise dynamic laws. (Mas-Colell et al. 1995: 620)

This is nonsense, and to give Mas-Colell his due I think he realizes it here. Economists model in equilibrium, not because they are 'good (or so we hope) at recognizing a state of equilibrium,' but simply because they can't get the results they want in dynamic analysis and have therefore not

made the leap from static to dynamic modeling that has occurred in all other disciplines.

Mas-Colell admits this when he discusses the attempts to generalize Walras's *tatonnement* process to a disequilibrium one. While he argues that a two-commodity exchange economy is stable,[6] he admits that this result does not generalize to three or more commodities: 'Unfortunately, as soon as [there are more than two goods] neither the local conclusions nor the global conclusions of the two-commodity case generalize' (ibid.: 622).

This may be unfortunate, but the correct reaction to it is to abandon static analysis and work in disequilibrium. This, clearly, is not what neoclassical economists have done – and unfortunately, economists of many other persuasions also use static analysis because they believe that equilibrium is the enduring state of the economy, while dynamics merely captures the transient moments between different equilibria. For example, a Sraffian economist defended static methodology in economics by arguing that '"static" analysis does not "ignore" time. To the contrary, that analysis allows enough time for changes in prime costs, markups, etc., to have their full effects' (Steedman 1992).

As this chapter shows, this confidence that 'the end point of a dynamic process is the state of static equilibrium' is false. Equally false was the belief of the founding fathers of economics, that dynamic analysis 'does not invalidate the conclusions of a static theory' (Clark 1898). But even if they were right, even if dynamic forces did lead, eventually, to static outcomes, it would still be invalid to model the economy using static techniques. Keynes put the case best in 1923, when he made his oft-quoted but rarely appreciated observation that 'in the long run we are all dead.' The full statement gives a rather better picture of his intent: 'But this long run is a misleading guide to current affairs. In the long run we are all dead. Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is long past the ocean is flat again' (Keynes 1971 [1923]).

Keynes was right: it is not valid to ignore the transient state of the economy. As Fisher later observed in very similar terms, equilibrium conditions in the absence of disturbances are irrelevant, because

disturbances will always occur. Whether equilibrium is stable or not, disequilibrium will be the state in which we live:

> We may tentatively assume that, ordinarily and within wide limits, all, or almost all, economic variables tend, in a general way, toward a stable equilibrium […]
>
> It follows that, unless some outside force intervenes, any 'free' oscillations about equilibrium must tend progressively to grow smaller and smaller, just as a rocking chair set in motion tends to stop.
>
> But the exact equilibrium thus sought is seldom reached and never long maintained. New disturbances are, humanly speaking, sure to occur, so that, in actual fact, any variable is almost always above or below the ideal equilibrium […]
>
> Theoretically there may be – in fact, at most times there must be – over- or under-production, over- or under-consumption, over- or under-spending, over- or under-saving, over- or under-investment, and over or under everything else. It is as absurd to assume that, for any long period of time, the variables in the economic organization, or any part of them, will 'stay put,' in perfect equilibrium, as to assume that the Atlantic Ocean can ever be without a wave. (Fisher 1933: 339)

We also live in a changing – and normally growing – economy. Surely we should be concerned, not with absolute levels of variables, but with their rates of change? Should not demand and supply analysis, for instance, be in terms of the rate of change of demand, and the rate of change of supply? Should not the outcome of supply and demand analysis be the rate of change of price and quantity over time, rather than static levels? Should not macroeconomics concern itself with the rate of change of output and employment, rather than their absolute levels?

Of course they should. As Keynes also once remarked, 'equilibrium is blither.' So why, fifty years after Keynes, are economists still blithering? Why do economists persist in modeling the economy with static tools when

dynamic ones exist; why do they treat as stationary entities which are forever changing?

There are many reasons, but the main one, as outlined in the previous chapter, is the extent to which the core ideological beliefs of neoclassical economics are bound up in the concept of equilibrium. As a by-product of this, economists are driven to maintain the concept of equilibrium in all manner of topics where dynamic, non-equilibrium analysis would not only be more relevant, but frankly would even be easier. This obsession with equilibrium has imposed enormous costs on economics.

First, unreal assumptions are needed to maintain conditions under which there will be a unique, 'optimal' equilibrium. These assumptions are often justified by an appeal to Friedman's methodological 'assumptions don't matter' argument, but as Chapter 8 pointed out, this notion is easily debunked. However, most economists take it as an article of faith, with insidious results. If you believe you can use unreality to model reality, then eventually your grip on reality itself can become tenuous – as Debreu's bizarre model of general equilibrium indicates.

Secondly, as shown in this chapter, even the unreal assumptions of general equilibrium theory are insufficient to save it from irrelevance, since even the model of general equilibrium has been shown to be unstable, so that no modeled or real economy could ever be in a state of equilibrium. Many of those who pioneered general equilibrium analysis are grudgingly conceding that these results require economics to radically alter direction. But they are also quite aware that lesser economists are, as Alan Kirman put it, 'not even concerned over the sea-worthiness of the vessel in which they are sailing' (Kirman 1989).

Thirdly, the emphasis on modeling everything as an equilibrium phenomenon has isolated economics from most if not all other sciences, where dynamic analysis – and in particular evolutionary analysis – is now dominant. Economists are now virtually the only 'scientists' who attempt to model a real-world system using static, equilibrium tools. As a result of this isolation, economists have been shielded from developments in mathematics and other sciences which have revolutionized how scientists perceive the world.

This isolation is to some extent fortuitous, because if economists really knew what is common knowledge in other sciences, then they would finally

have to abandon their obsession with equilibrium, and economics as outlined in this book would cease to exist. Most modern-day economists believe, as did the founding fathers of economics, that dynamic analysis would simply 'fill in the dots' between the static snapshots, thus replacing a series of still photographs with a moving picture. In fact, modern research in mathematics, physics, biology and many other disciplines has shown that dynamic analysis normally leads to results which contradict those of static analysis.

In the long run, we are all in the short run

Equilibrium can be the long-run destination of the economy only if it is stable – if any divergence sets up forces which will return the economy to equilibrium. Even after the proofs of the instability of general equilibrium, most economists believe that this is a non sequitur: surely, the equilibrium of any real-world system must be stable, since if it were unstable, wouldn't it break down? John Hicks articulated this view when he criticized one of the earliest dynamic models developed by an economist. He commented that Harrod (1939)

> welcomes the instability of his system, because he believes it to be an explanation of the tendency to fluctuation which exists in the real world. I think, as I shall proceed to show, that something of this sort may well have much to do with the tendency to fluctuation. But mathematical instability does not in itself elucidate fluctuation. A mathematically unstable system does not fluctuate; it just breaks down. The unstable position is one in which it will not tend to remain. (Hicks 1949)

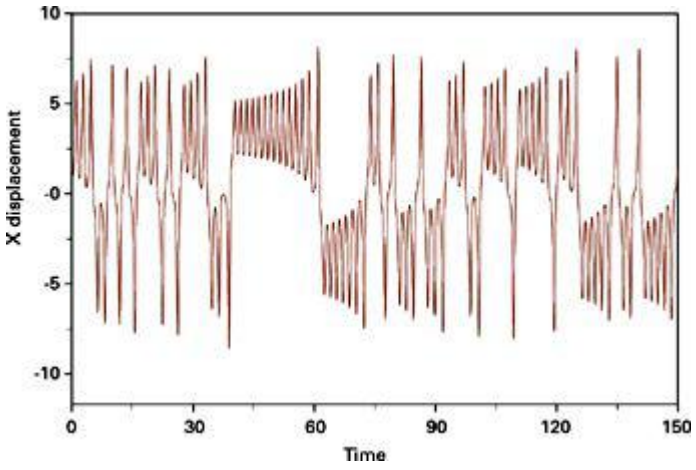
The modern discipline known colloquially as chaos theory has established that this belief, though still widespread among economists today, is quite simply wrong. The equilibrium of a real-world system can be unstable without the system itself breaking down.

The first and best illustration of this occurred, not in economics, but in meteorology. I'll give a brief exposition of this model, because it illustrates several ways in which the conventional economic understanding of dynamics is profoundly wrong. But first, we need a brief technical interlude

to explain the difference between the mathematical methods used in static analysis and those used in dynamics (you can skip to 'The weather and the butterfly' if you'd like to avoid mathspeak).

Straight lines and curved paths What static analysis means in technical terms is that the equations most neoclassical economists (and many non-orthodox economists) use in their mathematical models are 'algebraic' rather than 'differential.'

Algebraic equations are simply larger and more complicated versions of the equations we all did at school in geometry, when we were asked to work out the intersection of two lines. Given two equations for Y in terms of X, with different slopes and Y intercepts, we worked out the only X point where the two formulas gave the same Y point. Continuing with the geometry analogy, most of the equations used by economists use only straight lines, rather than more complicated shapes like parabolas, etc. Algebraic techniques with these equations scale indefinitely – you can have equations with hundreds of 'straight lines' and still get unique solutions.



9.2 The time path of one variable in the Lorenz model

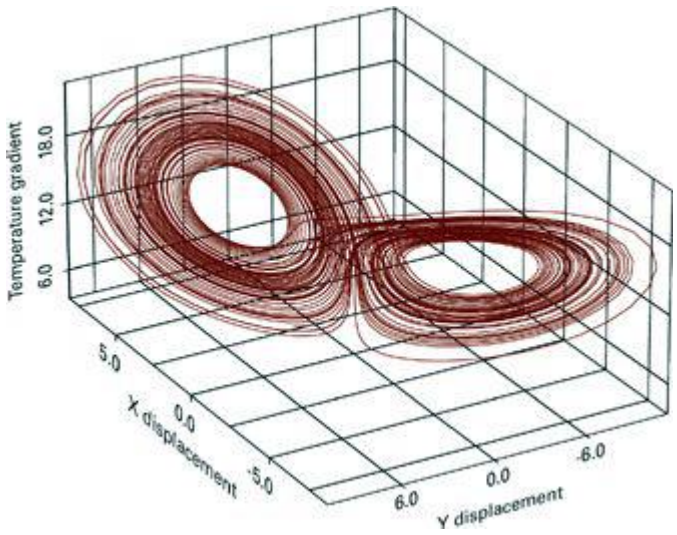
Differential equations, on the other hand, are more complicated descendants of the technique of differentiation, which you might have learnt if you did calculus at school or college. Rather than being expressed in terms of X and Y, these equations are expressed in terms of the rate of change of X and the rate of change of Y. While school calculus dealt only with 'the rate of change of Y with respect to X,' differential equations

typically are in terms of 'the rate of change of Y with respect to Y itself, other variables, and time.'

Most differential equation models also involve curved relationships between variables, rather than straight lines. A straight line is in fact the simplest type of relationship which can exist between two variables (other than that of no relationship at all). Straight-line relationships in differential equation models with unstable equilibria lead to ultimately absurd outcomes, such as negative prices, or cycles which approach infinite amplitude as time goes on. Nonlinear relationships, however, result in bounded behavior: the forces which repel the system when it is very close to equilibrium are eventually overwhelmed by attractive forces when the system is substantially distant from the equilibrium.

Unlike linear algebraic equations, nonlinear differential equations don't scale well. Only a very few simple nonlinear differential equations can be solved – the vast majority can't be solved at all. Once there are more than two variables in a system of nonlinear differential equations, there is in fact no analytic solution. Such systems must be simulated to see what is actually going on.

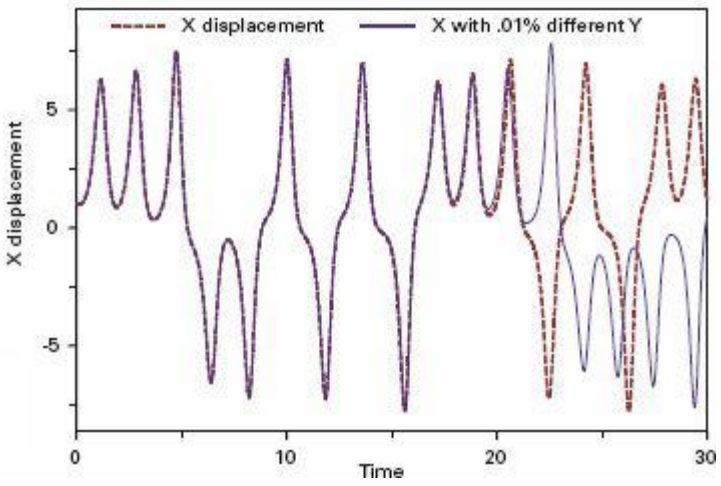
The weather and the butterfly In 1963, the meteorologist E. N. Lorenz devised a simple mathematical model of turbulent flow in a weather cell, using a simplified version of a well-known mathematical model of turbulent flow. His model had just three equations, with three variables and three constants. The first (x) equation described the intensity of convective motion, the second (y) the temperature difference between ascending and descending columns of air, and the third (z) described the divergence from linearity of the temperature gradient across the weather cell.[7]



9.3 Structure behind the chaos

It would be hard to think of a simpler set of three equations, and yet the behavior they generated was unbelievably complex. Figure 9.2 shows the time path of the east–west fluid displacement.

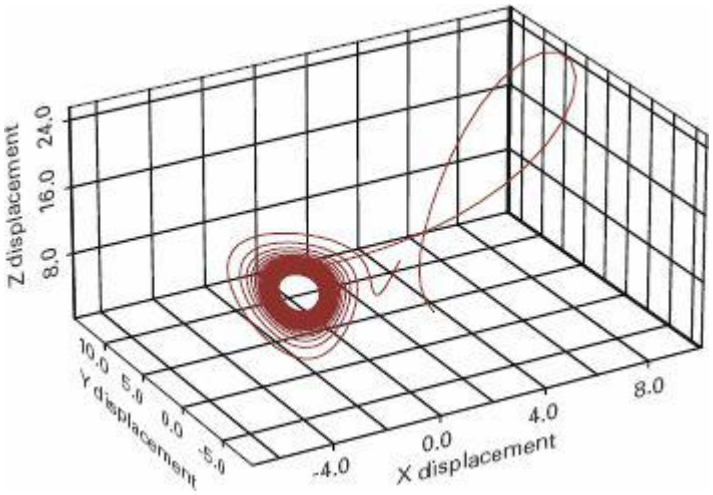
The y and z patterns were equally complex. Even more mysteriously, a tiny difference in the initial x, y or z values led, very quickly, to a totally different time path. It had been thought in the past that a tiny difference in any initial measurement would mean only a tiny error in predicting the future behavior of a variable. However, in this model, a tiny difference initially has no apparent effect, but then abruptly leads to a totally different outcome.



9.4 Sensitive dependence on initial conditions

Finally, though the pattern for any one variable appeared erratic, behind this apparent randomness lay a beautiful structure which is visible when the three variables are plotted on the one graph. Figure 9.3 shows the 'butterfly' behind the superficial chaos.

Detailed analysis of this system reveals that it has not one equilibrium, but three. *More importantly, all three equilibria are unstable*. A slight divergence from any equilibrium causes the system moving to move away from it very rapidly. A tiny divergence from one equilibrium point leads to the system instantly being propelled from that equilibrium. It then approaches another, only to be flung off to a third. It orbits that equilibrium, only to be eventually repelled from it. Finally, it approaches and then is repelled from the second equilibrium back towards the first.



9.5 Unstable equilibria

There are at least four lessons for economics in this model.

First, a system with unstable equilibria doesn't have to 'break down.' Instead, such a system can display complex cyclical behavior rather like that we see in real-world weather – and, more to the point, in real-world economies.

Secondly, if the equilibria of a model are unstable, then neither the initial nor the final position of the model will be equilibrium positions. The economic belief that dynamic analysis simply plots the movement between one equilibrium and another is therefore wrong. Instead, even simple dynamic models will display 'far from equilibrium' behavior. As a result,

rather than equilibrium being where the action is, equilibrium tells you where the model will never be.

Thirdly, extrapolating from models to the real world, actual economic variables are likely to always be in disequilibrium – even in the absence of external shocks (or 'exogenous' shocks, as economists prefer to call them), which is the usual economic explanation for cycles – and the conditions which economists have 'proved' apply at equilibrium will therefore be irrelevant in actual economies. In this sense, equilibrium truly is, as Keynes put it, 'blither.' Static economic analysis therefore can't be used as a simplified proxy for dynamic analysis: the two types of analysis will lead to completely different interpretations of reality. In all such cases, the static approach will be completely wrong and the dynamic approach will be at least partially right.

Finally, even as simple a system as Lorenz's, with just three variables and three constants, can display incredibly complex dynamics because the interactions between variables are nonlinear (if you check the equations in note 7, you will see terms like 'x times y'). As noted earlier, nonlinear relationships in differential equation models can lead to complex but bounded behavior.

From meteorology to economics

There are many models in economics which have properties akin to those of Lorenz's weather model – very few of which have been developed by neoclassical economists. Most were instead developed by economists who belong to alternative schools, in particular complexity theorists and evolutionary economists. One of the best-known such models, Goodwin's model of cyclical growth, put in mathematical form a model first suggested by Marx.

Marx argued that – in a highly simplified economy consisting of just capitalists and workers – there would be cycles in employment and income shares. In Marx's words:

> A rise in the price of labor, as a consequence of accumulation of capital [… means that] accumulation slackens in consequence of the rise in the price of labor, because the stimulus of gain is blunted. The rate of accumulation lessens; but with its lessening,

> the primary cause of that lessening vanishes, i.e., the disproportion between capital and exploitable labor-power.
>
> The mechanism of the process of capitalist production removes the very obstacles that it temporarily creates. The price of labor falls again to a level corresponding with the needs of the self-expansion of capital, whether the level be below, the same as, or above the one which was normal before the rise of wages took place […]
>
> To put it mathematically: the rate of accumulation is the independent, not the dependent, variable; the rate of wages, the dependent, not the independent, variable. (Marx 1867: ch. 25, section 1)[8]

In point form, the model is as follows:

- A high rate of growth of output led to a high level of employment.
- The high level of employment encouraged workers to demand large wage rises, which reduced profits.
- The reduced level of profits caused investment to decline, and growth to slow.
- The slower rate of growth led to increasing unemployment, which in turn led to workers accepting lower wages.
- Eventually the fall in workers' share of output restored profit to levels at which investment would resume, leading to a higher rate of growth and higher employment levels.
- This in time led to high wage demands once more, thus completing the cycle.

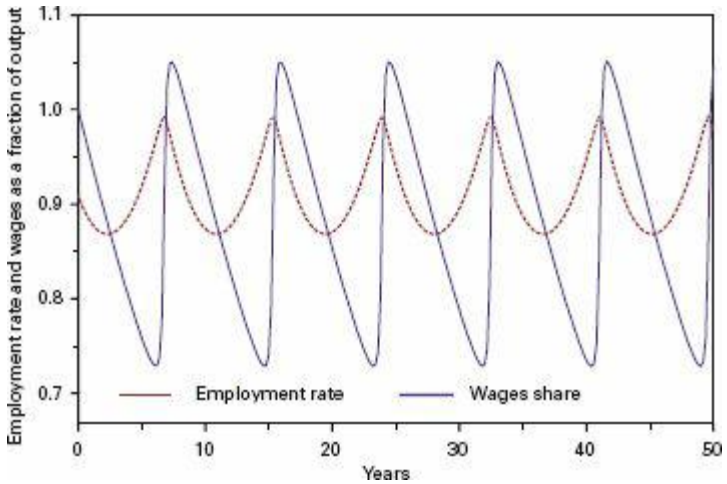
This cycle can also be stated in terms of causal relationships between key economic variables – the amount of capital, the level of output, and so on – which shows that the process Marx describes was based on an accurate view of the overall structure of the economy, and also an accurate deduction that this would lead to cycles in income distribution and employment, rather than either equilibrium or breakdown:

1 The amount of physical capital determines the amount of output.
2 Output determines employment.
3 The rate of employment determines the *rate of change* of wages (the 'Phillips Curve' relationship I discuss in the addendum to this chapter).
4 Wages times employment determines the wage bill, and when this is subtracted from output, profit is determined.
5 Profit determines the level of investment.
6 Investment determines the rate of change of capital – and this closes the causal loop of the model.

In mathematical form, this model reduces to two equations which are easily stated verbally:

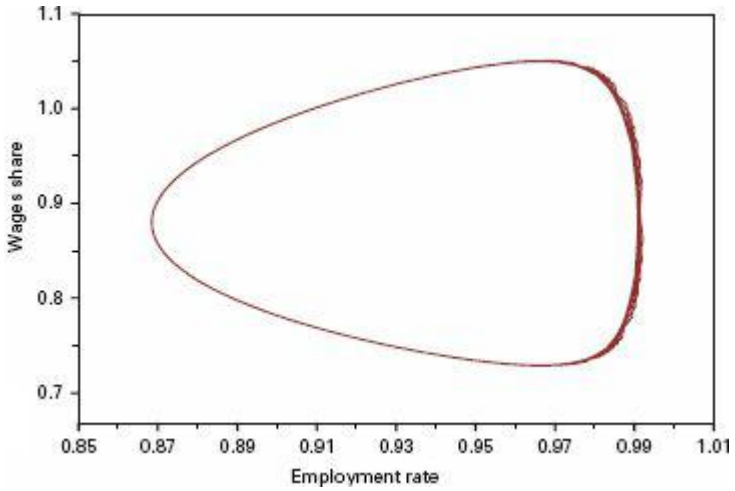
- The rate of change of workers' share of output equals workers' wage demands minus the rate of productivity growth.
- The rate of change of employment equals the rate of growth of output, minus population growth and technological change.[9]

This mathematical model generates the cycle envisaged by Marx. Rather than converging to equilibrium values, workers' share of output and the rate of employment both cycle indefinitely.



9.6 Cycles in employment and income shares

When wages share and employment are plotted against each other, the result is a closed loop. This is a far less complex structure than Lorenz's model, but it has one thing in common with it: *the model does not converge to its equilibrium* (which lies in the center of the loop), *but orbits around it indefinitely*.



9.7 A closed loop in employment and wages share of output

It is also easily extended to capture more aspects of the real world, and when this is done, dynamic patterns as rich as those in Lorenz's model appear – as I detail in Chapters 13 and 14.

Real-world phenomena therefore simply cannot be modeled using 'comparative statics' or equilibrium – unless we are willing to believe that cyclones are caused by something 'exogenous' to the weather, and stock market bubbles are caused by something outside the economy. Complexity theory has established that such phenomena can be modeled dynamically, so that abandoning static equilibrium analysis does not mean abandoning the ability to say meaningful things about the economy.

Instead, what has to be abandoned is the economic obsession with achieving some socially optimal outcome. As noted in this and the previous chapter, economists have conflated the concept of equilibrium with the vision of an 'economic utopia' in which no one could be made better off without making someone else worse off. But a free market economy could never remain in an optimal position, because economic equilibria are

unstable. The real question is whether we can control such an unstable system – whether we can constrain its instability within acceptable bounds.

This question was once at the heart of what is known as macroeconomics – the study of the entire economy and the attempt to control it using government fiscal and monetary policy. Unfortunately, as we shall see in the next chapter, neoclassical economists have emasculated this once virile area of analysis. As they did so, they ignored possibly the most important lesson to flow from the advances in dynamic analysis since Lorenz: the realization that complex systems have what are known as 'emergent behaviors' which mean that they cannot be understood by studying their constituent parts alone. This reality invalidates a key aspect of modern neoclassical macroeconomics: the attempt to derive models of the macroeconomy from microeconomic models of the behavior of individuals. A discussion of emergent behavior properly belongs in this chapter, but its neglect by neoclassical economists – and the practice of its opposite philosophy, 'reductionism' – has been so essential to the neoclassical ruination of macroeconomics that I have delayed a discussion of it until the next chapter.

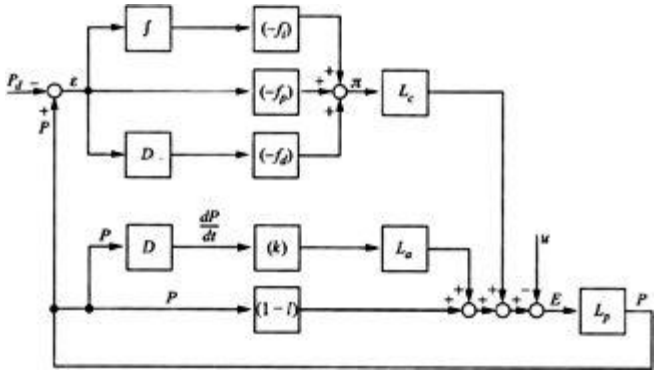
Before I move on, there is one other topic that also belongs in this chapter, rather than the next on macroeconomics, where it would normally be discussed in a conventional textbook: the 'Phillips Curve.' This is an alleged relationship between the level of unemployment and the rate of inflation that, though it is hotly disputed within economics, nonetheless plays a role in virtually every theory of macroeconomics, from Marx's at one extreme to neoclassical economics at the other.

It belongs in this chapter on dynamics, because the real objective of the person after whom it was named – the New Zealand-born engineer-turned-economist A. W. ('Bill') Phillips – was to persuade economists to abandon their static methods and embrace dynamic analysis. This is precisely what I am attempting to do now, so Phillips's work – including the 'Phillips Curve' – deserves to be discussed here as a valiant but unsuccessful previous attempt to shake economists out of their static straitjackets.

Addendum: Misunderstanding Bill Phillips, wages and 'the Phillips Curve'

Bill Phillips the man was undoubtedly one of the most dynamic human beings of all time. Compared to that of Phillips, the lives of most economists – even non-neoclassical ones – are as pale as the theories that neoclassical economists have concocted about the world. He left school at fifteen, worked as a crocodile hunter and gold miner in Australia, learnt engineering by correspondence, was awarded an MBE for his role in the defence of Singapore in 1942, and, as a prisoner of war, made a miniaturized radio from components he stole from the camp commander's radiogram. Despite the effects of malnutrition and abuse in the camp, within five years of the war finishing – and while still an undergraduate student of economics – he had his first paper published in a leading journal (Phillips 1950). The paper described an analog computer dynamic simulation model of the economy (MONIAC) that he constructed at a cost of £400, just three years after the first digital computer (ENIAC) had been constructed at a cost of US$500,000 (Leeson 1994, 2000).

MONIAC put into mechanical-hydraulic form the principles of dynamics that Phillips had learnt as an engineer, and it was this approach which he tried to communicate to economists, on the sound basis that their preferred methodology of comparative statics was inappropriate for economic modeling:



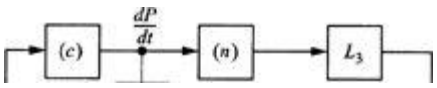
9.8 Phillips's functional flow block diagram model of the economy

> RECOMMENDATIONS for stabilizing aggregate production and employment have usually been derived from the analysis of multiplier models, using the method of comparative statics. This type of analysis does not provide a very firm basis for policy recommendations, for two reasons.

> First, the time path of income, production and employment during the process of adjustment is not revealed. It is quite possible that certain types of policy may give rise to undesired fluctuations, or even cause a previously stable system to become unstable, although the final equilibrium position as shown by a static analysis appears to be quite satisfactory.
>
> Second, the effects of variations in prices and interest rates cannot be dealt with adequately with the simple multiplier models which usually form the basis of the analysis. (Phillips 1954: 290)

Phillips instead proposed that economists should build dynamic models of the economy – models in which time was embraced rather than ignored via the device of comparative statics – and his underlying method here was the functional flow block diagram. This had been devised by engineers in the 1920s as a way to visually represent dynamic processes, which previously had been shown as either differential equations, or transformations of these equations into other mathematical forms.[10] Phillips drew such a diagrammatic representation of a simple dynamic economic model (ibid.: Fig. 10, p. 306; see Figure 9.8), with symbols to indicate operations like time lags, differentiation and integration with respect to time, addition and subtraction, etc. The model recast the standard comparative-static, multiplier-accelerator models of the time into dynamic form.

This model was only the starting point of a project to develop a complete dynamic model of the economy, in which the feedback effects and disequilibrium dynamics that were ignored by the conventional 'Keynesian' models of the time could be fully accounted for.



9.9 The component of Phillips's Figure 12 including the role of expectations in price setting

In particular, Phillips extended his model to consider the impact of expectations upon prices. Given how much his work has been falsely denigrated by neoclassical economists for ignoring the role of expectations

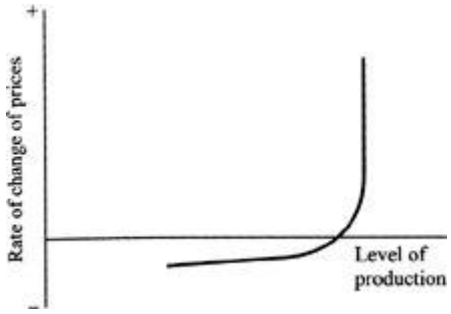
in economics, this aspect of his model deserves attention prior to considering the Phillips Curve itself:

> Demand is also likely to be influenced by the rate at which prices are changing […] this influence on demand being greater, the greater the rate of change of prices […] *The direction of this change in demand will depend on expectations about future price changes*. If changing prices induce expectations of further changes in the same direction, as will probably be the case after fairly rapid and prolonged movements, demand will change in the same direction as the changing prices […]
>
> *If, on the other hand, there is confidence that any movement of prices away from the level ruling in the recent past will soon be reversed, demand is likely to change in the opposite direction to the changing prices* […]. (Ibid.: 311; emphases added)

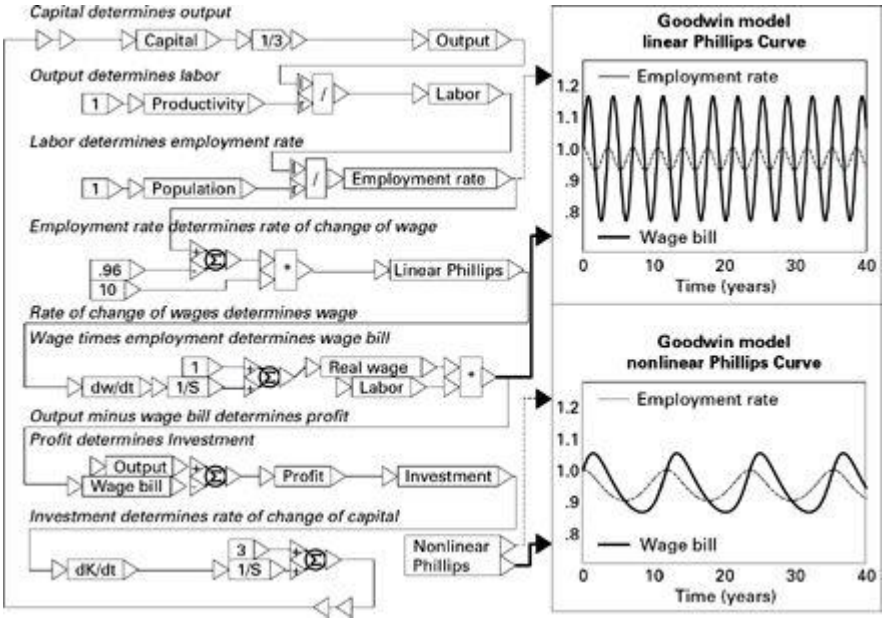
Phillips didn't merely talk about expectations: he extended his model to incorporate them – see Figure 9.9.

As part of this project, Phillips also hypothesized that there was a nonlinear relationship between 'the level of production and the rate of change of factor prices [labor and capital]' (ibid.: 308), and he sketched a hypothetical curve for this relationship – see Figure 9.10.

The role of this relationship in his dynamic model was to limit the rate at which prices would fall when unemployment was high, in line with 'the greater rigidity of factor prices in the downward than in the upward direction' (ibid.: 308). In a dynamic model itself, this does not lead to a stable trade-off between inflation and unemployment – which is the way his empirically derived curve was subsequently interpreted – but rather limits the volatility of the cycles that occur compared to what a linear relationship would yield.



9.10 Phillips's hand drawing of the output–pricechange relationship



9.11 A modern flow-chart simulation program generating cycles, not equilibrium

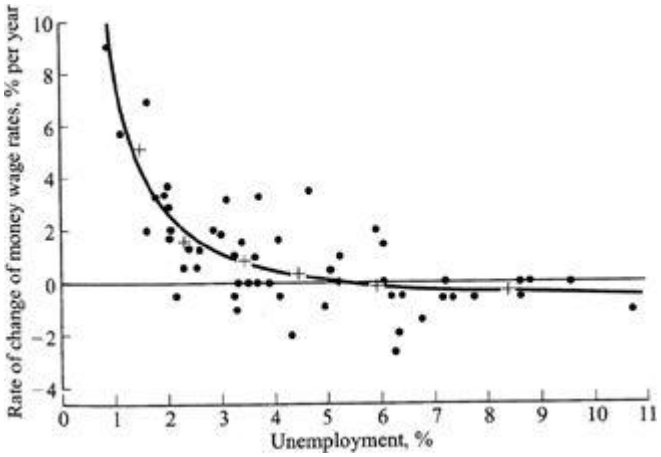
This was hard for Phillips to convey in his day, because then functional flow block diagrams were merely means to describe a dynamic model – they didn't let you simulate the model itself. But today, numerous computer programs enable these diagrams to be turned into active simulations. There is also an enormous analytic framework for analyzing stability and incomplete information supporting these programs: engineers have progressed dramatically in their capacity to model dynamic processes, while economics has if anything gone backwards.

Figure 9.11 illustrates both these modern simulation tools, and this difference between a linear and a nonlinear 'Phillips Curve' in Goodwin's growth cycle model. One of these programs (Vissim) turns the six-step verbal description of Marx's cycle model directly into a numerical

simulation, using a linear 'Phillips Curve.' This model cycles as Marx expected, but it has extreme, high-frequency cycles in both employment and wages share.

Embedded in the diagram is an otherwise identical model, which has a nonlinear Phillips Curve with the shape like that envisaged by Phillips. This has smaller, more realistic cycles and these have a lower frequency as well, closer to the actual frequency of the business cycle.

What this model doesn't have – and this is a very important point – is an equilibrium 'trade-off' between inflation (proxied here by the rate of change of wages) and unemployment. Instead the model economy is inherently cyclical, and Phillips's overall research agenda was to devise policy measures – inspired by engineering control theory – that might attenuate the severity of the cycles.



9.12 Phillips's empirically derived unemployment–money-wage-change relation

Had Phillips stuck with just a sketch of his hypothesized nonlinear relationship between the level of production and factor prices, it is possible that he would be known today only for these attempts to develop dynamic economic analysis – and possibly relatively unknown too, given how other pioneers of dynamics like Richard Goodwin (Goodwin 1986, 1990) and John Blatt (Blatt 1983) have been treated. Instead, he made the fateful decision to see whether he could find such a relationship in the UK data on unemployment and the rate of change of money wages.

This decision led to him being immortalized for work that he later told a colleague 'was just done in a weekend' while 'his best work was largely

ignored – his early control work' (Leeson 1994: 613).

To do his statistical analysis, Phillips assembled annual data for the UK from 1861 until 1957 from a range of sources. He then used the subset from 1861 till the outbreak of World War I to derive a nonlinear function that appeared to fit the data very tightly (see Figure 9.12). When he fitted the post-WWI data to this curve, the 'out of sample' data also had a relatively close fit to his equation (except for some deviations which he explained as due to negotiated inflation-wage deals between unions and employers, and the impact of World War II on forcing up agricultural prices in Britain).

He then summarized his results in the following accurate but poorly considered statement:

> Ignoring years in which import prices rise rapidly enough to initiate a wage-price spiral, which seem to occur very rarely except as a result of war, and assuming an increase in productivity of 2 per cent per year, it seems from the relation fitted to the data that if *aggregate demand were kept at a value which would maintain a stable level of product prices the associated level of unemployment would be a little under 2 per cent*.
>
> *If*, as is sometimes recommended, *demand were kept at a value which would maintain stable wage rates the associated level of unemployment would be about 5 per cent*. (Phillips 1954: 299; emphases added)

To actually achieve the preconditions that Phillips set out here – keeping aggregate demand 'at a value which would maintain a stable level of product prices' or 'at a value which would maintain stable wage rates' – would have required a whole host of control mechanisms to be added, even to Phillips's model of the economy, let alone the real economy itself. As the Goodwin model indicates, a dynamic model of the economy will have endogenous tendencies to cyclical behavior, and these in turn are merely a caricature of the cyclical nature of evolutionary change in a capitalist economy.

Developing these control mechanisms was, as noted, Phillips's main research agenda, but the economics profession at large, and politicians as well, latched on to this statement as if it provided a simple menu by which

the economy could be controlled. If you wanted stable prices (in the UK), just set unemployment to 2 percent; if you wanted stable money wages instead, set unemployment to 5 percent; and pick off any other combination you like along the Phillips Curve as well.

This simplistic, static 'trade-off' interpretation of Phillips's empirically derived curve rapidly came to be seen as the embodiment of Keynesian economics, and since the 1960s data also fitted the curve very well, initially this appeared to strengthen 'Keynesian' economics.

But in the late 1960s, the apparent 'trade-off' began to break down, with higher and higher levels of both inflation and unemployment. Since the belief that there was a trade-off had become equivalent in the public debate to Keynesian economics, the apparent breakdown of this relationship led to a loss of confidence in 'Keynesian' economics – and this was egged on by Milton Friedman as he campaigned to restore neoclassical economics to the position of primacy it had occupied prior to the Great Depression.

Phillips's empirical research recurs throughout the development of macroeconomics, as I am about to recount in the next chapter – as Robert Leeson observed: 'For over a third of a century, applied macroeconomics has, to a large extent, proceeded from the starting point of the trade-off interpretation of the work of A. W. H. "Bill" Phillips. It is hardly an exaggeration to say that any student destitute of the geometry of the Phillips curve would have difficulty passing an undergraduate macroeconomics examination' (Leeson 1997: 155).

However, even his empirical research has been distorted, since it has focused on just one of the factors that Phillips surmised would affect the rate of change of money wages – the level of employment. Phillips in fact put forward three causal factors:

> When the demand for a commodity or service is high relatively to the supply of it we expect the price to rise, the rate of rise being greater the greater the excess demand. Conversely when the demand is low relatively to the supply we expect the price to fall, the rate of fall being greater the greater the deficiency of demand. It seems plausible that this principle should operate as one of the factors determining the rate of change of money wage rates, which are the price of labor services.

> When the demand for labor is high and there are very few unemployed we should expect employers to bid wage rates up quite rapidly, each firm and each industry being continually tempted to offer a little above the prevailing rates to attract the most suitable labor from other firms and industries. On the other hand it appears that workers are reluctant to offer their services at less than the prevailing rates when the demand for labor is low and unemployment is high so that wage rates fall only very slowly. The relation between unemployment and the rate of change of wage rates is therefore likely to be highly non-linear.

Phillips then added that the *rate of change* of employment would affect the rate of change of money wages:

> It seems possible that a second factor influencing the rate of change of money wage rates might be the rate of change of the demand for labor, and so of unemployment. Thus in a year of rising business activity, with the demand for labor increasing and the percentage unemployment decreasing, employers will be bidding more vigorously for the services of labor than they would be in a year during which the average percentage unemployment was the same but the demand for labor was not increasing. Conversely in a year of falling business activity, with the demand for labor decreasing and the percentage unemployment increasing, employers will be less inclined to grant wage increases, and workers will be in a weaker position to press for them, than they would be in a year during which the average percentage unemployment was the same but the demand for labor was not decreasing.

Thirdly, he considered that there could be a feedback between the rate of inflation and the rate of change of money wages – though he tended to discount this except in times of war: 'A third factor which may affect the rate of change of money wage rates is the rate of change of retail prices, operating through cost of living adjustments in wage rates' (Phillips 1954: 283).

In subsequent work, Phillips went farther still, and considered that attempts to control the economy that relied upon the historically observed

relationship could change the relationship itself: 'In my view it cannot be too strongly stated that in attempting to control economic fluctuations we do not have two separate problems of estimating the system and controlling it, we have a single problem of jointly controlling and learning about the system, that is, a problem of learning control or adaptive control' (Phillips 1968: 164; Leeson 1994: 612, n. 13).

Phillips didn't consider the other two causal relationships in his empirical work because, at the time he did it, and with the computing resources available to him (a hand-operated electronic desk calculator), quite simply, it was impossible to do so. But today it is quite feasible to model all three causal factors, and adaptive learning as well, in a modern dynamic model of the kind that Phillips had hoped to develop.

Unfortunately, Phillips's noble intentions resulted in a backfire: far from helping wean economists off their dependency on static methods, the misinterpretation of his simple empirical research allowed the rebirth of neoclassical economics and its equilibrium methodology – and ultimately, the reduction of macroeconomics to applied microeconomics.

10 | WHY THEY DIDN'T SEE IT COMING

Why the world's leading macroeconomists were the last ones capable of realizing that a major economic crisis was imminent

Proverbs become proverbs because they succinctly state a profound truth, and no proverb better describes the state of neoclassical macroeconomics before the Great Recession than 'Pride goes before the fall.' The full proverb puts it even better: 'Pride goes before Destruction, and a Haughty Spirit before a Fall.' Before the 'Great Recession' (as the sudden economic downturn that began in 2007 is known in America), a popular '*topic du jour*' in the leading macroeconomic journals of the world (which are dominated by neoclassical economists) was explaining 'The Great Moderation'– the apparent decline in both the levels and volatility of unemployment and inflation since 1990. It was a trend they expected to see continue, and they were largely self-congratulatory as to why it had come about: it was a product of their successful management of the economy.

Few were more prominent in promulgating this view than Federal Reserve chairman Ben Bernanke. In 2004, while a member of the board of governors of the Reserve,[1] Bernanke gave a speech with precisely that title, in which he observed that:

> One of the most striking features of the economic landscape over the past twenty years or so has been a substantial decline in macroeconomic volatility […] the variability of quarterly growth in real output […] has declined by half since the mid-1980s, while the variability of quarterly inflation has declined by about two thirds. Several writers on the topic have dubbed this remarkable decline in the variability of both output and inflation 'the Great Moderation.' (Bernanke 2004b)

He nominated three possible causes of this phenomenon: 'structural change, improved macroeconomic policies, and good luck.' While he conceded that a definitive selection could not be made between the three factors, he argued that 'improved monetary policy' deserved more credit than it had received to date:

> improved monetary policy has likely made an important contribution not only to the reduced volatility of inflation (which is not particularly controversial) but to the reduced volatility of output as well. Moreover, because a change in the monetary policy regime has pervasive effects, I have suggested that some of the effects of improved monetary policies may have been misidentified as exogenous changes in economic structure or in the distribution of economic shocks. This conclusion on my part makes me optimistic for the future, because I am confident that monetary policymakers will not forget the lessons of the 1970s. (Ibid.)

Equally confident that neoclassical economics had delivered a better world was Nobel laureate Robert Lucas, who is one of the key architects of modern neoclassical macroeconomics. In his Presidential Address to the American Economic Association in 2003, he went even farther in his optimism than Bernanke, to assert that macroeconomic theory had made another depression impossible:

> Macroeconomics was born as a distinct field in the 1940's, as a part of the intellectual response to the Great Depression. The term then referred to the body of knowledge and expertise that we hoped would prevent the recurrence of that economic disaster. My thesis in this lecture is that macroeconomics in this original sense has succeeded: *Its central problem of depression prevention has been solved, for all practical purposes, and has in fact been solved for many decades*. (Lucas 2003: 1; emphasis added)

They had no idea of what was about to happen. And fundamentally, they had no one but themselves to blame for their ignorance.

The kernel

Macroeconomics, the study of the behavior of the entire economy, was once an area of economic research independent from microeconomics, the study of individual markets. However, working with a cavalier ignorance of the many flaws in microeconomics, economists reshaped macroeconomics, not to increase its relevance to the economy, but to make it a branch of microeconomics. Today, macroeconomics is based on propositions which have been shown to be untenable in the preceding chapters. This process of decay was set in train first by Keynes's incomplete escape from conventional theory at the time he wrote *The General Theory of Employment, Interest and Money*, and accelerated by Hicks's dubious interpretation of Keynes as a marginalist.

From Hicks's IS-LM (investment and savings–liquidity and money) model on, the road was cleared for the novelty in macroeconomics to be eliminated, and for the key conclusion of pre-Keynesian economics – that a market economy could not experience a depression – to be restored, just in time for the next depression to occur.

The roadmap

This is a complicated chapter, and not merely because the subject matter itself is difficult. An additional complication comes from the way in which the version of neoclassical theory taught to undergraduates is very different to that taught to students in PhD programs.

Undergraduate courses teach what is known as the IS-LM (and/or AS-AD, aggregate supply–aggregate demand) model of macroeconomics, which is presented as a précis of Keynes's theory, but in reality was devised by Keynes's contemporary and intellectual rival John Hicks. PhD students, on the other hand, learn a class of models that goes by the grandiose – and utterly misleading – name of 'dynamic stochastic general equilibrium' (DSGE) models.

Both IS-LM and DSGE models are derived from the microeconomic concepts that I have shown are fallacious in the preceding chapters. They

differ only in how extreme their reliance is on microeconomic theory, and on the presumption that everything happens in equilibrium. But they are also very different models, and therefore they have to be discussed independently, so in some ways there are two chapters in this one.

I precede these mini-chapters with a discussion of the fallacy they have in common: the belief that macroeconomics can and should be derived from microeconomic theory.

Then in the first mini-chapter I outline Keynes's critique of 'Say's Law,' the argument that 'supply creates its own demand,' and embellish it by comparing it to Marx's critique of the same proposition. I next argue that a key concept in all of neoclassical economics, 'Walras's Law,' is simply Say's Law in a more formal guise, and that it is false in a credit-driven economy. Keynes's and Marx's critiques of the conventional economics of their day are therefore still applicable to modern economics. Hicks's reinterpretation of Keynes as a 'marginalist' is debunked. Finally I detail Hicks's late realization that his interpretation of Keynes was untenable once uncertainty was taken into account as a key determinant of the level of investment.

In the second mini-chapter I cover the manner in which the DSGE approach to neoclassical macroeconomics overthrew the IS-LM model, and show that the key motivations for this were the desire to reduce macroeconomics to applied microeconomics, and to prove that there was a natural rate of unemployment that could not be altered by government policy. The inevitable intrusion of realism into this story led to the dominance of what is called the 'New Keynesian' faction of neoclassical macroeconomists. Neoclassical economists were confident that they had finally managed to reconcile Walras with Keynes, and this confidence made them optimistic about the economic future.

Then the Great Recession hit.

Macroeconomics and the reductionist fallacy

Humanity made great progress in understanding reality by ignoring the overwhelming complexity of the universe, and focusing on small components of it in isolation from each other. Compare, for example, the ancient vision of the physical world of consisting of four elemental factors,

earth, water, air and fire, to our understanding of the periodic table and the quantum mechanical factors beneath that today. We would not have got from the ancient view of the world to the modern without ignoring the overall complexity of the universe and focusing on individual components of it, in isolation from all others.

The success of this approach – known as 'reductionism' – once led to the belief that there was a hierarchical ranking of sciences, in which more complex areas were merely simplified manifestations of the underlying fundamental determinants. For example, the biological processes in living organisms were thought to be merely a surface manifestation of the underlying chemical processes, and they in turn were just surface manifestations of the quantum mechanics that ruled chemical interactions. This attitude, known as 'strong reductionism,' argued that, ultimately, all sciences could be reduced to physics.

This belief was best put by the man who first showed its true limits, Henri Poincaré:

> This conception was not without grandeur; it was seductive, and many among us have not finally renounced it; they know that one will attain the ultimate elements of things only by patiently disentangling the complicated skein that our senses give us; that it is necessary to advance step by step, neglecting no intermediary; that our fathers were wrong in wishing to skip stations; but they believe that when one shall have arrived at these ultimate elements, there again will be found the majestic simplicity of celestial mechanics. (Poincaré 1956 [1905]: 166)

In turn, strong reductionism implied that all large-scale systems could be understood by working up from the small-scale. In the case of economics, this implied that the behavior of the macroeconomy should be derived directly from microeconomics, and this belief indeed dominated the development of macroeconomic theory from shortly after the publication of Keynes's *General Theory*. Today, neoclassical macroeconomics truly is applied microeconomics.

In the physical sciences, a very different development occurred. Poincaré showed that there were limits to reductionism in 1899, when he

proved that, while a gravitational system with two celestial bodies (one sun and one planet) was utterly predictable, it was impossible to predict the behavior of a solar system with more than one planet. Reductionism still dominated the physical sciences for another seventy years, however, until these limits became apparent with the advent of the computer.

Before the computer, reductionism had a natural ally in the inability of researchers to analyze nonlinear relationships between variables. Strong reductionism implies that the behavior of any complex system can be entirely understood by considering the behavior of its constituents, and then summing their effects: 'the whole is the sum of the parts.' This belief was consistent with the limitations of linear algebra, which was relatively easy to do before computers.

Then the number-crunching power of computers enabled researchers to consider systems with nonlinear relations between variables – as with Lorenz's model of the weather, where two of the three variables are multiplied by each other in two of the three equations – and they consistently observed a remarkable result: in systems where variables interact in nonlinear ways, 'the whole is *more than* the sum of its parts,' and behaviors will occur at the aggregate level that cannot be found at the level of the system's elementary components. This phenomenon, the occurrence of behaviors at the aggregate level that could not be explained by behaviors at the component level, was christened 'emergent properties.'

Scientists then reconsidered the role of reductionism. It still had its place, but they were now aware of the fallacy in the belief that the best way to understand any systems was from the bottom up. In a paper tellingly entitled 'More is different,' the Physics Nobel laureate Philip Anderson called this fallacy 'constructionism.' It had two manifestations. First, even if a reductionist vision of a particular system was correct, the belief that the best way to understand the system was to construct it from its constituent parts was false:

> The main fallacy in this kind of thinking is that the reductionist hypothesis does not by any means imply a 'constructionist' one: The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. In fact, the more the elementary particle physicists tell

> us about the nature of the fundamental laws the less relevance they seem to have to the very real problems of the rest of science, much less to those of society. (Anderson 1972: 393)

The second was that larger systems turned out to have behaviors which were unique to their scale: *scale itself resulted in new behaviors which could not be deduced from the behavior of isolated components of a system*:

> The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other. (Ibid.: 393)

Anderson was willing to entertain the proposition that there was a hierarchy to science, so that: 'one may array the sciences roughly linearly in a hierarchy, according to the idea: "The elementary entities of science X obey the laws of science Y"' (Table 10.1).

But he rejected the idea that any science in the X column could simply be treated as the applied version of the relevant science in the Y column: 'But this hierarchy does not imply that science X is "just applied Y." At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as in the previous one. Psychology is not applied biology, nor is biology applied chemistry' (ibid.: 393).

TABLE 10.1 Anderson's ranking of sciences

X	Y
Solid state or many-body physics	Elementary particle physics
Chemistry	Many-body physics
Molecular biology	Chemistry
Cell biology	Molecular biology
...	...
Psychology	Physiology
Social sciences	Psychology

The physical sciences embraced this discovery of emergent behavior, and what was first dubbed 'chaos theory' (Li and Yorke 1975) and is now known as 'complexity theory' (May and Oster 1976) is a fertile aspect of research in fields as diverse as physics and biology.

Among neoclassical economists, however, the reductionist fallacy held sway, and this is nowhere more evident than in the deliberate reduction of macroeconomics to applied microeconomics *in the confident but false belief that this was possible*.

Ironically, despite its adherence to strong reductionism, neoclassical economics provides one of the best examples of emergent phenomena ever: the 'Sonnenschein-Mantel-Debreu conditions' that were discussed in Chapter 3. This research proved that a market demand curve derived from the preferences of individual consumers who in isolation obeyed the Law of Demand – i.e. they had 'downward-sloping demand curves' – will not itself obey the Law of Demand: a market demand curve can have any shape at all.[2]

This is emergence par excellence: a behavior which, under the assumptions of revealed preference, is provably absent from individual consumers – demand curves that can rise as well as fall when price increases – can occur at the level of single markets in a multiple-consumer, multiple-commodity economy.

The correct inference from that research is that, not only is macroeconomics not applied microeconomics, but even microeconomics itself can't be based on a simple extrapolation from the alleged behavior of individual consumers and firms. Thus, even within microeconomics, the study of markets cannot be reduced to the analysis of individual behaviors, while under no circumstances can macroeconomics be derived from microeconomics.

However, with some honorable exceptions (Kirman 1989, 1992), neoclassical economists resiled from this discovery of emergent properties within economics. The result was misinterpreted, and buried in poor pedagogy, so that three generations of post-WWII neoclassical economists continued to believe in the reductionist fallacy.

In this, they continued the behavior of their pre-WWII forebears. Ever since Adam Smith's *Wealth of Nations*, the dominant tendency in

economics has been to analyze the economy from the perspective of the behavior of individual rational agents, and to derive from this the inference that, so long as prices are flexible, there can be no macroeconomic problems. Thinkers who took a different perspective, such as Malthus in his debates with Ricardo, or Marx and other critics, were driven to the periphery of economics.

In the language of the nineteenth century, the mainstream economists of that time argued that there could be no 'general glut': while individual markets might have more supply than demand, in the aggregate there had to be other markets where there was more demand than supply. Therefore, while there could be problems in individual markets, the entire economy should always be in balance, because a deficiency in one market would be matched by an excess in another. All that would be required to correct the imbalance would be to let the market mechanism work, so that the price of the good with excess demand would rise while the one with excess supply would fall. Macroeconomics, as we call it today, was seen as unnecessary.

Prior to the 1870s, this belief that there could be no macroeconomic problem involved a strange mishmash of ideas, because the classical school of thought that dominated economics 'proved' the absence of macroeconomic problems by borrowing arguments from Jean Baptiste Say, who was effectively an early neoclassical. After the 1870s, there was no such disconnect, as the neoclassical revolution led by Menger, Walras and Marshall swept away the old classical school. Economists continued to be confident that there could never be a general glut, and this macroeconomic belief was now derived from a consistent microeconomic theory.

Then the Great Depression began. As unemployment relentlessly climbed to 25 percent of the American workforce, and fascism broke out in Europe, neoclassical economists of the day were in disarray. Into this breach stepped Keynes. With the publication of *The General Theory of Employment, Interest and Money* in 1936, Keynes effectively invented macroeconomics as a separate sub-discipline within economics.

From that point on, neoclassical economists attempted to undermine it.

Say, Walras, and the self-equilibrating economy …

The *General Theory* was conceived and published during capitalism's greatest slump, the Great Depression, when America's output fell by 30 percent in four years, stock prices fell by 90 percent, commodity prices fell by almost 10 percent a year in its first two years, and unemployment remained above 15 percent for a decade.

Prior to then, mainstream economists did not believe there were any intractable macroeconomic problems. Individual markets might be out of equilibrium at any one time – and this could include the market for labor or the market for money – but the overall economy, the sum of all those individual markets, was bound to be balanced.

The basis for this confidence was the widespread belief, among economists, in what Keynes termed Say's Law. As Keynes described it, this was the proposition that 'supply creates its own demand' (Keynes 1936). Some economists dispute Keynes's rendition of Say's Law (Kates 1998), and I concur that in several ways Keynes obscured what Say actually meant. So it is appropriate to turn to the horse's mouth for a definition:

> Every producer asks for money in exchange for his products, only for the purpose of employing that money again immediately in the purchase of another product; for we do not consume money, and it is not sought after in ordinary cases to conceal it: thus, when a producer desires to exchange his product for money, he may be considered as already asking for the merchandise which he proposes to buy with this money. It is thus that the producers, though they have all of them the air of demanding money for their goods, do in reality demand merchandise for their merchandise. (Say 1967 [1821])

Say's core proposition is that overall balance is assured because, to quote Steve Kates, the strongest modern-day proponent of Say's Law: '[t]he sale of goods and services to the market is the source of the income from which purchases are financed' (Kates 1998).

This, according to the 'classical' economists from whom Keynes hoped to distinguish himself,[3] meant that there could never be a slump due to an overall deficiency in demand. Instead, slumps, when they occurred, were due to sectoral imbalances.

If the demand for one market – such as labor – was too low relative to supply, this was because demand exceeded supply in one or more other markets. The solution was for sellers in the market suffering from excess supply – workers – to accept a lower price for their commodity.

Money was also treated as a commodity in the pre-Keynesian model, and it was possible that, at some point in time, many people would want to hold money and very few would want goods. There could then be a serious slump, as producers of goods found that people did not want to part with their money. Physical commodity markets and the labor market could then be in excess supply – with unsold goods and unemployed workers – but this would be because of the excess demand for money, and not because of any overall deficiency of aggregate demand. In the aggregate, demand and supply would be in balance.

Keynes's attempt to refute this notion was, to put it kindly, rather confusing, and on this basis alone I can to some extent understand the inability of many neoclassical economists to comprehend his theory. I reproduce Keynes's argument in its entirety in the next quote, and if you don't comprehend it completely on a first reading, don't worry – in fact, I'd worry if you did comprehend it! After you've waded through this, I'll provide a far clearer explanation that Keynes was aware of at the time he wrote the *General Theory*, but which – probably for political reasons – he chose not to use.

OK: take a good swig of coffee, a deep breath, and read on:

> This theory can be summed up in the following propositions:
>
> 1 In a given situation of technique, resources and costs, income (both money-income and real income) depends on the volume of employment N.
>
> 2 The relationship between the community's income and what it can be expected to spend on consumption, designated by D1, will depend on the psychological characteristic of the community, which we shall call its propensity to consume. That is to say, consumption will depend on the level of aggregate income and, therefore, on the level of employment N, except when there is some change in the propensity to consume.

3 The amount of labor N which the entrepreneurs decide to employ depends on the sum (D) of two quantities, namely D1, the amount which the community is expected to spend on consumption, and D2, the amount which it is expected to devote to new investment. D is what we have called above the effective demand.

4 Since D1 + D2 = D = f(N), where f is the aggregate supply function, and since, as we have seen in (2) above, D1 is a function of N, which we may write c(N), depending on the propensity to consume, it follows that f(N) – c(N) = D2.

5 Hence the volume of employment in equilibrium depends on (i) the aggregate supply function, (ii) the propensity to consume, and (iii) the volume of investment, D2. This is the essence of the *General Theory of Employment*.

6 For every value of N there is a corresponding marginal productivity of labor in the wage-goods industries; and it is this which determines the real wage. (5) is, therefore, subject to the condition that N cannot exceed the value which reduces the real wage to equality with the marginal disutility of labor. This means that not all changes in D are compatible with our temporary assumption that money-wages are constant. Thus it will be essential to a full statement of our theory to dispense with this assumption.

7 On the classical theory, according to which D = f(N) for all values of N, the volume of employment is in neutral equilibrium for all values of N less than its maximum value; so that the forces of competition between entrepreneurs may be expected to push it to this maximum value. Only at this point, on the classical theory, can there be stable equilibrium.

8 When employment increases, D1 will increase, but not by so much as D; since when our income increases our consumption increases also, but not by so much. The key to our practical problem is to be found in this psychological law. For it follows from this that the greater the volume of employment the greater will be the gap between the aggregate supply price (Z) of the corresponding output and the sum (D1) which the entrepreneurs can expect to get back out of the expenditure of consumers. Hence, if there is no change in the propensity to consume, employment cannot increase, unless at the same time D2 is increasing so as to fill the increasing gap between Z and D1. Thus – except on the

> special assumptions of the classical theory according to which there is some force in operation which, when employment increases, always causes D2 to increase sufficiently to fill the widening gap between Z and D1 – the economic system may find itself in stable equilibrium with N at a level below full employment, namely at the level given by the intersection of the aggregate demand function with the aggregate supply function. (Keynes 1936: 28–9; emphasis added)

You got that? Oh, come on, pull the other one! It's far more likely that your head is still spinning after reading that extract, and the same applied to the handful of leading neoclassical economists who read Keynes, and tried to work out how he could argue that aggregate demand could be deficient.

It's a tragedy that what Keynes himself described as 'the essence of the General Theory of Employment' was expressed in such a convoluted and turgid fashion – especially given his capacity for brilliant prose. It is therefore not altogether amazing that neoclassical economists believed that Keynes either misunderstood what they believed was the true concept at the heart of Say's Law, or that he intended to refute the clearly incorrect belief that overall balance meant that there could never be involuntary unemployment – whereas Say's Law allowed for involuntary unemployment as a by-product of sectoral imbalances.

The upshot is that the essence of Say's Law lives on in modern economics, though it now goes under the more respectable name of 'Walras's Law' (or, in some circles, 'Say's Principle'). Its modern definition is that 'the sum of all notional excess demands is zero,' and this proposition is accepted as valid – indeed as irrefutable – by modern-day economists.

However, I argue that this is precisely the concept which Keynes intended to refute, and that he was right to do so.

Say no more? The modern attempt to reconcile Keynes with Say and Walras (Leijonhufvud 1968; Clower and Leijonhufvud 1973)[4] starts from the proposition that, on the average, agents in a market economy are neither thieves (who want to take more than they give) nor philanthropists (who want to give more than they get). Therefore the normal agent will intend to have balanced supplies and demands: the value of what he wishes to sell

will equal the value of what he wishes to buy, so that 'the sum of his notional excess demands is zero.'

The excess demand for any single product by a single agent can be positive – so that the agent wishes to be a net buyer of that product – or negative – so that the agent wishes to be a net seller. However, in sum, his excess demands will be zero.

This balance at the level of the individual agent necessarily carries over to the aggregate of all agents: if the intended excess demands of each individual agent sum to zero, then the intended excess demands of all agents sum to zero.

However, this identity of aggregate supply and aggregate demand at the overall market level doesn't necessarily translate to identity at the level of each individual market. In particular, as noted earlier, it is possible for excess demand for money to be positive – in which case commodity markets would be 'glutted.' Excess demand for labor can also be negative – the supply of workers can exceed the demand for them – so that there will be involuntarily unemployed workers (and also a notional excess demand for the products that the unemployed workers intended to buy).

These two circumstances are both explanations of a depression. The former would involve a 'rising price' for money – or in other words 'deflation' as the money price of all other commodities fell. The latter would involve a falling price for labor – falling wages. However, both these forces would make a depression a temporary phenomenon. As Dixon puts it:

> [F]ollowers of Walras would say that involuntary unemployment cannot persist in a market economy with flexible wages and prices. They would argue that if the commodities market has excess demand then the prices of commodities will tend to rise and this will tend to reduce the level of excess demand in that market. In the labor market, where there is excess supply, they would assert that money wages will tend to fall. The joint effect of the rising price together with a falling money wage is that the real wage will tend to drop thus reducing (and eventually removing entirely) the excess supply in the labor market.

> As a consequence of the above, many would see the pronouncements of Keynes that the economy could find itself with an excess supply of labor and yet, in all (other) respects be in 'equilibrium,' as being in conflict with Walras' Law and therefore wrong or 'bad' in theory and so inadmissible. (Dixon 2000b)

Keynes's critique Let's now simplify Keynes's argument from that pivotal passage to see whether it's consistent with the way in which neoclassical economists later interpreted it. Keynes divided all output into two classes: consumption and investment. If the economy was in equilibrium, then Say's Law would argue that excess demand for consumption goods would be zero, and likewise for investment goods.

Keynes then imagined what would happen if demand for consumption goods fell, so that excess demand for consumption goods was negative (supply exceeded demand).[5] Say's Law would argue that demand for investment goods would rise to compensate: notional excess demand for investment goods would be positive.

However, as Keynes argued extensively throughout the *General Theory*, demand for investment goods is driven by expectations of profit, and these in turn depend heavily upon expected sales to consumers. A fall in consumer demand now could lead entrepreneurs to expect lower sales in the future – since in an uncertain environment 'the facts of the existing situation enter, in a sense disproportionately, into the formation of our long-term expectations; our usual practice being to take the existing situation and to project it into the future' (Keynes 1936: 148).

Dampened expectations would therefore lead entrepreneurs to reduce their demand for investment goods in response to a reduced demand for consumer goods. Thus a situation of negative excess demand for consumer goods could lead to a state of negative excess demand for investment goods too – a general slump.

This clearly contradicts Walras's Law. Since economists regard Walras's Law as irrefutable, this led some economists to ridicule Keynes's argument, and others to attempt to find how Keynes's argument could be reconciled with Walras's Law. The most widely accepted reconciliation was achieved by Robert Clower and Axel Leijonhufvud.

Say's Principle Clower and Leijonhufvud asserted that Keynes and Walras *were* compatible, because Walras's Law applied effectively only in equilibrium. Out of equilibrium, then, though the sum of *notional* excess demands was still zero, the sum of *effective* demands could be negative.

For example, if there was negative excess demand in the labor market – so that some workers were involuntarily unemployed – then it didn't help that these unemployed workers wanted to buy commodities. Without employment, their notional demands remained just that. Though they might want to buy commodities, without a wage their notional demand had no impact upon actual sales of commodities. Actual negative excess demand for labor might therefore not be balanced by actual positive excess demand for commodities, so that overall, the sum of excess demand could be negative. Keynes was vindicated as a disequilibrium theorist.[6] Keynes and Walras were reconciled.

But were they? Prior to the publication of the *General Theory*, Keynes indicated that he rejected the very basis of Walras's Law – the proposition that the sum of notional excess demands is zero – when he praised the author of what he had once described as an 'obsolete economic textbook which I know to be not only scientifically erroneous but without interest for the modern world' (Keynes 1925): Karl Marx.

The circuit of capital Marx's critique of Say's Law went to the heart of Walras's Law (and Say's Law). Marx rejected Say's initial proposition that '[e]very producer asks for money in exchange for his products, only for the purpose of employing that money again immediately in the purchase of another product' (Say 1967 [1821]). Instead, Marx pointed out that this notion asserted that no one in a market economy wished to accumulate wealth. If there was never any difference between the value of commodities someone desired to sell and buy on the market, then no one would ever desire to accumulate wealth. But an essential feature of capitalism is the existence of a group of agents with precisely that intention.

Believers in Say's Principle or Walras's Law might find these agents rather bizarre, since in their terms these agents are 'thieves,' who wish to take more than they give. However, far from being bizarre, these agents are an essential part of a market economy. They are known as capitalists. Far from their behavior being aberrant in a market economy, it is in fact the

essence of capitalism – and according to Marx, they do this without being thieves.

Whereas both Say's Law and Walras's Law assert that people simply desire to consume commodities, Marx asserted that an essential aspect of capitalism is the desire to accumulate. He derided Say's belief that the ultimate objective of every agent in a market economy was simply consumption – which is still generally accepted by economists today, as well as the economists of Marx's time – as an ideologically convenient but misleading fiction which obscures the actual dynamics of capitalism:

> It must never be forgotten, that in capitalist production what matters is not the immediate use-value but the exchange-value, and, in particular, the expansion of surplus-value. This is the driving motive of capitalist production, and it is a pretty conception that – in order to reason away the contradictions of capitalist production – abstracts from its very basis and depicts it as a production aiming at the direct satisfaction of the consumption of the producers. (Marx 1968 [1861]: ch. 17, section 6)

Capitalists are clearly fundamental to capitalism, and their behavior directly contradicts the Walras's and Say's Law presumption that every agent's intended excess demand is zero. As Marx put it:

> The capitalist throws less value in the form of money into the circulation than he draws out of it […] Since he functions […] as an industrial capitalist, his supply of commodity-value is always greater than his demand for it. If his supply and demand in this respect covered each other it would mean that his capital had not produced any surplus-value […] His aim is not to equalize his supply and demand, but to make the inequality between them […] as great as possible. (Marx 1885: ch. 4, section 'The meeting of demand and supply')

The dilemma for Marx was to explain how this inequality could be achieved without 'robbing' other participants in the market, and without

violating the principle that commodities were bought and sold at fair values. His solution points out the fallacy underlying the economist's superficially appealing arguments in Say's Law, Walras's Law, and Say's Principle.

This was that the market process had to include a production stage where the quantity and value of output exceeded the value of inputs – in Marx's terms and in Sraffa's (discussed in Chapter 6), a surplus is produced. The capitalist pays a fair price for his raw materials, and a fair wage to his employees. They are then combined in a production process which generates commodities for sale where the physical quantity of commodities and their monetary value exceed the quantity and value of inputs. The commodities are then sold for more than the cost of the raw materials and workers' wages, yielding a profit. The profit allows the capitalist to fulfill his desire to accumulate wealth, without robbing any other market participants, and without having to buy commodities below their value and sell them above it.[7]

Say's Law and Walras's Law, on the other hand, begin from the abstraction of an exchange-only economy: an economy in which goods exist at the outset, but where no production takes place (production is shoehorned into the analysis at a later point, but unsatisfactorily, as I outline below). The market simply enables the exchange of pre-existing goods. In such an economy, surplus in Marx's sense would be impossible. Equally, if one agent desired to and did accumulate wealth, that would necessarily involve theft in the Say's Principle sense. However, this condition does not hold when we move from the fiction of an exchange-only economy to the reality of a production and exchange economy. With production, it is possible for agents to desire to accumulate wealth without therefore aspiring to be thieves.

Marx formalized this analysis in terms of two 'circuits,' the 'Circuit of Commodities' and the 'Circuit of Capital.'

In the Circuit of Commodities, people come to market with commodities, which they exchange for money in order to buy other commodities. Marx stylized this as C – M – C:

$$\text{Commodity} \rightarrow \text{Money} \rightarrow \text{Commodity}$$

Though Marx discussed various ways in which this circuit could fail – owing primarily to delays between the sale of one commodity and the purchase of the next – generally speaking it obeys Walras's Law. Each 'agent' desires to convert commodities of a given value into different commodities of equivalent value.

However, in the Circuit of Capital, people came to market with money, with the intention of turning this money into more money. These agents buy commodities – specifically, labor and raw materials – with money, put these to work in a factory to produce other commodities, and then sell these commodities for (hopefully) more money, thus making a profit. Marx stylized this as M – C – M+:

> Money→Commodity→More money

The complete circuit, and the one which emphasizes the fallacy behind Walras's Law, was M – C(L, MP) … P … C+c – M+m:

> Money→Labor and means of production … Production … Different commodities, of greater value than paid for the labor and means of production→Sale of commodities to generate more money

This circuit specifically violates Say's Principle and Walras's Law. Rather than simply wanting to exchange one set of commodities for another of equivalent value, the agents in this circuit wish to complete it with more wealth than they started with. If we focus upon the commodity stages of this circuit, then, as Marx says, these agents wish to supply more than they demand, and to accumulate the difference as profit which adds to their wealth. Their supply is the commodities they produce for sale. Their demand is the inputs to production they purchase – the labor and raw materials. In Say's Principle's terms, the sum of these, their excess demand, is negative. When the two circuits are added together, the sum of all excess demands in a capitalist economy is likewise negative (prior to the introduction of credit, which we consider below).

This explanation of why Say's Law and Walras's Law don't apply to a market economy is far clearer than Keynes's, and the great pity is that

Keynes didn't use it in the *General Theory*, because it was in his 1933 draft. In this draft, Keynes observes that Marx made the 'pregnant observation' that:

> [T]he nature of production in the actual world is not C – M – C', i.e. of exchanging commodity (or effort) for money in order to obtain another commodity (or effort). That may be the standpoint of the private consumer. But it is not the attitude of business, which is a case of M – C – M', i.e., of parting with money for commodity (or effort) in order to obtain more money. (Dillard 1984: 424, citing Keynes's Collected Works, vol. 29, p. 81)

Keynes continued in a footnote that this vision of capitalism as having two circuits, one of which was motivated solely by the desire to accumulate wealth, in turn implied the likelihood of periodic crises when expectations of profit were not met:

> Marx, however, was approaching the intermediate truth when he added that the continuous excess of M' [over M] would be inevitably interrupted by a series of crises, gradually increasing in intensity, or entrepreneur bankruptcy and underemployment, during which, presumably M must be in excess. My own argument, if it is accepted, should at least serve to effect a reconciliation between the followers of Marx and those of Major Douglas, leaving the classical economics still high and dry in the belief that M and M' are always equal. (Ibid.: 424, citing Keynes's Collected Works, vol. 29, p. 82n.)

Unfortunately, Keynes later substituted his own convoluted reasoning for Marx's, I expect for two reasons. First, his argument was an attempt to put Marx's logic into the Marshallian framework in which Keynes was educated; secondly, he probably made a political judgment, at a time when Stalin's power was rising and communism had great political appeal, not to acknowledge the 'father of communism' in his critique of conventional economics.

Had Marx's clear logic been brought to center stage by Keynes, it is feasible that the 'neoclassical counter-revolution' initiated by Hicks might not have even commenced, because the fact that Keynes rejected Walras's Law, and his sound reasons for doing so, would have been so much clearer. So although Keynes's decision can be understood in the context of his times, with the benefit of hindsight, it was a serious mistake. Keynes's obscure and confusing argument allowed economists to continue believing that Walras's Law was an irrefutable truth. Only those working outside the neoclassical mainstream realized otherwise.

Credit and the fallacy of Walras's Law

Minsky, like Keynes before him, also omitted any reference to Marx in his own work, but his reasons for doing so are far easier to accept: given that he was an American academic during the McCarthyist period, any acknowledgment of Marx would have seriously impeded his academic career, if not ended it altogether.[8] However, he was strongly influenced by Marx's analysis, and took Marx's logic one step farther. He pointed out that since there is a buyer for every seller, and since accounting demands that expenditure must equal receipts, and yet growth also occurs over time, *then credit and debt must make up the gap*. Credit and debt are therefore fundamental to capitalism:

> If income is to grow, the financial markets, where the various plans to save and invest are reconciled, must generate an aggregate demand that, aside from brief intervals, is ever rising. For real aggregate demand to be increasing, […] it is necessary that current spending plans, summed over all sectors, be greater than current received income and that some market technique exist by which aggregate spending in excess of aggregate anticipated income can be financed. *It follows that over a period during which economic growth takes place, at least some sectors finance a part of their spending by emitting debt or selling assets*. (Minsky 1982 [1963]: 6; emphasis added)

Minsky's insight here points out the pivotal blind spot in thinking which leads neoclassical and Austrian economists to believe respectively in

Walras's Law and Say's Law: *they fail to consider the role of credit in a capitalist economy*.

Say, banker, can you spare a dime? Say's Law and Walras's Law envisage a world in which commodities are purchased only from the proceeds of selling other commodities, and in which commodities are the only things that are bought and sold. As Kates put it:

> According to the law of markets [Say's Law], aggregate demand was a conception unnecessary for a proper understanding of the cyclical behavior of economies. There were, of course, purchases and sales, and one could add together in some way everything bought during a period of time and describe this as aggregate demand […] [but] demand was not thought of as independent of supply. Instead, demand was constituted by supply; one could not demand without first having produced. Or to be more precise, demand in aggregate was made up of supplies in aggregate. (Kates 2003: 73–4)

In contrast to the position put by Kates, the world in which we live is one in which goods are purchased using both the proceeds of selling other goods and credit, while what is bought and sold includes existing assets as well as newly produced goods.

Aggregate demand is therefore aggregate supply *plus the change in debt*, while aggregate demand is expended on both commodities and assets (shares and property).[9] This guarantees the overall accounting balance that is an integral part of both Say's Law and Walras's Law, but it includes both the role of credit and the role of asset sales in a capitalist economy, which both of those 'laws' omit. Those 'laws' are thus relevant only to a world of either pure exchange or simple commodity production – the world that Marx characterizes as C→M→C – but are not relevant to the (normally) growing capitalist world in which we actually live.

The Say's Law/Walras's Law fallacy of ignoring the role of credit is the foundation of the neoclassical (and Austrian) argument that 'general gluts' and depressions are impossible, and that all crises are really sectoral imbalances which can be corrected by price adjustments alone. Once this

fallacy is removed, depressions or 'general gluts' (and general booms) are possible, and the contraction of credit plays a key role in them. But credit which is not backed by existing goods is also an essential feature of an expanding economy as well, as Schumpeter explains more clearly than either Minsky or Marx.

Schumpeter focused upon the role of entrepreneurs in capitalism, and made the point that an entrepreneur is someone with an idea but not necessarily the finance needed to put that idea into motion.[10] The entrepreneur therefore must borrow money to be able to purchase the goods and labor needed to turn his idea into a final product. This money, borrowed from a bank, adds to the demand for existing goods and services generated by the sale of those existing goods and services.

> The fundamental notion that the essence of economic development consists in a different employment of existing services of labor and land leads us to the statement that the carrying out of new combinations takes place through the withdrawal of services of labor and land from their previous employments […] this again leads us to two heresies: first to the heresy that money, and then to the second heresy that also other means of payment, perform an essential function, hence that processes in terms of means of payment are not merely reflexes of processes in terms of goods. In every possible strain, with rare unanimity, even with impatience and moral and intellectual indignation, a very long line of theorists have assured us of the opposite […]
>
> *From this it follows, therefore, that in real life total credit must be greater than it could be if there were only fully covered credit*. The credit structure projects not only beyond the existing gold basis, but also beyond the existing commodity basis. (Schumpeter 1934: 95, 101; emphasis added)

This Marx-Schumpeter-Minsky perspective thus integrates production, exchange and credit as holistic aspects of a capitalist economy, and therefore as essential elements of any theory of capitalism. Neoclassical economics, in contrast, can only analyze an exchange or simple commodity

production economy in which money is simply a means to make barter easier.

Say's Principle, which insists that the sum of all notional excess demands is zero, is a model of a capitalist economy without production and, most importantly, without capitalists.

Walrasian rejoinders?

There are a number of objections which economists could make to this Marx-Schumpeter-Minsky model of a monetary production economy.

First, Marx's circuits clearly cover not one market, but two: one when the capitalist buys his inputs, the other when he sells his outputs. Since these are two distinct markets in time, there is no reason, even under Walras's Law, why demands in one should equal supplies in the other. However, in each market, Walras's Law will apply.

Secondly, it is incorrect to conflate the exchange process with the production process. It is quite possible that agents could purchase inputs to production in one market, then combine them in production, produce a larger value of commodities and subsequently bring those commodities to sale at a subsequent market.

Thirdly, Marx's notion of a surplus implies that there are some commodities which can be purchased and, through production, turned into a larger value of commodities. This implies a 'free lunch.' If such a possibility ever existed, it would have long ago been 'arbitraged' away by the price of these commodities rising, or the price of the outputs falling.

Fourthly, Marx neglects the concept of a rate of time discount. Though some agents may appear to want to accumulate over time, if we discount future incomes to reflect the fact that the commodities that income will enable you to buy will be consumed in the future, then overall these agents are simply maintaining their level of satisfaction over time.

Taking the first and second hypothetical objections together, one of the strengths of Marx's approach is that his model covers a process through time, rather than merely considering an instant in time. In reality, at the aggregate level, exchange and production occur simultaneously. Factories are continuously producing commodities, sales rooms continually moving recently produced stock, workers are being paid wages, and spending them

on consumer goods. Marx's circuits analysis captures the organic nature of the production and exchange processes of a market economy, whereas the neoclassical approach artificially separates these into distinct stages.

This organic approach therefore enables Marx to consider the economy as a dynamic process, in which growth is an integral aspect of a capitalist economy. As part of this process, there are some agents who are continually accumulating wealth (when economic conditions are favorable), and others who are continually simply maintaining their level of economic well-being (though they can also gain in wealth if real wages are increasing, and if the wage exceeds subsistence).

Walras's Law, on the other hand, is best suited to the economic irrelevance of an exchange-only economy, or a production economy in which growth does not occur (which Marx called simple commodity production). If production and growth do occur, then they take place outside the market, when ironically the market is the main intellectual focus of neoclassical economics. Conventional economics is thus a theory which suits a static economy – and which can only be adapted to a dynamic economy with great difficulty, if at all – when what is needed are theories to analyze dynamic economies. Marx's 'through time' model of circuits is thus better suited to the analysis of a market economy than the 'moment in time' model of Walras's Law.

Marx's model of capitalist expectations is also far more valid than Walras's. A capitalist might well have his purchases and supplies balanced in any one market, as Walras's Law requires. However, purchases in this period are undertaken with the intention of selling a greater quantity in the next market. Marx's notion of the circuit of capital thus provides a link between one 'market period' and the next, which Walras's Law does not.

Since Say's Law and Walras's Law are in fact founded upon the hypothesized state of mind of each market participant at one instant in time, and since at any instant in time we can presume that a capitalist will desire to accumulate, then the very starting point of Say's/Walras's Law is invalid. In a capitalist economy, the sum of the intended excess demands at any one point in time will be negative, not zero. Marx's circuit thus more accurately states the intention of capitalists by its focus on the growth in wealth over time, than does Walras's Law's dynamically irrelevant and factually incorrect instantaneous static snapshot.

The arbitrage argument highlights the difference between neoclassical theory and Marx's theory of value,[11] which I discuss in more detail in Chapter 17 (where, I had better point out, I reject the 'labor theory of value' which conventional Marxists regard as integral to Marx's analysis). Neoclassical theory basically argues that the average rate of profit is driven down to the marginal productivity of capital, so that profit simply reflects the contribution which capital makes to output. This rate of profit is then called 'normal profit,' treated as a cost of production, and notionally set as the zero mark. Only profit above this level, called super-normal profit, is formally acknowledged in the theory, and in the pervasive theory of perfectly competitive equilibrium, super-normal profit is zero, so that profit fails to appear as a variable in economic theory.

The notion that profit is determined by the marginal product of capital was debunked in Chapter 7. Marx's theory of value, on the other hand, sees profit as a surplus of sales over the cost of production, allows for a positive rate of profit, and makes the rate of profit an integral part of the theory of production and exchange.

The time discount argument, that people are simply maintaining their level of satisfaction over time, has problems on at least two fronts. First, it is very hard to believe, for example, that Warren Buffett would feel that his level of wealth in 2011 was equivalent to his wealth in 1970. Successful capitalists would clearly feel that they have gained in wealth over time – and unsuccessful capitalists would definitely know that they have lost or at least failed to gain. Secondly, all this argument does is move the zero position when calculating whether someone is accumulating, staying the same, or losing out. If the normal rate of time discount is, say, 2 percent, then anyone who is accumulating wealth at more than 2 percent per annum is increasing their wealth – the sum of their time-discounted excess demands is negative.

So what?

The Walrasian argument that the sum of all excess demands is zero provides an apparent center of gravity for the economy. Rather like a seesaw, if one sector of the economy is down, then another sector is necessarily up. Furthermore, economists postulate that there are

countervailing forces at work: a 'down' sector will have the price of its output driven down, thus increasing demand and restoring balance, and vice versa for an 'up' sector. The seesaw will ultimately return to balance.

A negative sum for aggregate excess demand – and the requirement that this be made up for by borrowing money from banks – moves that center of gravity. Instead of the economy behaving like a seesaw where the pivot is carefully placed at the center of gravity, it behaves like one where the pivot is instead off-center, and can move abruptly one way or the other. A down sector is not necessarily offset by an up sector, so that, contrary to Walras's Law, the entire economy can remain down – or up – for an indefinite period.

In particular, a general slump is feasible. As Keynes argued, a decline in spending on consumption by consumers could lead investors to also reduce their demand for investment goods, so that the economy could remain in a situation of inadequate excess demand.

The key destabilizing force is investment. As both Keynes and Marx emphasize, investment is undertaken not for its own sake, but to yield a profit. If expectations of profit evaporate, then so too will investment spending, and the economy will be thrown into a general slump. Equally, if expectations of profit become too euphoric, investment can be overdone, and the economy can be thrown into an unsustainable boom – in that the profits expected by the investors will not be realized, and the boom will give way to bust. The non-Say's/Walras's Law vision of the economy shared by Marx, Schumpeter, Keynes and Minsky thus accords with the manifest instability of the macroeconomy, whereas Walras's Law asserts that, despite appearances to the contrary, the macroeconomy really is stable.

At the same time, this potential for instability is also a necessary aspect of the potential for growth. Instability, in and of itself, is not a bad thing, but in fact is fundamental to any dynamic, growing system. To extend the seesaw analogy, the fact that the real-world economic seesaw is not in equilibrium means that the only way to stop it tipping over is to keep the seesaw itself moving in the direction of its imbalance.[12] The neoclassical obsession with equilibrium is therefore a hindrance to understanding the forces that enable the economy to grow, when growth has always been a fundamental aspect of capitalism.[13]

Unfortunately, this perspective on Keynes's *General Theory* was buried beneath economists' mistaken belief that Walras's Law was incontrovertible. By forging a reconciliation between Keynes and Walras, the resulting 'Keynesian economics' was not Keynes's economics at all. It is little wonder that this Keynesian 'straw man' was so easily deconstructed by its conservative critics.

Say no more! Though Keynes unintentionally obscured Marx's critique of Say's Law, he also provided an eloquent explanation of why this shallow, simplistic notion held, and continues to hold, such a strong grip upon the minds of economists:

> That it reached conclusions quite different from what the ordinary uninstructed person would expect, added, I suppose, to its intellectual prestige. That its teaching, translated into practice, was austere and often unpalatable, lent it virtue. That it was adapted to carry a vast and consistent logical superstructure, gave it beauty. That it could explain much social injustice and apparent cruelty as an inevitable incident in the scheme of progress, and the attempt to change such things as likely on the whole to do more harm than good, commended it to authority. That it afforded a measure of justification to the free activities of the individual capitalist, attracted to it the support of the dominant social force behind authority. (Keynes 1936)

However, Keynes continued, this contrariness had, by the time of the Great Depression, led to a diminution of economics in the eyes of the public:

> But although the doctrine itself has remained unquestioned by orthodox economists up to a late date, its signal failure for purposes of scientific prediction has greatly impaired, in the course of time, the prestige of its practitioners. For professional economists, after Malthus, were apparently unmoved by the lack of correspondence between the results of their theory and the facts of observation – a discrepancy which the ordinary man has not failed to observe, with the result of his growing unwillingness

> to accord to economists that measure of respect which he gives to other groups of scientists whose theoretical results are confirmed by observation when they are applied to the facts. (Ibid.)

Despite Marx's and Keynes's critiques, Say's Law and Walras's Law lived on, and still dominate economic thinking today. The attempts by well-meaning economists like Clower and Leijonhufvud to reconcile Keynes with Walras's Law thus robbed Keynes of a vital component of his argument, making 'Keynesian economics' a severely emasculated version of Keynes's thought. But this was far from the only way in which Keynesian economics became a travesty of Keynes's original vision.

Hamlet without the prince

Rather as the Bible is for many Christians, the *General Theory* is the essential economics reference which few economists have ever read – including the vast majority of those who call themselves Keynesian economists.

There are many reasons for this.

One is that the *General Theory* is a difficult book. There are at least two roots to this difficulty. The good root is that Keynes was so much more insightful than most other economists that the concepts in the *General Theory* are difficult for more ordinary mortals to grasp; the bad root is that, as Keynes himself acknowledged, the book was replete with concepts from the very school of economics which he was hoping to overthrow. As cited previously, in the Preface Keynes observed that

> The composition of this book has been for the author a long struggle of escape, and so must the reading of it be for most readers if the author's assault upon them is to be successful – a struggle of escape from habitual modes of thought and expression. The ideas which are here expressed so laboriously are extremely simple and should be obvious. The difficulty lies, not in the new ideas, but in escaping from the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds. (Ibid.: xxiii)

The second and most important reason is this transcendental truth: neoclassical economists don't believe that macroeconomics should exist. The attitude of strong reductionism is so strong in neoclassical economics that the very existence of macroeconomics as an independent field of research within economics was an affront to them. Neoclassical economists could read the *General Theory* and find it incomprehensible, because the concepts they expect – utility-maximizing consumers, profit-maximizing producers, equilibrium and so on – are not the foundations of Keynes's thought.[14]

A final reason for not reading it is laziness: it is much easier to read the 'Reader's Digest' version given in a textbook than it is to slog through the unabridged original. As a result, many economists were inclined to rely upon summaries, rather than reading the original. Keynes obliged by providing his own summary, of just fifteen pages, in 1937.

Keynes and uncertainty The key concept in Keynes's summary was the impact of expectations upon investment, when those expectations were about what might happen in an uncertain future.

Investment is undertaken to augment wealth, and yet the outcome of any investment depends upon economic circumstances in the relatively distant future. Since the future cannot be known, investment is necessarily undertaken on the basis of expectations formed under uncertainty. Keynes was at pains to distinguish the concept of uncertainty from the simpler concept of risk.

Risk occurs when some future event can only be one of a number of already known alternatives, and when there is a known history of previous outcomes which enables us to assign a reliable and definite probability to each possible outcome. A dice roll is an example of risk. The dice can land only on one of six sides, and therefore only one of six numbers will turn up. If they are fair dice, each number has a 1 in 6 chance of turning up. The theory of probability can then be used to help predict the chances of various patterns of numbers occurring in future rolls of the dice.

Uncertainty is fundamentally different, and it has proved to be a difficult concept for economists before and after Keynes to grasp. Keynes gave several examples. Neither roulette, nor life expectancy, nor even the

weather qualified. Instead, uncertainty referred to such things as the chance that war might break out (this was in 1937, not long before Chamberlain's 'peace in our time' deal with Hitler), the rate of interest twenty years in the future, or when some invention would become obsolete. I gave a more positive and I hope evocative example of uncertainty in Chapter 8.

Probability theory cannot be used to help guide us in these circumstances because there is no prior history to go on, and because the outcomes are not constrained to any known finite set of possibilities. As Keynes put it, 'About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know' (Keynes 1937: 214).

Faced with this uncertainty, and yet compelled to act in spite of it, we develop conventions to help us cope. In marked contrast to his clumsy critique of Say's Law in the *General Theory*, Keynes, in explaining these conventions, was at his eloquent best:

> How do we manage in such circumstances to behave in a manner which saves our faces as rational, economic men? We have devised for the purpose a variety of techniques, of which much the most important are the three following:
>
> 1. We assume that the present is a much more serviceable guide to the future than a candid examination of past experience would show it to have been hitherto. In other words we largely ignore the prospect of future changes about the actual character of which we know nothing.
> 2. We assume that the existing state of opinion as expressed in prices and the character of existing output is based on a correct summing up of future prospects, so that we can accept it as such unless and until something new and relevant comes into the picture.
> 3. Knowing that our own individual judgment is worthless, we endeavor to fall back on the judgment of the rest of the world which is perhaps better informed. That is, we endeavor to conform with the behavior of the majority or the average. The psychology of a society of individuals each of whom is endeavoring to copy the others leads to what we may strictly term a conventional judgment. (Ibid.: 214)

Keynes notes that expectations formed in this manner are certain to be disappointed, but there is no other way in which to form them. Expectations are therefore bound to be fragile, since future circumstances almost inevitably turn out to be different from what we expected. This volatility in expectations will mean sudden shifts in investor (and speculator) sentiment, which will suddenly change the values placed on assets, to the detriment of anyone whose assets are held in non-liquid form.

As a consequence, money plays an essential role in a market economy because of its instant liquidity. The extent to which we desire to hold our wealth in the form of non-income-earning money, rather than income-earning but illiquid assets, 'is a barometer of the degree of our distrust of our own calculations and conventions concerning the future' (ibid.: 216).

This 'liquidity preference,' Keynes argued, determines the rate of interest: the less we trust our fragile expectations of the future, the higher the rate of interest has to be to entice us to sacrifice unprofitable but safe cash for potentially profitable but volatile assets.

In assets themselves investors face two broad alternatives: lending money at the prevailing rate of interest (effectively purchasing bonds), or buying shares which confer part-ownership of capital assets. Both these activities are effectively 'placement,' as Blatt (1983) put it, rather than investment proper, however, which is the building of new capital assets (Boyd and Blatt 1988).[15]

New capital assets are produced not for their own sake, but in expectation of profits, and profits will come in the form of capital gain if their market prices (the result of placement activity) exceed their costs of construction. Physical investment is, therefore, also extremely volatile because 'it depends on two sets of judgments about the future, neither of which rests on an adequate or secure foundation – on the propensity to hoard [the flip side of liquidity preference] and on opinions of the future yield of capital-assets' (Keynes 1937: 218). These two factors, which play a key role in determining how much investment takes place, are likely to feed upon and destabilize each other: if we become more pessimistic about the future prospects of investments, we are likely to want to hoard more, not less.

Having explained why expectations are so important in economic practice and economic theory, why uncertainty makes expectations so fragile and volatile, and how these factors affect the rate of interest and the level of investment, Keynes returned once more to an attack on Say's Law. He divided expenditure into consumption – which is relatively stable – and investment – which is highly volatile – and emphasized that investment is the key determinant of the level and rate of change of output (and hence employment). His theory was, therefore, a theory 'of why output and employment are so liable to fluctuation' (ibid.: 221). In contrast to the unintelligible summary in the *General Theory* itself, Keynes gave a relatively pithy summary in which expectations, investment and uncertainty had pivotal roles:

> The theory can be summed up by saying that, given the psychology of the public, the level of output and employment as a whole depends on the amount of investment. I put it in this way, not because this is the only factor on which aggregate output depends, but because it is usual in a complex system to regard as the causa causans that factor which is most prone to sudden and wide fluctuation.
>
> More comprehensively, aggregate output depends on the propensity to hoard, on the policy of the monetary authority as it affects the quantity of money, on the state of confidence concerning the prospective yield of capital-assets, on the propensity to spend and on the social factors which influence the level of the money-wage. But of these several factors it is those which determine the rate of investment which are most unreliable, since it is they which are influenced by our views of the future about which we know so little. (Ibid.: 221; emphasis added)

Keynes peppered this paper with observations about how conventional economics ignored the issue of uncertainty, and how expectations are formed under uncertainty, by simply assuming the problem away: 'I accuse the classical economic theory of being itself one of these pretty, polite techniques which tries to deal with the present by abstracting from the fact that we know very little about the future' (ibid.: 215).

Finally, in a departure from the *General Theory* of just a year earlier, Keynes criticized the concept of the 'marginal efficiency of capital' – the ratio of the yield of newly produced capital assets to their price. Whereas he used this concept extensively in the *General Theory*, here he argued that it is indeterminate, since the price of capital assets is so volatile, and there will be a different 'marginal efficiency of capital' for every different level of asset prices. Rather than being a determinant of investment, the 'marginal efficiency of capital' might simply be a by-product of the level of investment and current expectations.

These are all difficult concepts, especially for economists who were bred in the neoclassical tradition in which 'at any given time facts and expectations were assumed to be given in a definite and calculable form; and risks, of which, though admitted, not much notice was taken, were supposed to be capable of an exact actuarial computation' (ibid.: 213).

But if Keynes had truly unleashed a revolution in economic thought, then economists should have attempted to get their minds around these difficult concepts, and fought to escape from the 'habitual modes of thought and expression' which had gripped them prior to the calamity of the Great Depression.

Did economists do so? Some did, but the majority did not – and for that reason the profession bifurcated into two camps: a minority which swore fealty to Keynes's revolutionary vision (who generally call themselves 'post-Keynesian'), and a majority which paid lip-service to some of Keynes's words, but which rapidly fell back into old, familiar ways. These economists ignored Keynes's *General Theory*, and even his pithy summary, instead clutching at another alleged summary by one J. R. Hicks.

Slimming Keynes down to size: the IS-LM model Hicks's 'Mr. Keynes and the Classics' purported to be a book review of the *General Theory*. Hicks began by disputing that neoclassical economists held quite the views Keynes alleged that they held, and therefore tried to construct a more typical classical theory 'in a form similar to that in which Mr. Keynes sets out his own theory […] Thus I assume that I am dealing with a short period in which the quantity of physical equipment of all kinds available can be taken as fixed' (Hicks 1937: 148).

Was this really the manner in which Keynes set out his own theory? Not according to Keynes, who criticized 'the classics' (by which he meant what we today call neoclassical economists) for working with a model in which 'the amount of factors employed was given' (Keynes 1937: 212). Nonetheless, Hicks continued. He summarized the 'typical Classical theory' in three equations, which argued that:

- the amount of money determined total output (output was some constant times the money stock);
- the rate of interest determined the level of investment; and
- the rate of interest determined the level of savings (and savings equaled investment).[16]

The first equation determines total output and total employment,[17] while the second two simply determine how much of output is devoted to investment, and how much to current consumption. If the savings rate increases, then so does investment. Increasing money wages 'will necessarily diminish employment and raise real wages' (Hicks 1937), while the obverse policy – cutting money wages – will necessarily increase employment and reduce the real wage. Decreasing the money supply directly decreases income and employment, and is the main explanation for economic downturns (an argument which Milton Friedman later revived).

Clearly, Keynes's theory was substantially different from this. But how did Hicks summarize Keynes? In three more equations, where:

- the demand for money depends upon the rate of interest (in place of the 'classical' fixed relationship between money and output);
- investment is a function of the rate of interest; and
- savings is a function of income.

Hello? What happened to uncertainty, expectations, liquidity preference determining the rate of interest, speculative capital asset prices, and so on? They are nowhere to be seen. Sometime later, Hyman Minsky commented

that ‘Keynes without uncertainty is rather like Hamlet without the Prince’ (Minsky 1975: 75), but this is what Hicks served up as Keynes. Even the *Reader’s Digest* would draw the line at this level of abridging, but this was not the end of Hicks’s rephrasing of Keynes’s Shakespearean sonnets into schoolyard doggerel.

He next argued that ‘Keynes’s’ first equation omitted the impact of income on demand for money. This was the traditional ‘transactions demand for money’ argument that some level of money was needed to finance everyday transactions, so that an increase in income would generate an increase in the demand for money. To be truly general, said Hicks, the ‘general theory’ should include the impact of income on the demand for money, as well as the impact of the rate of interest.

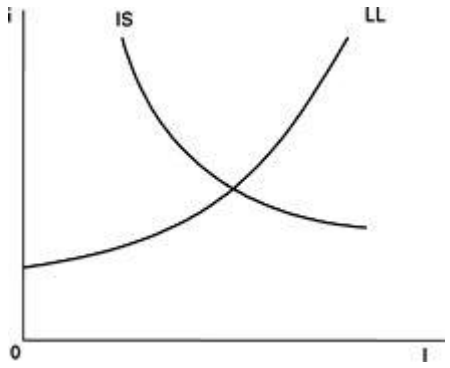
Keynes had omitted discussion of the transactions demand for money because this demand was relatively stable, and therefore less important than the more important demand set by liquidity preference. But Hicks believed that ‘however much stress we lay upon the “speculative motive,” the “transactions motive” must always come in as well’ (Hicks 1937: 153). So he proposed a revised set of equations in which the demand for money depends upon two variables – the rate of interest and the level of income – though not, as Keynes had it, on ‘the degree of our distrust of our own calculations and conventions concerning the future’ (Keynes 1937: 216).

With this revision, Keynes, who was at such pains to distinguish himself from his predecessors – primarily though not exclusively on the basis of the importance he attached to uncertainty and expectations – is pushed back into the camp from which he desired to escape. As Hicks put it: ‘With this revision, Mr. Keynes takes a big step back to Marshallian orthodoxy, and his theory becomes hard to distinguish from the revised and qualified Marshallian theories, which, as we have seen, are not new. Is there really any difference between them, or is the whole thing a sham fight? Let us have recourse to a diagram’ (Hicks 1937: 153).

Hicks’s diagram explains why his rendition of Keynes was so readily accepted by economists, while Keynes’s own summary was ignored (see Figure 10.1). It was the old familiar totem of two intersecting curves, though now relabeled to reflect its somewhat different derivation: in place of ‘S’ and ‘D’ for supply and demand, we now had ‘IS’ and ‘LM.’ The ‘Totem of the Micro,’ as Leijonhufvud satirized the supply and demand

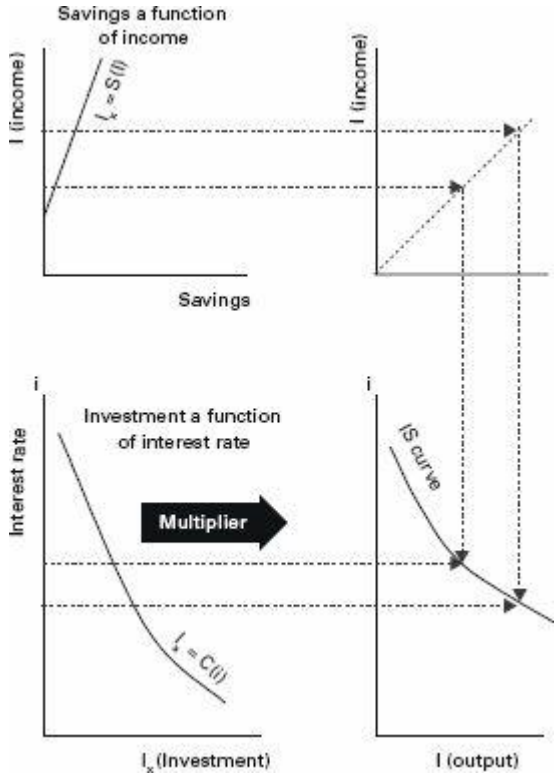
diagrams of Marshallian microeconomics, now had a bigger sibling for macroeconomics – though it was not derived in a way that microeconomists would accept, nor did it reach conclusions about the macro economy with which they would agree, as we shall see later.

The downward-sloping curve, the equivalent of the microeconomic demand curve, was derived from the investment and savings relations in Hicks's model. The upward-sloping curve, the equivalent of the microeconomic supply curve, was derived from the money demand relation (on the assumption that the money supply was controlled by the monetary authorities, and was therefore determined outside the model).

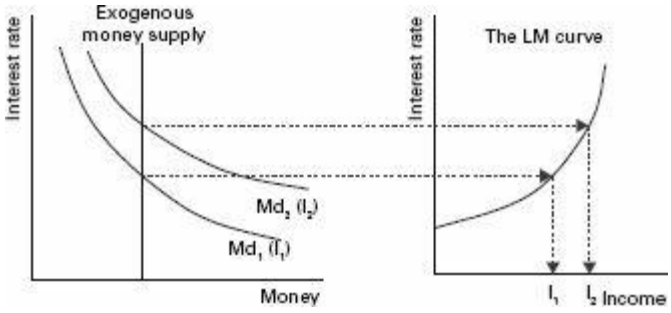


10.1 Hicks's model of keynes

The IS curve showed all those combinations of the rate of interest (i) and the level of income (I) which yielded equilibrium in the goods market. The LL curve (which economists today call the LM curve) showed all those combinations of the rate of interest and the level of income which gave equilibrium in the money market.



10.2 Derivation of the downward-sloping IS curve



10.3 Derivation of the upward-sloping LM curve

Here, at last, in comparison to the strange concepts of Keynes, economists were back on familiar ground. As Hicks put it:

> Income and the rate of interest are now determined together at P, the point of intersection of the curves LL and IS. They are determined together; just as price and output are determined together in the modern theory of demand and supply. Indeed, Mr. Keynes's innovation is closely parallel, in this respect, to the innovation of the marginalists. (Ibid.)

One problem with this 'general theory,' however, was that many of Keynes's conclusions could not be derived from it – something which would not have surprised Keynes a great deal, since this model omitted his key concepts of uncertainty and expectations. But Hicks had an apparent dilemma:

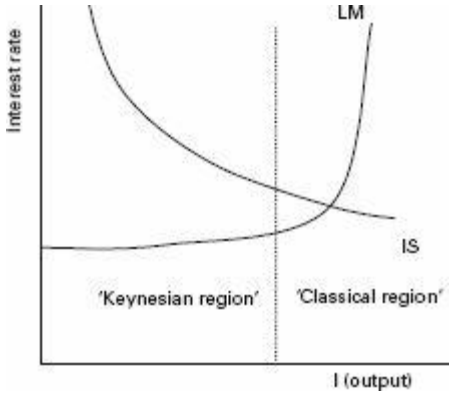
> But if this is the real 'General Theory,' how does Mr. Keynes come to make his remarks about an increase in the inducement to invest not raising the rate of interest? It would appear from our diagram that a rise in the marginal-efficiency-of-capital schedule must raise the curve IS; and, therefore, although it will raise income and employment, it will also raise the rate of interest. (Ibid.: 154)

To Keynes, the reason why an increased desire to invest would not necessarily raise the rate of interest is because the latter was determined by liquidity preference, which 'is a barometer of the degree of our distrust of our own calculations and conventions concerning the future.' In a depressed economy, an increase in investment could well reduce the 'degree of distrust,' leading to a fall in the rate of interest rather than a rise. But with Hicks's picture of Keynes shorn of uncertainty, conventions and expectations, there were no such mechanisms to draw upon. Fortunately, Hicks's model provided a simple and far more conventional solution: simply bend the curves:

> This brings us to what, from many points of view, is the most important thing in Mr. Keynes's book. It is not only possible to show that a given supply of money determines a certain relation between Income and interest (which we have expressed by the curve LL); it is also possible to say something about the shape of the curve. It will probably tend to be nearly horizontal on the left, and nearly vertical on the right. This is because there is (1) some minimum below which the rate of interest is unlikely to go, and (though Mr. Keynes does not stress this) there is (2) a maximum to the level of income which can possibly be financed with a

> given amount of money. If we like we can think of the curve as approaching these limits asymptotically. (Ibid.)

This 'liquidity trap' enabled Hicks to provide an explanation for the Great Depression, and simultaneously reconcile Keynes with 'the Classics.' Keynes was consigned to one end of the LM curve, where the liquidity trap applied, and 'the Classics' to the other, where full employment was the rule (see Figure 3.1). In the 'classical' range of the LM curve, conventional economics reigned supreme: there was a maximal, full employment level of income, where any attempts to increase output would simply cause a rising interest rate (or inflation, in extensions of the IS-LM model). In the 'Keynesian' region, monetary policy (which moved the LM curve) was ineffective, because the LM curve was effectively horizontal, but fiscal policy (which moved the IS curve) could generate greater output – and hence employment – without increasing interest rates. A higher level of government expenditure could shift the IS curve to the right, thus moving the point of intersection of the IS and LM curves to the right and raising the equilibrium level of output.



10.4 'Reconciling' Keynes with 'the Classics'

Hicks put the position pithily. In the 'Keynesian region' of his model, a depression can ensue because traditional monetary policy is ineffective – but Keynes's prescription of fiscal policy can save the day: 'So the General Theory of Employment is the economics of Depression' (ibid.: 155).

Hicks next proposed that, for reasons of mathematical elegance rather than economic relevance, all three variables (demand for money, investment

and savings) should be made functions of both income and the rate of interest (though not uncertainty or expectations):

> In order to elucidate the relation between Mr. Keynes and the 'Classics,' we have invented a little apparatus. It does not appear that we have exhausted the uses of that apparatus, so let us conclude by giving it a little run on its own.
>
> With that apparatus at our disposal, we are no longer obliged to make certain simplifications which Mr. Keynes makes in his exposition. We can reinsert the missing i in the third equation, and allow for any possible effect of the rate of interest upon saving; and, what is much more important, we can call in question the sole dependence of investment upon the rate of interest, which looks rather suspicious in the second equation. Mathematical elegance would suggest that we ought to have I and i in all three equations, if the theory is to be really General. (Ibid.: 156)

Economists, having been threatened by Keynes with the need to completely retrain themselves, could now engage in their favorite game of tobogganing up and down one curve, moving another to the left or right, just as they did in microeconomics. It is little wonder that this Hicksian IS-LM model was adopted as the basis for 'Keynesian' economics, and equally little wonder that, many years later, macroeconomics was converted to a subset of microeconomics.

The true origins of IS-LM Though 'cutting-edge' economic analysis has left Hicks's model behind, most macroeconomists still think in IS-LM terms, and this model is still the common fodder served up to undergraduate students as Keynesian economics. It therefore still has pedagogic and disciplinary relevance. So the question arises: from where did this model emanate? It clearly was not derived from Keynes's *General Theory*, apart from the adoption of some of Keynes's terminology. The mystery of its origins was finally solved by one Sir John Hicks – an older, be-nighted and somewhat wiser J. R. Hicks.

ISLM: an apology Hicks's detective work was published in a paper entitled 'IS-LM: an explanation,'[18] but in many ways it was an apology. Published in the non-orthodox *Journal of Post Keynesian Economics* in 1980, the paper's opening sentence was: 'The IS-LM diagram, which is widely, though not universally, accepted as a convenient synopsis of Keynesian theory, is a thing for which I cannot deny that I have some responsibility' (Hicks 1981: 141).

Even after this rueful opening, Hicks clung to a very Walrasian vision of Keynes, and elsewhere he described the IS-LM diagram as 'a product of my Walrasianism' (Hicks 1979: 990). But he conceded that his rendition had erroneously omitted any discussion of uncertainty or expectations. His explanation as to how he could have missed so fundamental an aspect of Keynes's thought was that, shortly before the *General Theory* was published, he had published a paper which, he believed, had strong similarities to Keynes's argument (Hicks 1935).[19] What he then published as a review of Keynes was actually a restatement of his own model, using some of Keynes's terminology.

Hicks saw two key problems in cross-dressing as Keynes. The first was that his model was 'a flexprice model […] while in Keynes's the level of money wages (at least) was exogenously determined' (Hicks 1981: 141);[20] the second 'more fundamental' problem was that Hicks's model used a period of a single week, while Keynes used 'a "short-period," a term with connotations derived from Marshall; we shall not go far wrong if we think of it as a year' (Hicks 1980).

Discussing the second problem, Hicks argued that the difference in period length had a drastic impact upon the relevance of expectations. With a time period of just a week, it is not unreasonable to keep expectations constant – and therefore to ignore them. But keeping expectations constant over a year in an IS-LM model does not make sense, because 'for the purpose of generating an LM curve, which is to represent liquidity preference, it will not do without amendment. For there is no sense in liquidity, unless expectations are uncertain' (Hicks 1981: 152).

This was precisely the point Keynes himself made, in ironic form, in 1937:

> Money […] is a store of wealth. So we are told, without a smile on the face. But in the world of the classical economy, what an insane use to which to put it! For it is a recognized characteristic of money as a store of wealth that it is barren; whereas practically every other form of storing wealth yields some interest or profit. Why should anyone outside a lunatic asylum wish to use money as a store of wealth?
>
> Because, partly on reasonable and partly on instinctive grounds, our desire to hold money as a store of wealth is a barometer of the degree of our distrust of our own calculations and conventions concerning the future […] The possession of actual money lulls our disquietude; and the premium which we require to make us part with money is the measure of the degree of our disquietude. (Keynes 1937: 215–16)

Thus, without uncertain expectations, there is no sense in liquidity preference, and Hicks cannot justify the LM half of his IS-LM model. But with uncertain expectations, there is no sense in equilibrium analysis either, since equilibrium can be maintained only if expectations are continually being fulfilled. Hicks concluded that the equilibrium/constant expectations framework of the IS-LM model was theoretically unsound, and practically irrelevant to the problems of the macroeconomy:

> I accordingly conclude that the only way in which IS-LM analysis usefully survives – as anything more than a classroom gadget, to be superseded, later on, by something better – is in application to a particular kind of causal analysis, where the use of equilibrium methods, even a drastic use of equilibrium methods, is not inappropriate […]
>
> When one turns to questions of policy, looking towards the future instead of the past, the use of equilibrium methods is still more suspect. For one cannot prescribe policy without considering at least the possibility that policy may be changed. There can be no change of policy if everything is to go on as expected – if the economy is to remain in what (however

> approximately) may be regarded as its existing equilibrium. (Hicks 1981)

There is one more, crucial weakness in Hicks's model that he touched upon but did not consider properly, and which would invalidate his model even if an LM curve could be derived: his use of 'Walras's Law' to reduce the number of markets in the model from three to two: 'Keynes had three elements in his theory: the marginal efficiency of capital, the consumption function, and liquidity preference. The market for goods, the market for bonds, and the market for money […]' (ibid.: 142).

He then explained that he dropped the second of these markets by applying Walras's Law: 'One did not have to bother about the market for "loanable funds," since it appeared, on the Walras analogy, that if these two "markets" were in equilibrium, the third must be also. So I concluded that the intersection of IS and LM determined the equilibrium of the system as a whole' (ibid.: 142).

Next he noted that there was in fact one other market that should be part of the model: the labor market – which was, of course, an integral part of Keynes's analysis in the *General Theory* itself: 'In strictness, we now need four markets, since labor and goods will have to be distinguished […]' (ibid.: 142–3).

He went on to argue that its omission was justified, but here his neoclassical fixation with equilibrium analysis led him astray, because – ignoring for the moment that Walras's Law is false in a capitalist economy – Walras's Law allows you to drop one market only when all other markets are in equilibrium.[21] In the IS-LM model, this applies only where the two curves cross: where the combination of GDP and the interest rate is such that both the goods market (the IS curve) and the money market (the LM curve) are in equilibrium. Then, in a three-market model – goods, money, and labor – if the money and goods markets are in equilibrium, then so too must be the labor market.

However, if the combination of the interest rate and the GDP are such that one of these two markets is out of equilibrium, then so too must be the labor market. Therefore only in equilibrium can the labor market be ignored. At any other location, the labor market must also be considered –

and therefore the IS-LM model is incomplete. Everywhere except at the point of intersection of IS and LM, it needs to be the IS-LM-'LSLD' model (where 'LS' and 'LD' refer to labor supply and labor demand respectively).

Furthermore, since at anywhere except the intersection of IS and LM at least one of those two markets is in disequilibrium, the third, ignored 'LSLD' market must also be in disequilibrium: wages must be higher (or lower) than the level that will clear the labor market. Therefore price-setting in this market – and the other one that is in disequilibrium – must be a dynamic, disequilibrium process, not a simple calculation of the equilibrium wage. Even Hicks's emasculated version of Keynes's macroeconomics must employ dynamic, disequilibrium analysis, in contrast to the comparative static mode in which the IS-LM model is normally applied.

This IS-LM model is thus invalid, even on its own terms, if it is pushed anywhere beyond working out what rate of interest and GDP combination represent equilibrium in the economy. To be used as a model of economic dynamics, it must become a three-equation model, and these must all be disequilibrium equations. This is not how IS-LM is taught, or used.[22]

But in its heyday, the IS-LM model gave economists something they had never really had previously: a framework on which to build models that were not merely drawings, or symbolic equations, but numerical equations that they could use to predict the future course of the economy.

The age of large-scale econometric models

Hicks's model and the later development of the 'Aggregate Supply-Aggregate Demand' model set off the heyday of attempts by economists to turn these models into numerical simulations of the economy, using the newly developed tool of the computer.

With a careful choice of parameter values, these models could generate a reasonable fit between the inputs ('exogenous variables') and the variables like future output and employment levels ('endogenous variables'). If a model's fit to the data wasn't too good, it could be improved by fine-tuning the parameters, or adding more variables, and as a result most of these models 'grew like Topsy.' One of the earliest such model, developed by Lawrence Klein (Klein 1950; Renfro 2009), had just six

equations; eventually models with thousands of equations were developed – and many are still in use.

There were five aspects of these models that made them easy to simulate, but which also made them fundamentally unsuited for economic analysis.

First, the models were frequently linear – variables in the equations were multiplied by constants, and added together to produce predictions for other variables – in contrast to the nonlinear models outlined in Chapter 9, so they couldn't develop interactions between variables that caused cyclical behavior, let alone complex behavior as in Lorenz's weather model.

Secondly, even when nonlinearities existed – when employment was divided by population to calculate an employment rate, for example, or when logarithms of variables were used rather than the raw variables – the model was solved as if these nonlinearities did not affect the system's tendency towards equilibrium (McCullough and Renfro 2000). Simulations therefore worked out what the equilibrium of the model would be, and their predictions had the economy converging to this point over time (see, for example, Renfro 2009: 46).[23]

Thirdly, the models effectively assumed that the economy's dynamics involved movements from one equilibrium position to another, with movement being caused by 'exogenous shocks' – events external to the economy (such as damaging floods or unexpected bountiful harvests). This continued the convention in econometrics of seeing fluctuations in the economy having non-economic causes: 'The majority of the economic oscillations which we encounter […] seem to be explained by the fact that certain exterior impulses hit the economic system and thereby initiate more or less regular oscillations' (Frisch 1933: 171 [1]).[24] Even though this argument was made during the Great Depression, where no 'external impulse' could be blamed for the crisis, and when economists like Schumpeter, Keynes and Fisher were arguing that cycles and possibly breakdowns were endemic to capitalism, this belief became the standard view in numerical simulations of the economy.[25]

Fourthly, they were based on a neoclassical vision of the economy, and therefore omitted the credit and debt variables that we now know are crucial to macroeconomics.

Finally, they omitted any consideration of how expectations are formed under pervasive uncertainty, a key aspect of Keynes's vision of the macroeconomy that was lacking in the parent IS-LM model.

There were therefore many good grounds on which these models could have been criticized. However, the one focused on by economists was something entirely different: they objected to these numerical models simply because, as with Hicks's stylized IS-LM model from which they were derived, they argued that there could be involuntary unemployment, and that the level of unemployment could be affected by government demand-management policies – conclusions which neoclassical economists mistakenly believed contradicted neoclassical microeconomics.

From IS-LM to the representative agent

Hicks's critical epiphany about the IS-LM model came far too late to stop the revisionist juggernaut he had set in motion by reinterpreting Keynes as a Walrasian back in 1937. His recantation in 1981 was generally ignored by economists, who – if they were aware of it at all – would have been more inclined to put his views down to approaching senility than to any blinding logical revelation. In any case, the gradual demolition of IS-LM by economists was substantially advanced by 1980.

This demolition began back in the 1950s with the 'strong reductionist' critique that Hicks's 'Keynesian' model did not have good microeconomic foundations, by which neoclassical economists meant that it was not possible to derive results that IS-LM could generate – such as the economy settling into a less than full-employment equilibrium – from standard microeconomics.

Of course, in making this critique they were profoundly ignorant of the aggregation errors in the theory itself that I have outlined in preceding chapters. Properly understood, it is possible to derive results like involuntary unemployment from a neoclassical model. A properly derived market demand curve can have any shape at all (Chapter 3), leading to a market marginal revenue curve that would therefore intersect the constant or falling marginal cost curve (Chapters 4 and 5) in its market in multiple locations. Complexities in distribution and production covered in Chapters 6 and 7 would complicate the outcome even further, while price-setting

would have to be done in dynamic disequilibrium, raising the specter of nonlinear dynamics and chaos (Chapter 9).

A macroeconomic model derived properly from neoclassical foundations would probably be more chaotic than the real world itself, even without introducing the complications the neoclassical model omits by improperly excluding money and debt from its analysis.[26]

However, all this was unknown to the neoclassicals, who continually chipped away at the IS-LM model and its cousin the AS-AD model ('Aggregate Supply-Aggregate Demand'), and even to the many defenders of these models. Non-orthodox economists, who were aware of these issues, watched on in bemused horror as a model that was already a bastardization of Keynes's analysis[27] was further emasculated over time. The extent to which this was an agenda driven by ignorance rather than wisdom can be seen in the memoir of Nobel laureate Robert Lucas, one of the key actors in this process, when he delivered the keynote address to the 2003 History of Political Economy conference.

He began by asserting stridently that he was once a Keynesian:

> My credentials? Was I a Keynesian myself? Absolutely. And does my Chicago training disqualify me for that? No, not at all […] Our Keynesian credentials, if we wanted to claim them, were as good as could be obtained in any graduate school in the country in 1963.

Then he continued:

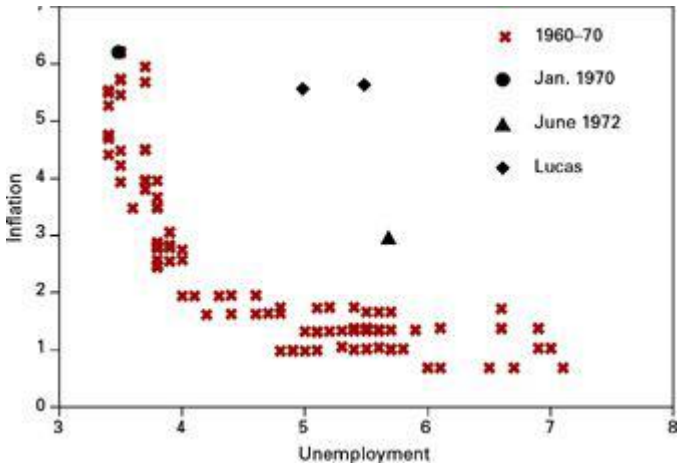
> I thought when I was trying to prepare some notes for this talk that people attending the conference might be arguing about Axel Leijonhufvud's thesis that IS-LM was a distortion of Keynes, but I didn't really hear any of this in the discussions this afternoon. So I'm going to think about IS-LM and Keynesian economics as being synonyms.
>
> I remember when Leijonhufvud's book came out and I asked my colleague Gary Becker if he thought Hicks had got the *General Theory* right with his IS-LM diagram. Gary said, 'Well, I don't know, but I hope he did, because if it wasn't for Hicks I never would have made any sense out of that damn book.' That's

> kind of the way I feel, too, so I'm hoping Hicks got it right. (Lucas 2004: 13–14; emphases added)

This was over twenty years after Hicks himself said that he had got it wrong! And Lucas had the hide to call himself a Keynesian, when he admits that 'if it wasn't for Hicks,' both he and fellow Nobel laureate Gary Becker 'never would have made any sense out of that damn book'? This is one reason I bridle when I hear the comment that 'Keynesian economics has failed'; what most self-described Keynesians in economics mean by the word 'Keynesian' is the economics of Hicks and Samuelson, not Keynes.

Starting from the false belief that Hicks had accurately summarized Keynes, Lucas then conformed to the unfortunate rule within economics, that poor scholarship is built upon poor scholarship. He played a crucial role in undermining IS-LM analysis itself in the early 1970s, first with the development of 'rational expectations macroeconomics' and then with what became known as 'the Lucas critique' – an attack on using numerical macroeconomic models as a guide to policy. These developments led to the final overthrow of any aspect of Hicksian, let alone 'Keynesian,' thought from mainstream macroeconomics. In the ultimate fulfillment of the program of strong reductionism, macroeconomics was reduced to no more than applied microeconomics – and based on the premise that all the concepts that I have shown to be false in the preceding chapters were instead true.

Lucas's assault on IS-LM[28] With Hicks's IS-LM model accepted as providing a mathematical expression of Keynes, Lucas (Lucas 1972) focused on models that economists had constructed using Hicks's model as a foundation, which concluded that macroeconomic policy could alter the level of economic activity. He began by conceding that most economists believed that the 'Phillips Curve' accurately described the 'trade-off' society faced between inflation and unemployment. He also conceded that the statistical evidence certainly showed a negative relationship between inflation and unemployment: when: 'It is an observed fact that, in U.S. time series, inflation rates and unemployment are negatively correlated' (ibid.: 50).



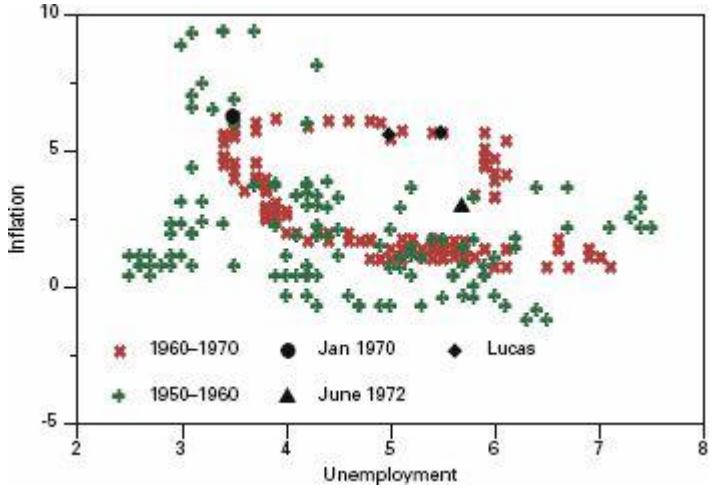
10.5 Unemployment-inflation data in the USA, 1960–70

The 'Phillips Curve trade-off' interpretation of these statistics turned an empirical regularity into a guide for policy. Since the statistics implied that unemployment and inflation moved in opposite directions, it seemed that the government could choose the level of employment it wanted by manipulating aggregate demand (so long as it was willing to tolerate the inflation rate that went with it). This 'rule of thumb' policy conclusion was also consistent with the results of the large-scale econometric models derived from Hicks's IS-LM model.

However, Lucas put himself in the skeptics' camp, and argued instead in favor of what he called the 'Natural Rate Hypothesis,' that there was no such trade-off – instead, that the economy had a natural rate of employment towards which it tended, and any attempt to increase employment above this rate would simply increase the rate of inflation, without altering employment. He defined the 'Natural Rate Hypothesis' as: 'the hypothesis that different time paths of the general price level will be associated with time paths of real output that do not differ on average' (ibid.: 50).

This, in a convoluted way, asserted the pre-Great Depression neoclassical belief that the economy tended toward an equilibrium in which relative prices were stable, and any attempt to increase the number of people employed would simply cause inflation. Lucas's problem, in asserting this belief, was the evidence. He presented this paper before the 'stagflation' of the 1970s, when inflation and unemployment both rose at the same time, and the evidence of the period from 1960 to 1970 showed a clear trade-off between inflation and unemployment – see Figure 10.5.

Though the inflation-unemployment data at the precise date at which he spoke had a much higher unemployment level than had been experienced at a comparable rate of inflation in the 1960s, shortly after he spoke (in October 1970) the inflation rate plunged in an apparent lagged response to the rise in unemployment during the 1969–70 recession.



10.6 Unemployment–inflation data in the USA, 1950–72

How then to justify skepticism about what seemed an obvious reality? He argued that the 'Phillips Curve' was simply an artifact of how 'agents form and respond to price and wage expectations,' and that attempting to exploit this curve for policy reasons would destroy the apparent trade-off, because agents would change their expectations: 'The main source of this skepticism is the notion that the observed Phillips curve results from the way agents form and respond to price and wage expectations, and that attempts to move along the curve to increase output may be frustrated by changes in expectations that shift the curve' (ibid.: 50).

Lucas thus accepted the empirical evidence of the negative relationship between inflation and unemployment – in that a higher level of inflation was statistically correlated with a lower level of unemployment. However, he argued that this could not be used as a policy tool, alleging that attempts to drive unemployment down by driving inflation up would simply result in higher inflation at the same rate of unemployment.

This was not an entirely new argument – Friedman had made a similar assertion two years earlier (Friedman 1968), using what became known as

'Adaptive Expectations' (Friedman 1971: 331). But Milton's model wasn't good enough for Lucas – though not for the reasons you might expect.

Helicopter Milton Ben Bernanke copped the nickname 'Helicopter Ben' for his observation that a deflation could always be reversed by the government 'printing money':

> the U.S. government has a technology, called a printing press, that allows it to produce as many U.S. dollars as it wishes at essentially no cost. By increasing the number of U.S. dollars in circulation under a fiat (that is, paper) money system, a government should always be able to generate increased nominal spending and inflation […] and sufficient injections of money will ultimately always reverse a deflation. (Bernanke 2002a)

However, the 'Helicopter' part of the nickname alluded not to work by Bernanke, but by his intellectual mentor Milton Friedman, who, more than any other neoclassical, was responsible for the overthrow of the IS-LM model and its replacement by a resurgent neoclassical orthodoxy.

In any sane discipline, Friedman's starting point for his dismantling of the then Keynesian orthodoxy would have been good enough reason to ignore him completely – if not recommend he see a psychiatrist. A key aspect of the neoclassical model is the proposition known as 'money neutrality': that the nominal quantity of money has no effect on the real performance of the macroeconomy, apart from causing inflation. Friedman reasserted that belief, but also clearly stated the condition required for it to operate in reality. The condition was that, if the quantity of money in circulation increased by some factor, then all nominal quantities including the level of debts was also increased by the same factor:

> *It is a commonplace of monetary theory that nothing is so unimportant as the quantity of money expressed in terms of the nominal monetary unit* – dollars, or pounds, or pesos. Let the unit of account be changed from dollars to cents; that will multiply the quantity of money by 100, but have no other effect. Similarly, *let the number of dollars in existence be multiplied by 100; that, too, will have no other essential effect, provided that all other nominal*

> *magnitudes (prices of goods and services, and quantities of other assets and liabilities that are expressed in nominal terms) are also multiplied by 100*. (Friedman 1969: 1; emphases added)

This condition is so clearly not fulfilled in reality that the opposite conclusion therefore applies: since the value of assets and liabilities is not adjusted when inflation occurs, therefore the nominal quantity of money in circulation is important. However, Friedman, who had already given us the 'assumptions don't matter' methodological madness, continued straight on as if it didn't matter that this condition was not fulfilled in reality.

Friedman's next counterfactual assertion was that, left to its own devices, a free market economy with no growth and a constant stock of money would settle into an equilibrium in which supply equaled demand in all markets, and all resources including labor were fully employed (where full employment was defined as supply equaling demand at the equilibrium real wage):[29] 'Let us suppose that these conditions have been in existence long enough for the society to have reached a state of equilibrium. Relative prices are determined by the solution of a system of Walrasian equations' (ibid.: 3).

He then considered what would happen to money prices in such a situation if there was a sudden increase in the money supply: 'Let us suppose now that one day a helicopter flies over this community and drops an additional $1,000 in bills from the sky, which is, of course, hastily collected by members of the community. Let us suppose further that everyone is convinced that this is a unique event which will never be repeated' (ibid.: 4–5).

If you are gobsmacked by this absurd vision of how money is created – dropped from the air like manna from heaven – brace yourself: ideas even more absurd that this are about to come your way.

Friedman's 'helicopter' is of course a parable for the behavior of a central bank (which is not a market actor) that injects money into the system – as Bernanke has himself done twice already, though during the Great Recession rather than when the economy was in 'a state of equilibrium.'[30] But it is a parable which takes for granted that the money supply is completely under the Fed's control – that it is 'exogenous' in the

parlance of economics. In contrast, the empirically derived 'endogenous' theory of money I'll outline in Chapter 14 argues that the money supply is largely outside the Fed's control.

However, with his simplistic model of money creation, Friedman decided that the consequence of doubling the money supply would be that nominal prices would ultimately double. Relative prices and real output would be unaffected in the long run, but – in an important qualification compared to Lucas's later analysis – Friedman conceded that in the interim there could be disturbances to relative prices and the levels of output and employment:

> It is much harder to say anything about the transition. To begin with, some producers may be slow to adjust their prices and may let themselves be induced to produce more for the market at the expense of non-market uses of resources. Others may try to make spending exceed receipts by taking a vacation from production for the market. Hence, measured income at initial nominal prices may either rise or fall during the transition. Similarly, some prices may adjust more rapidly than others, so relative prices and quantities may be affected. There might be overshooting and, as a result, a cyclical adjustment pattern […]. (Ibid.: 6)

Friedman then extended this 'one-off' thought experiment to a theory of inflation by assuming that this 'helicopter drop' of money becomes a continuous process:

> Let us now complicate our example by supposing that the dropping of money, instead of being a unique, miraculous event, becomes a continuous process, *which, perhaps after a lag, becomes fully anticipated by everyone*. Money rains down from heaven at a rate which produces a steady increase in the quantity of money, let us say, of 10 per cent per year. (Ibid.: 8; emphasis added)

The highlighted phrase in the preceding quote is what Friedman later called 'Adaptive Expectations': people form expectations of what will

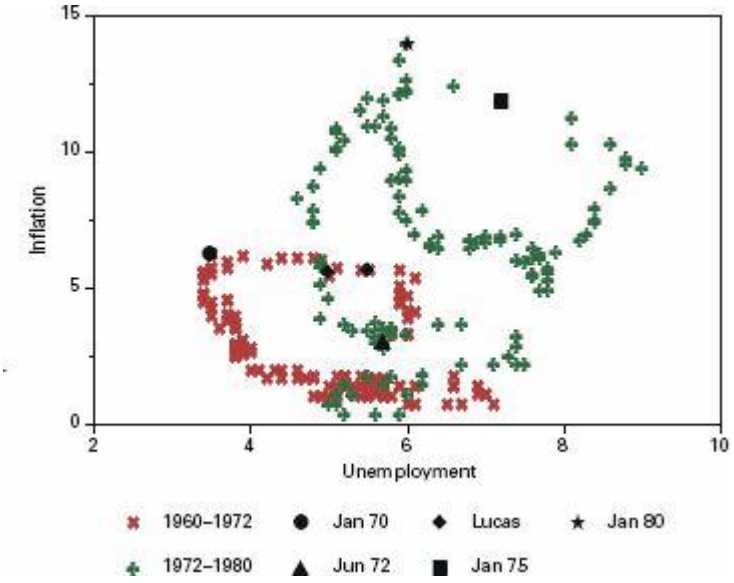
happen in the future based on experience of what has happened in the recent past. He also considered that there could be disturbances in the short term in this new situation of a permanent 10 percent per annum increase in the money supply: 'If individuals did not respond instantaneously, or if there were frictions, the situation would be different during a transitory period. The state of affairs just described would emerge finally when individuals succeeded in restoring and maintaining initial real balances' (ibid.: 10).

However, in the long run, these disturbances dissipate and the economy settles into a long-run equilibrium where all 'real magnitudes' (relative prices, output, employment) are the same as before, but the absolute price level is rising at 10 percent per annum. This occurs not because markets are in disequilibrium with demand exceeding supply, causing prices to rise, but because of the expectations all agents have formed that prices always rise by 10 percent per annum. It is thus expectations which cause prices to rise, rather than disequilibrium: 'One natural question to ask about this final situation is, "What raises the price level, if at all points markets are cleared and real magnitudes are stable?" The answer is, "Because everyone confidently anticipates that prices will rise"' (ibid.).

This was the basis for Friedman's argument against Keynesian demand-management policies, which attempted to exploit the apparent negative relationship between unemployment and the rate of inflation: though a higher rate of growth of the money supply could in the transition cause employment to rise, ultimately the economy would return to its equilibrium level of employment, but at a higher rate of inflation. This was characterized as the 'short-run Phillips Curve' 'moving outwards' – the temporary trade-off between higher inflation and lower unemployment in the transition involved higher and higher levels of inflation for the same level of unemployment – while the 'long-run Phillips curve' was vertical at the long-run equilibrium level of unemployment.

Though Friedman's model was highly simplistic, his vigorous promotion of his 'monetarist' theories just preceded the outbreak of stagflation during the 1970s, giving an apparent vindication of his position. There did indeed seem to be an outward movement of the negative relationship between unemployment and inflation, while there appeared to be a 'long-run' rate of unemployment the economy kept tending towards, at

about a 6 percent rate of unemployment compared to the level of below 4 percent that had been achieved in the 1960s.



10.7 Unemployment–inflation data in the USA, 1960–80

Friedman's monetarism thus defeated Keynesian demand management both inside the academic profession, and in public policy, with central banks trying to limit the growth of the money supply in order to reduce the inflation rate.[31] The period of 'stagflation' – rising unemployment and rising inflation – thus sounded the death-knell for 'Keynesian' economics within the academic profession. However, monetarism's defeat of 'Keynesian' theory wasn't enough for Lucas, since monetarism still implied that the government could alter the level of employment.

'Improving' on Friedman The problem with monetarism, as Lucas saw it, was Friedman's admission that in the short run, a boost to the money supply could have real effects. Lucas began by stating the paradox – for a neoclassical economist – that in neoclassical theory there should be *no* relationship between inflation and employment: changes in aggregate demand caused by changes in the money supply should simply alter the price level while leaving supply unchanged:

> It is natural (to an economist) to view the cyclical correlation between real output and prices as arising from a volatile aggregate demand schedule that traces out a relatively stable,

> upward-sloping supply curve. This point of departure leads to something of a paradox, *since the absence of money illusion on the part of firms and consumers appears to imply a vertical aggregate supply schedule*, which in turn implies that aggregate demand fluctuations of a purely nominal nature should lead to price fluctuations only. (Lucas 1972: 51; emphasis added)

Lucas's comment about 'money illusion' shows that, though he criticized Friedman, it was because Friedman was not neoclassical enough for him – Friedman's macroeconomics was not sufficiently based upon neoclassical microeconomic theory. Since microeconomics predicted that changing all prices and incomes wouldn't affect the output decision of a single consumer, macroeconomics had to conclude that the aggregate rate of unemployment couldn't be altered by monetary means:

> On the contrary, as soon as Phelps and others made the first serious attempts to rationalize the apparent trade-off in modern theoretical terms, the zero-degree homogeneity of demand and supply functions was re-discovered in this new context (as Friedman predicted it would be) and re-named the 'natural rate hypothesis.' (Lucas 1976: 19)

After discussing models used to explain the perceived inflation–unemployment trade-off based on adaptive expectations, Lucas observed that under Adaptive Expectations, it was possible that actual inflation (which was driven by the actual rate of growth of the money supply at a given time) might differ from expected inflation (which was based on people's experience of past inflation that adjusted 'after a lag' to the current rate of inflation). This in turn would mean that, if actual inflation exceeded expected inflation, then there could be 'unlimited real output gains from a well-chosen inflationary policy. Even a once-and-for-all price increase, while yielding no output expansion in the limit, will induce increased output over the (infinity of) transition periods. Moreover, a sustained inflation will yield a permanently increased level of output' (Lucas 1972: 53).

But herein lay a dilemma: Lucas's logic had revealed that the only way to conclude that there was a natural rate of employment was to assume that expected inflation always equaled actual inflation, which in turn means assuming that people can accurately predict the future.

Obviously Lucas couldn't assume this.

Well, obviously, if he wasn't a neoclassical economist! Because that's precisely what he did assume. His way of stating this was obtuse, but nonetheless unmistakable:

> In the preceding section, the hypothesis of adaptive expectations was rejected as a component of the natural rate hypothesis on the grounds that, under some policy [the gap between actual and expected inflation] is non-zero. If the impossibility of a non-zero value […] is taken as an essential feature of the natural rate theory, *one is led simply to adding the assumption that [the gap between actual and expected inflation] is zero as an additional axiom* […]. (Ibid.: 54; emphasis added)

Such an 'axiom' is transparently nonsense – something that might have led a sensible person to stop at this point. But instead Lucas immediately moved on to an equivalent way of stating this 'axiom' that wasn't so obviously absurd: 'or to assume that expectations are rational in the sense of Muth' (ibid.).

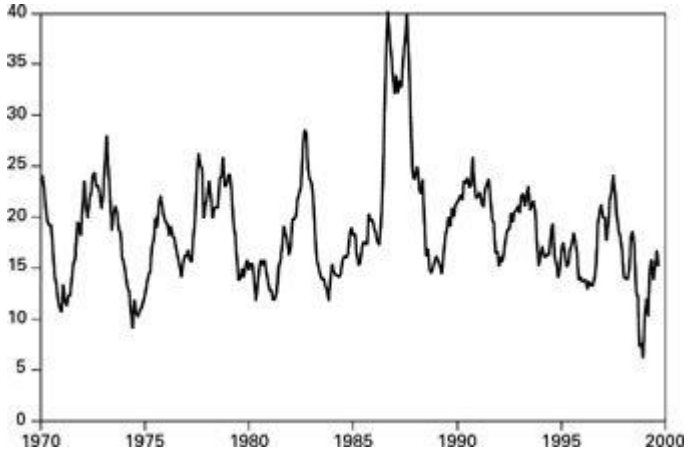
Thus neoclassical macroeconomics began its descent into madness which, thirty-five years later, left it utterly unprepared for the economic collapse of the Great Recession.

Expectations and rationality Decades before, when the Great Depression also forced economists to consider reality rather than their largely verbal models of equilibrium, Keynes made a similar point to Lucas's, that expectations about the future affect decisions today, and he pilloried the neoclassical theorists of his day for ignoring this.

Keynes welded the role of expectations in economics with uncertainty about the future, and considered how people still manage to make decisions despite uncertainty. Thirty-five years later, Lucas reintroduced expectations into macroeconomics, but with the assumption that people could accurately

predict the future and thus eliminate uncertainty – an even more absurd position than that of his pre-Great Depression predecessors, whom Keynes merely accused of 'abstracting from the fact that we know very little about the future.'

It is one of the greatest abuses of language committed by neoclassical economists that a proposition which in any other discipline would be deemed as insane – that on average, people's expectations about the future are accurate – goes under the name of 'rational expectations' in economics. That the idea could even be countenanced shows the extent to which neoclassical economics is driven by a teleological desire to prove that capitalism is fundamentally stable, rather than by a desire to understand the empirical record of the actual economy.



10.8 The hog cycle (hog/corn price)

The paper that initially developed the concept of 'rational expectations' (Muth 1961) applied it to microeconomics, to develop a critique of a simplified theory of price cycles in agricultural markets known as 'the Cobweb model.' Agricultural products like pork were subject to irregular cycles in prices – see Figure 10.8 – and one explanation that microeconomists developed was that time lags in production generated the cycles.

The Cobweb cycle model argued that suppliers would take prices one season as a guide to how many hogs to breed in the next season. When prices were high, many hogs would be raised the subsequent season, which would cause prices to crash the season after; while when prices were low, few hogs would be raised the next season, which would cause prices to rise.

Prices thus fluctuated in disequilibrium over time, overshooting and undershooting the equilibrium price.

The Cobweb assumed the existence of standard Marshallian supply and demand curves – something we have debunked in Chapters 3–5 – and also had a hard time explaining the lengthy cycles that could occur, which were measured in multiples of the breeding cycle itself.[32] Seizing on the latter weakness, Muth proposed that farmers' price expectations were not simply that last year's prices would be next year's, but that they would be to some degree informed by experience – a sensible observation in itself. However, he extrapolated from this to the following hypothesis: 'I should like to suggest that expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory' (ibid.: 316).

That is, he assumed that farmers formed their expectations of next year's price by assuming that it would be the equilibrium price as given by the Marshallian model of supply and demand, and that these expectations were correct – they were what would happen because the model itself was accurate: 'The hypothesis can be rephrased a little more precisely as follows: that expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the "objective" probability distributions of outcomes)' (ibid.).

Not only did Muth believe that the predictions of the theory were that price would equal marginal cost in equilibrium (erroneously, as we saw in Chapter 4), he also assumed that the producers had implicit knowledge of the market's supply and demand functions, and would form their expectations accordingly and therefore correctly anticipate the future.

Muth's rationality was thus rationality on steroids – not only did people know of and behave in their own best interests, they also knew how the system in which they were bit players actually behaved. This is not mere utility-maximizing rationality with respect to one's own interests (something I showed was computationally impossible in Chapter 3), but 'meta-rationality' – knowledge of how the entire system in which we are embedded works which is so good that the average expectation of the future will be correct.

This is the opposite of the realistic concept of uncertainty that Keynes had tried, unsuccessfully, to introduce into economic theory. Muth introduced expectations into his model in a manner that neutralized uncertainty.

Though there were some nuances later in the article which made it somewhat less unrealistic – including that expectations might 'consistently over- or under-discount the effect of current events' (ibid.: 321), the impact of inventories, speculators and so on[33] – the impact of this 'rational expectations hypothesis' on the model of price fluctuations in an agricultural market was that the expected market price was the equilibrium price, and all fluctuations about this price were caused by random shocks.

This is a familiar tune in neoclassical economics: whenever an attempt to incorporate a more realistic vision of how the economy functions results in a need to think in a disequilibrium way, economists dream up ways of relegitimizing equilibrium analysis once more. This is accepted within neoclassical economics itself, even if it involves doing severe damage to realism – as the assumption that the future can be (on average) accurately predicted surely does – and even if it involves an obvious contradiction of other parts of neoclassical economics.

Muth committed such a contradiction when he put forward as a justification for assuming rational expectations at the market level the proposition that: 'Information is scarce, and the economic system generally does not waste it' (ibid.: 316).

Leaving aside the very concept of information about the future, this assertion within neoclassical economic theory leads to the conclusion that expectations should be less than rational.

If information is scarce, then it should have a price, and a rational agent should purchase information (about the future …) up until the point at which the marginal cost of this information equals the marginal benefit from acquiring it. This would necessarily occur before enough information (about the future …) was purchased to allow completely rational expectations (about the future …) to be formed, so that actual expectations should be less than fully 'rational.'

No such limit occurred to Muth, however, let alone to Lucas, who appropriated this concept from the model of a single market to apply it at

the level of the entire economy.

The macroeconomics of Nostradamus The argument that producers in a given market have at least some idea of how that market works, and can therefore produce slightly informed predictions of what next season's price might be, given this season's outcome, is not entirely unreasonable. But the argument that agents in a macroeconomy can know how the macroeconomy works and therefore correctly anticipate the future course of macroeconomic variables like inflation is simply absurd.

However, this absurdity was in fact a necessity for neoclassical economics. If it were to maintain the belief that the economy was fundamentally stable, then expectations of the future had to be either ignored or tamed.

In Keynes's day, as he himself noted, neoclassical economics did the former. After Keynes, expectations were again ignored in Hicks's development of the IS-LM model, and then the numerical forecasting models derived from it. Then, in one of the greatest travesties in the history of economic thought, Muth and Lucas could claim that they were introducing expectations into economic theory, because they were clearly unaware of Keynes's earlier insistence on the importance of expectations in the context of uncertainty about the future.

However, here they were constrained by the dilemma that Keynes observed afflicted his neoclassical contemporaries, when he noted that they attempted 'to deal with the present by abstracting from the fact that we know very little about the future' (Keynes 1937: 215). Neoclassical economics could only maintain its belief that the economy was in equilibrium if actions today, taken on the basis of how conditions were expected to be in the future, were correct. So the choice that neoclassical economics faced was between ignoring the future, or pretending that it could be accurately foreseen.

Keynes's contemporaries chose the former route; Lucas and modern neoclassicals instead embraced the latter – and had the hide to call such a view 'rational.' In reality, 'rational expectations' was a device, not to introduce expectations into economic modeling, but to keep time and uncertainty about the future out of it. In place of dealing with the present 'by abstracting from the fact that we know very little about the future,'

rational expectations deals with the present 'by pretending that we can predict the future.'

Microeconomic macroeconomics The concept that agents in a complex system like the macroeconomy can accurately predict its future should have been rejected on first sight. Not only does it ignore uncertainty, even prediction of what a model itself will do in the future is only possible if the model is 'ergodic' – meaning that the past history of the model is a reliable guide to its future behavior.

The complex dynamic models we considered in Chapter 9, such as Lorenz's model of atmospheric turbulence, are non-ergodic.[34] The past history of a complex model is not a reliable guide to its future behavior, because where the model will evolve to is dependent on where it starts from – the so-called 'Butterfly Effect' applies. Two situations with differences in initial conditions that are too small to be distinguished from each other will have drastically different outcomes in the future: they will be similar for a short while (which is why weather forecasting is accurate only about a week in advance) but then diverge completely.

Only if models of the economy are not of this class are 'rational expectations' possible even within the model. The easiest way to make rational expectations work within a model is to make it linear – and this is what Muth did in his first model:

> For purposes of analysis, we shall use a specialized form of the hypothesis. In particular, we assume:
>
> 1. The random disturbances are normally distributed.
> 2. Certainty equivalents exist for the variables to be predicted.
> 3. The equations of the system, including the expectations formulas, are linear.
>
> These assumptions are not quite so strong as may appear at first because any one of them virtually implies the other two. (Muth 1961: 317)

Though some subsequent 'rational expectations' models used in macroeconomics had nonlinearities, they continued to make Muth's second assumption – that the 'exogenous shocks,' which are the only explanation these models have for cyclical behavior, are 'normally distributed' – and as Muth observes, this is effectively the same as having a linear model.

However, 'rational expectations' makes no sense in non-ergodic models: any predictions made from within such a model about the model's future behavior would be wrong (let alone predictions made about the economy the model is alleged to simulate). Crucially, the errors made by agents within that model would not be 'normally distributed' – they would not be neatly distributed around the model's mean as in the classic 'Bell Curve.' Instead the distribution would be 'chaotic,' with lots of what Nassim Taleb labeled 'Black Swan events' (Taleb 2007). It would be futile to have 'rational expectations' in such a model, because these would be misleading guides to the model's future. The model's future would be uncertain, and the best thing any agent in such a model could do would be to project forward its current trajectory, while also expecting that expectation to be wrong.

What applies to a model applies in extremis to the real world, and parallels Keynes's observations about how people in a market economy actually behave: they apply conventions, the most common of which is to extrapolate forward current conditions, even though 'candid examination of past experience' (Keynes 1937: 214) would show that these conditions did not persist.

Keynes remarked that superficially this might appear irrational, but there is no better course of action when the future is uncertain. One of Keynes's observations, highlighted in the next quote, directly contradicts the key assumption of rational expectations, which is that on average people's expectations about the future will be correct:

> *It would be foolish, in forming our expectations, to attach great weight to matters which are very uncertain. It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and scanty.*

For this reason the facts of the existing situation enter, in a sense disproportionately, into the formation of our long-term expectations; our usual practice being to take the existing situation and to project it into the future […]

The essence of this convention […] lies in assuming that the existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change. This does not mean that we really believe that the existing state of affairs will continue indefinitely. We know from extensive experience that this is most unlikely. The actual results of an investment over a long term of years very seldom agree with the initial expectation.

Nor can we rationalize our behavior by arguing that to a man in a state of ignorance errors in either direction are equally probable, so that there remains a mean actuarial expectation based on equi-probabilities. For it can easily be shown that the assumption of arithmetically equal probabilities based on a state of ignorance leads to absurdities.

We are assuming, in effect, that the existing market valuation, however arrived at, is uniquely correct in relation to our existing knowledge of the facts which will influence the yield of the investment, and that it will only change in proportion to changes in this knowledge; though, philosophically speaking, it cannot be uniquely correct, since our existing knowledge does not provide a sufficient basis for a calculated mathematical expectation […]

In abnormal times in particular, when the hypothesis of an indefinite continuance of the existing state of affairs is less plausible than usual even though there are no express grounds to anticipate a definite change, *the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for a reasonable calculation*. (Keynes 1936: 148, 152, 154; emphasis added)

The concept of rational expectations should therefore have died at birth, but because it let neoclassical economists return to their pre-Keynesian practice of arguing that the economy was self-regulating and always either in or tending toward equilibrium, rational expectations was instead embraced. Lucas and his colleagues Thomas Sargent, Neil Wallace, Edward Prescott, Leonard Rapping, and several others produced a series of papers that developed models of the macroeconomy that extrapolated directly from the alleged behavior of a single utility-maximizing and profit-maximizing agent who was endowed, via 'rational expectations,' with the capacity to accurately predict the future.

One of these predictions was that increasing the money supply would cause inflation. In a model without 'rational expectations,' if the government increased the money supply in order to reduce unemployment, there would be a lag between when the money supply was increased, and when the inflation actually occurred. In the meantime, the increased money supply would have the impact desired by the government, of increasing economic activity – and hence reducing unemployment. This was Friedman's adaptive expectations, leading to the undesirable result – from the point of view of neoclassical economists – that the government could reduce the unemployment rate below equilibrium via a policy of permanent accelerating inflation.

The twist of adding expectations into the model, when expectations were identical to the prediction of the model, was that inflation would occur instantly, rather than with a lag. This is because, since everyone expects an increased money supply to cause inflation, everyone instantly puts their prices up as soon as the money supply rises. The lag between an increase in the money supply and an increase in prices is eliminated, and with it disappears any temporary impact of the money supply on unemployment. In one of the pivotal papers in this literature, Sargent and Wallace put it this way:

> The public knows the monetary authority's feedback rule and takes this into account in forming its expectations [… therefore] unanticipated movements in the money supply cause movements in [output], but anticipated movements do not […]

> [R]emoving the assumption that the authority can systematically trick the public eliminates the implication that there is an exploitable tradeoff between inflation and unemployment in any sense pertinent for making policy. *The assumption that the public's expectations are 'rational' and so equal to objective mathematical expectations accomplishes precisely this*.
>
> In this system, there is no sense in which the authority has the option to conduct countercyclical policy. To exploit the Phillips Curve, it must somehow trick the public. But by virtue of the assumption that expectations are rational, there is no feedback rule that the authority can employ and expect to be able systematically to fool the public. This means that *the authority cannot expect to exploit the Phillips Curve even for one period*. Thus, combining the natural rate hypothesis with the assumption that expectations are rational transforms the former from a curiosity with perhaps remote policy implications into an hypothesis with immediate and drastic implications about the feasibility of pursuing countercyclical policy. (Sargent and Wallace 1976: 173, 176, 177–8; emphases added)

Not surprisingly, this doctrine was termed the 'policy ineffectiveness proposition.' If anything that was consciously done by policymakers to manipulate the economy led instantly to countervailing behavior by people in the economy, then nothing the government could do would alter the rate of unemployment. Instead, all the government could do was cause inflation.

This doctrine also provided a basis on which to attack the strongest edifices of macroeconomics at the time, the large-scale numerical simulations of the economy derived from Hicks's IS-LM model.

The Lucas critique These numerical simulations had two roles: providing a means to organize economic statistics from the past, and providing a means to forecast what might happen to the economy if a new government policy were implemented. Lucas's critique focused on this second role, by arguing that the parameters in the models' equations reflected the expectations that

agents in the economy had under past policies. A new policy would evince new reactions from agents within the economy, thus altering the parameters and rendering projected economic outcomes based on them invalid. As Lucas put it:

> The thesis of this essay is that […] the 'theory of economic policy' […] is in need of major revision. More particularly, I shall argue that the features which lead to success in short-term forecasting are unrelated to quantitative policy evaluation, that the major econometric models are (well) designed to perform the former task only, and that simulations using these models can, in principle, provide no useful information as to the actual consequences of alternative economic policies. (Lucas 1976: 19–20)

Leaving aside the absurdity of using this critique to justify the assumption of rational expectations, Lucas's general point was valid: one of the many things that an economic model should incorporate is the possibility that the behavior of the economy could alter in response to a change in government policy.

However, it is a wild extrapolation to then argue that the change would be sufficient to completely neutralize the policy, as rational expectations exponents contended. It is also committing the fallacy of strong reductionism to believe that this justifies overthrowing explicitly macroeconomic models and replacing them with ones in which macroeconomics is directly extrapolated from microeconomics.

The applicability of the Lucas critique to the existing IS-LM-based macroeconomic modeling tradition was also a matter of degree, as Gordon argued at the same conference:

> While I am prepared to grant the validity of the proposition that the mechanical extrapolation of a model with fixed parameters cannot provide useful information on the effects of all policy changes, on the other hand the effects of some policy changes can be determined if parameter shifts are allowed and are either (a) estimated from the response of parameters to policy changes

> within the sample period or (b) are deduced from a priori theoretical consideration. (Gordon 1976: 47)

However, Lucas and the Rational Expectations Mafia[35] weren't interested in nuances: their objective was the elimination of macroeconomics as a separate discipline, and the replacement of IS-LM-based macroeconomic models with models that extrapolated the neoclassical microeconomics to an analysis of the entire economy. Manifestos to this effect are spread throughout the economic literature.

The microeconomic manifesto The belief that macroeconomics should be applied microeconomics was an article of faith for neoclassical economists, and this faith was radiantly on display at Lucas's keynote speech to the History of Political Economy conference in the year in which he became president of the American Economic Association. In this memoir, he reiterated the view that macroeconomics had to be based on Walrasian microeconomics:

> I think Patinkin was absolutely right to try and use general equilibrium theory to think about macroeconomic problems. Patinkin and I are both Walrasians, whatever that means. I don't see how anybody can not be.
>
> I also held on to Patinkin's ambition somehow, that the theory ought to be microeconomically founded, unified with price theory. I think this was a very common view […] Nobody was satisfied with IS-LM as the end of macroeconomic theorizing. The idea was we were going to tie it together with microeconomics and that was the job of our generation. Or to continue doing that. That wasn't an anti-Keynesian view. You can see the same ambition in Klein's work or Modigliani's. (Lucas 2004: 16, 20)

Today, macroeconomic textbooks start from the presumption that macroeconomics must have microeconomic foundations. Ljungqvist and Sargent's 2004 text gives a typical justification for this:

> This book is about micro foundations for macroeconomics. [There are] two possible justifications for putting microfoundations underneath macroeconomic models. The first is aesthetic and pre-empirical: models with micro foundations are by construction coherent and explicit. And because they contain descriptions of agents' purposes, they allow us to analyze policy interventions using standard methods of welfare economics. Lucas […] gives a distinct second reason: a model with micro foundations broadens the sources of empirical evidence that can be used to assign numerical values to the model's parameters […] We don't think that the clock will soon be turned back to a time when macroeconomics was done without micro foundations. (Ljungqvist and Sargent 2004: xxvi–xxvii)

The problem for early would-be neoclassical macroeconomists was that, strictly speaking, there was no microeconomic model of macroeconomics when they began their campaign. So they developed a neoclassical macro model from the foundation of the neoclassical growth model developed by Nobel laureate Robert Solow (Solow 1956) and Trevor Swan (Swan 2002). They interpreted the equilibrium growth path of the economy as being determined by the consumption and leisure preferences of a representative consumer, and explained deviations from equilibrium – which the rest of us know as the business cycle – by unpredictable 'shocks' to technology and consumer preferences.

This resulted in a model of the macroeconomy as consisting of a single consumer, who lives for ever, consuming the output of the economy, which is a single good produced in a single firm, which he owns and in which he is the only employee, which pays him both profits equivalent to the marginal product of capital and a wage equivalent to the marginal product of labor, to which he decides how much labor to supply by solving a utility function that maximizes his utility over an infinite time horizon, which he rationally expects and therefore correctly predicts. The economy would always be in equilibrium except for the impact of unexpected 'technology shocks' that change the firm's productive capabilities (or his consumption preferences) and thus temporarily cause the single capitalist/worker/consumer to alter his

working hours. Any reduction in working hours is a voluntary act, so the representative agent is never involuntarily unemployed, he's just taking more leisure. And there are no banks, no debt, and indeed no money in this model.

You think I'm joking? I wish I was. Here's Robert Solow's own summary of these models – initially called 'real business cycle' models, though over time they morphed into what are now called 'Dynamic Stochastic General Equilibrium' models:

> The prototypical real-business-cycle model goes like this. There is a single, immortal household – a representative consumer – that earns wages from supplying labor. It also owns the single price-taking firm, so the household receives the net income of the firm. The household takes the present and future wage rates and present and future dividends as given, and formulates an optimal infinite-horizon consumption-saving (and possibly labor-saving) plan […] The firm looks at the same prices, and maximizes current profit by employing labor, renting capital and producing and selling output […] (Solow 2001: 23)
>
> In the ordinary way, an equilibrium is a sequence of inter-temporal prices and wage rates that makes the decisions of household and firm consistent with each other. This is nothing but the neoclassical growth model […]
>
> The theory actually imagines that the model economy is disturbed from time to time by unforeseeable shocks to the technology and the household's tastes […] There is thus nothing pathological or remediable about observed fluctuations. Unforeseeable disturbances are by definition unforeseen; after one of them has happened, the economy is already making optimal adjustments, given its technology and the inter-temporal preferences of its single inhabitant or identical inhabitants. There is no role for macroeconomic policy in this world […] the best it [the government] can do is to perform its necessary functions in the most regular, predictable way, so as not to add unnecessary variance to the environment. (Ibid.: 23–4)

If you get the feeling that Solow – a neoclassical economist par excellence and, as noted, the author of the growth model from which real business cycle models were derived – is not happy with the microeconomic takeover of macroeconomics, you'd be right. Though microeconomics masquerading as macroeconomics took over PhD programs across the USA, and it is all the current crop of neoclassicals really knows, there has always been opposition to this approach to macroeconomics from within the neoclassical school itself. Solow's own reactions are the most notable, since Solow's growth model is acknowledged by Finn Kydland and Edward Prescott, the originators of these models, as its fountainhead (Kydland and Prescott 1991: 167–8).

Solow's reaction to the fact that his growth model was used as the basis of modern neoclassical macroeconomics was one of bewilderment:

> The puzzle I want to discuss – at least it seems to me to be a puzzle, though part of the puzzle is why it does not seem to be a puzzle to many of my younger colleagues – is this. More than forty years ago, I […] worked out […] neoclassical growth theory […] [I]t was clear from the beginning what I thought it did not apply to, namely short-run fluctuations in aggregate output and employment […] the business cycle […]
>
> [N]ow […] if you pick up an article today with the words 'business cycle' in the title, there is a fairly high probability that its basic theoretical orientation will be what is called 'real business cycle theory' and the underlying model will be […] a slightly dressed up version of the neoclasssical growth model. The question I want to circle around is: how did that happen? (Solow 2001: 19)

Solow inadvertently provided one answer to his own question when he discussed the preceding IS-LM model:

> For a while the dominant framework for thinking about the short run was roughly 'Keynesian.' I use that label for convenience; *I have absolutely no interest in 'what Keynes really meant*.' To be

> more specific, the framework I mean is what is sometimes called 'American Keynesianism' as taught to many thousands of students by Paul Samuelson's textbook and a long line of followers. (Ibid.: 21)

How bizarre! Solow is decrying that poor scholarship led to his growth cycle model being used for a purpose for which it was not designed, and yet he is blasé about whether or not the models of the economy he helped develop, and which he labels Keynesian (albeit with the qualifier 'American'), have anything to do with Keynes's ideas.

The old saying 'As ye sow, so shall ye reap' applies here. The poor scholarship that let American economists delude themselves into believing that they were Keynesians, when in fact they were extending models originated by – and later disowned by – John Hicks, now let them use Solow's growth model as a foundation for models of the business cycle, even though Solow himself disowned the enterprise on two very valid grounds.

The first is that the limitations of IS-LM modeling pointed out in the Lucas critique did not justify modeling the entire macroeconomy as a single representative agent. Unlike many neoclassicals, Solow was aware that the Sonnenschein-Mantel-Debreu conditions discussed in Chapter 3 invalidate attempts to model the entire economy by extrapolating from microeconomic theory about the behavior of individual consumers:

> the main argument for this modeling strategy has been a more aesthetic one: its virtue is said to be that it is compatible with general equilibrium theory, and thus it is superior to ad hoc descriptive models that are not related to 'deep' structural parameters. The preferred nickname for this class of models is 'DSGE' (dynamic stochastic general equilibrium). I think that this argument is fundamentally misconceived […] *The cover story about 'microfoundations' can in no way justify recourse to the narrow representative-agent construct* […]

He also supplied a simple analogy as to why the valid criticism of IS-LM models – that they don't consider that economic agents may change

their behavior when government policies change – does not justify the strong reductionist approach of reducing macroeconomics to applied microeconomics:

> The nature of the sleight-of-hand involved here can be made plain by an analogy. I tell you that I eat nothing but cabbage. You ask me why, and I reply portentously: I am a vegetarian! But vegetarianism is reason for a meatless diet; it cannot justify my extreme and unappetizing choice. Even in growth theory (let alone in short-run macroeconomics), reasonable 'micro-foundations' do not demand implausibility; indeed, they should exclude implausibility. (Solow 2007: 8; emphasis added)

Solow's second point is a practical one: the standard fare of macroeconomics – booms and slumps, inflation and deflation, unemployment rising as people are sacked during recessions – cannot occur in pure DSGE models. They are therefore a particularly useless foundation from which to analyze such phenomena. In a paper tellingly entitled 'Dumb and dumber in macroeconomics,' Solow observed that, though 'The original impulse to look for better or more explicit micro foundations was probably reasonable […]'

> What emerged was not a good idea. The preferred model has a single representative consumer optimizing over infinite time with perfect foresight or rational expectations, in an environment that realizes the resulting plans more or less flawlessly through perfectly competitive forward-looking markets for goods and labor, and perfectly flexible prices and wages.
>
> *How could anyone expect a sensible short-to-medium-run macroeconomics to come out of that set-up?* My impression is that this approach (which seems now to be the mainstream, and certainly dominates the journals, if not the workaday world of macroeconomics) has had no empirical success; but that is not the point here. *I start from the presumption that we want macroeconomics to account for the occasional aggregative pathologies that beset modern capitalist economies, like*

> *recessions, intervals of stagnation, inflation, 'stagflation,' not to mention negative pathologies like unusually good times. A model that rules out pathologies by definition is unlikely to help*. (Solow 2003: 1; emphases added)

In typical neoclassical fashion, Solow's legitimate complaints about 'micro-foundations-based representative agent macroeconomics' have been ignored. The accepted wisdom within neoclassical economics remains that macro models had to have 'good microeconomic foundations,' and the only dispute, prior to the Great Recession, was over what constituted good foundations. This led to a bifurcation within neoclassical macroeconomics into two camps, one of which preferred to model the entire economy as a single agent existing in a perfectly competitive general equilibrium, the other of which modeled the economy as one (and occasionally more than one) agent existing in a state of imperfectly competitive general equilibrium.

It was a sham dichotomy, because they both shared the vision that, if the neoclassical fantasy of perfect competition applied, there would be no macroeconomic problems. They differed only on whether they believed that the neoclassical fantasy could be assumed to apply in reality or not. As the end of the first decade of the twenty-first century approached, they had largely reached a rapprochement. And then the Great Recession crushed both their visions.

Much ado about almost nothing: freshwater versus saltwater macroeconomics Nobel laureate Paul Krugman popularized the monikers 'freshwater' and 'saltwater' economists for these two approaches to economics, and makes much of their differences (Krugman 2009a, 2009b). But the reality is that what they share is far more important than their slight differences, because they are both neoclassical theories in which macroeconomic problems arise only if there are microeconomic 'imperfections,' and they both believe that a perfectly competitive economy with flexible prices is the definition of perfection.

As I explained in Chapters 3–5 of this book, this vision even of their own model is fundamentally wrong. Demand curves derived from aggregating the individual demand of 'perfectly rational' consumers could

have any shape at all. Competitive firms would not produce where marginal cost equals price, but where marginal revenue equals demand, and set price above this level. Market demand curves would intersect with the marginal revenue curves of the industry's suppliers in multiple locations, making the very notion of an equilibrium price in a single market problematic. Incorporating the issues covered in subsequent chapters results in even more of a mess. Not even microeconomic analysis can be based on neoclassical microeconomics – let alone the analysis of an entire economy.

Both saltwater and freshwater economists were therefore up Strong Reductionism Creek without a paddle when the Great Recession hit. I would prefer to leave them there, but since their squabbling and mea culpas dominate even today's debate about where macroeconomics should go from now, I have to detail how they got there in the first place, and why they remain lost in an irrelevant intellectual tributary of their own making when the real world is drowning in the flood of the Great Recession.

From Keynes to freshwater macroeconomics Both freshwater and saltwater macroeconomics had their genesis in the pre-Keynesian belief that all dilemmas at the level of the overall economy must instead be signs of malfunctioning in particular markets – and normally the labor market. As Coddington noted in the very early days of the neoclassical revolt against Keynesian macroeconomics, Keynes's neoclassical predecessors – whom he labeled 'Classical' – had precisely the same view, and it was based on a reductionist vision of how economics should be done:

> Keynes attacked a body of theory that he designated 'Classical.' [… and] called into question the method of analysis by which this system was constructed […] this method consisted of analyzing markets on the basis of the choices made by individual traders […] This method of analysis […] I will refer to as 'reductionism,' on the grounds that the central idea is the reduction of market phenomena to (stylized) individual choices. (Coddington 1976: 1258)

This pre-Keynesian vision was reconstructed by neoclassical macroeconomists after Keynes. Their starting point was the key implication

of 'The predominant theory of markets, namely the Walrasian or Arrow Debreu model of general competitive equilibrium,' which was

> that unemployment never appears and that economic policy never has universally good effects. First, it postulates that the supply and demand by price-taking agents equilibrates in the market for any commodity, including labor. Hence, no unemployment occurs. Second, Walrasian equilibria are efficient, as anticipated by Adam Smith's 'invisible hand' […] Thus, either economic policy has no effects or it hurts some group of citizens. (Silvestre 1993: 105)

This pre-Keynesian attitude was reborn with the development of what Coddington termed 'Reconstituted Reductionism,' because believers in neoclassical economics could give Keynes's work any intellectual credence only if it were seen as a statement of what would happen out of equilibrium, since in general equilibrium, Walras's Law would apply and there could be no macroeconomic problems. As Robert Clower put it, Keynes either had such a hypothesis 'at the back of his mind, or most of the *General Theory* is theoretical nonsense' (Clower 1969: 290).

As I've explained above, Walras's Law itself is a theoretical nonsense that ignores the role of credit in a market economy. However, the belief that Walras's Law was a universal truth, and that any deviation from its fundamental result – that macroeconomic crises were in fact manifestations of disequilibrium in individual markets – must be a fallacy was shared by both sides of the neoclassical saltwater/freshwater divide. Coddington correctly noted that both saltwater and freshwater economists assumed that economics had to be conducted from a reductionist perspective.

> the claim that equilibrium theorizing must be abandoned in order to accommodate Keynesian ideas postulates that theorizing must be carried out in accordance with the reductionist program. (Coddington 1976: 1269)

> To ask this question, one needs a construction in which prices adjust less than instantaneously to economic circumstances, so that at any point in time the prices may be effectively providing

> incentives to act, but the information they reflect will not be appropriate for the equilibrium that is being approached. (Ibid.: 1270)

Saltwater economists were willing to abandon equilibrium (or at least perfectly competitive equilibrium) but still believed they had to reason in a reductionist way. Freshwater economists clung to modeling the economy as if it were always in equilibrium, which gave rise to the problem for them of the historical fact that unemployment occurred – or, in their terms, that economic statistics reported that, on occasions, lots of people were not working. But according to their theory, if all markets including labor were in equilibrium apart from the impact of unexpected shocks, unemployment in general could not exist. How, then, to interpret past instances when high levels of unemployment were recorded – like, for example, the Great Depression?

Their interpretation, in a nutshell, was that the Great Depression was an extended holiday: something happened that caused workers to decide to work less, and this increase in leisure was recorded by the statistical agencies as an increase in unemployment. This something was a change in government policy that made it rational for workers to voluntarily reduce their working hours in order to maximize their lifetime utility.

You think I'm joking? Consider these statements by the doyen of the freshwater or 'New Classical' faction of neoclassical macroeconomists, Nobel laureate Edward Prescott:

> the key to defining and explaining the Great Depression is the behavior of *market hours worked per adult* […] there must have been a fundamental change in labor market institutions and industrial policies that lowered steady-state, or normal, market hours […]
>
> [T]he economy is continually hit by shocks, and what economists observe in business cycles is the effects of past and current shocks. A bust occurs if a number of negative shocks are bunched in time. A boom occurs if a number of positive shocks are bunched in time. Business cycles are, in the language of Slutzky, the 'sum of random causes.'

> The fundamental difference between the Great Depression and business cycles is that market hours did not return to normal during the Great Depression. Rather, market hours fell and stayed low. In the 1930s, labor market institutions and industrial policy actions changed normal market hours. I think these institutions and actions are what caused the Great Depression […]
>
> *From the perspective of growth theory, the Great Depression is a great decline in steady-state market hours*. I think this great decline was the unintended consequence of labor market institutions and industrial policies designed to improve the performance of the economy. Exactly what changes in market institutions and industrial policies gave rise to the large decline in normal market hours is not clear […]
>
> The Marxian view is that capitalistic economies are inherently unstable and that excessive accumulation of capital will lead to increasingly severe economic crises. Growth theory, which has proved to be empirically successful, says this is not true. *The capitalistic economy is stable, and absent some change in technology or the rules of the economic game, the economy converges to a constant growth path with the standard of living doubling every 40 years*. In the 1930s, there was an important change in the rules of the economic game. This change lowered the steady-state market hours. The Keynesians had it all wrong. *In the Great Depression, employment was not low because investment was low. Employment and investment were low because labor market institutions and industrial policies changed in a way that lowered normal employment*. (Prescott 1999: 1–3; emphases added)

Prescott's culprit for these changes, predictably, is the government: 'government policies that affect TFP [total factor productivity] and hours per working-age person are the crucial determinants of the great depressions of the 20th century […]' (Kehoe and Prescott 2002: 1).

The reason that Prescott and his fellow freshwater economists were led to such a frankly crazy interpretation of the Great Depression is that their model allowed no other alternative.

As a reminder, their model, in a nutshell, is the following. There is a single consumer, endowed with rational expectations, who aims to maximize his utility from consumption and leisure over the infinite future. His income emanates from the profits of the single firm in the economy, of which he is the sole owner, and in which he is the sole worker, where the profits he receives are the marginal product of capital times the amount of capital employed by the firm, and his wages are the marginal product of labor times the hours he works in the firm. The output of the firm determines consumption and investment output today, and today's investment (minus depreciation) determines tomorrow's capital stock. The single consumer/capitalist/worker decides how much of current output to devote to investment, and how many hours to work, so that the discounted expected future value of his consumption plus leisure plan is maximized. Technology enables expanding production over time, with productivity growing at a constant rate but subject to random shocks, and these shocks cause the equilibrium levels of labor and investment chosen by the consumer/capitalist/worker to alter – but the choices made are always equilibrium choices.

With that bizarre vision of a market economy, while standard business cycle fluctuations in employment can be explained as a rational response by workers to work less today – because productivity has increased owing to a series of positive technology shocks – the only explanation for the sustained decline in employment that occurs during a depression is that it is a rational response by the household sector to a change in government policy to take more leisure.

The saltwater–freshwater dialectic Saltwater neoclassicals like Krugman, Stiglitz and so on can at least be congratulated for being realistic enough to reject this extreme Panglossian view of how the economy operates. But the dilemma for them is that the freshwater vision is more faithful to the underlying neoclassical vision of the economy that they share with the freshwaters.

Herein lies the dialectic that has defined the development of neoclassical macroeconomics over time, between theoretical purity on the one hand and reality on the other. To a neoclassical, theoretical purity involves reducing everything to the Walrasian vision of a perfectly equilibrating economy – in which case no macroeconomic crises can occur (since price movements will rapidly eliminate any macro imbalances caused by disequilibria in individual markets). Reality introduces the vexing counterpoint that recessions do occur, and persist for an inordinate period of time, and that it simply beggars belief that the dole queues of the 1930s – and the massive unemployment of the Great Recession – are manifestations of workers voluntarily taking more leisure.

This in turn leads to a dialectical division of labor within neoclassical economics. Ideologues who are most committed to the vision of the free market as the perfect system were the first to respond to any challenge to this vision – thus firstly Friedman, then Lucas, Prescott and the other freshwater economists led the revolt against IS-LM Keynesianism, and the Real Business Cycle/DSGE approach to economics evolved.

Then the liberals or comparative realists within neoclassical economics – Stiglitz, Krugman, Woodford and the like – reacted to the unrealism that the extreme purity approach embodies, though at the same time they took this perspective as the proper point from which to commence macroeconomic modeling. So they embellished the purist model with deviations from microeconomic perfection, and generated a model that can more closely emulate the economic data on which they focus – predominantly the rates of real economic growth, employment and inflation. This became known as the 'New Keynesian' or saltwater approach to economics, in contrast to the 'New Classical' or freshwater approach: start from precisely the same vision of a macroeconomy that would be in perfect equilibrium with no involuntary unemployment if all consumers were homogeneous, markets were perfect and prices adjusted instantly to any shocks; then add in maybe two types of agents, imperfect competition and other deviations from perfection to generate inflation and involuntary unemployment.

The founding editor of the American Economic Association's specialist macroeconomics journal, Olivier Blanchard (Blanchard 2008, 2009) described the basic or 'toy' saltwater/New Keynesian' model as starting

from the freshwater/New Classical model without capital, to which it added two 'imperfections': monopolistic competition and inflation caused expectations of future inflation plus a gap between what output actually is and the higher level that neoclassical theory says it would be if there were no 'imperfections.'[36] It then added monetary policy conducted by a central bank using the Taylor Rule,[37] with which it attempts to control inflation by setting the real interest rate on the basis of the rate of inflation and the output gap.

This results in a model that can be expressed in three equations – one for consumption or aggregate demand as a function of the real interest rate and (rationally) expected future output, another for inflation, and a third for the central bank's interest-rate-setting policy. Blanchard stated that the model was

> simple, analytically convenient, and has largely replaced the IS-LM model as the basic model of fluctuations in graduate courses (although not yet in undergraduate textbooks). Similar to the IS-LM model, it reduces a complex reality to a few simple equations. Unlike the IS-LM model, it is formally, rather than informally, derived from optimization by firms and consumers. (Blanchard 2009: 214–15)

The weaknesses in the model[38] are addressed by adding yet more microeconomic imperfections. These include adding the reality that the labor market is not homogeneous to explain involuntary unemployment – 'One striking (and unpleasant) characteristic of the basic NK model is that there is no unemployment!' (ibid.: 216) – and using the concept of asymmetric information to explain problems in credit markets. This saltwater approach necessarily achieved a better fit to the data than the extreme neoclassical vision of the freshwater faction, but for reasons that are hardly exemplary, as Solow observed:

> The simpler sort of RBC model that I have been using for expository purposes has had little or no empirical success, even with a very undemanding notion of 'empirical success.' As a result, some of the freer spirits in the RBC school have begun to

> loosen up the basic framework by allowing for 'imperfections' in the labor market, and even in the capital market […]
>
> The model then sounds better and fits the data better. *This is not surprising: these imperfections were chosen by intelligent economists to make the models work better* […] (Solow 2001: 26; emphasis added)

Nonetheless, the better apparent fit to the data from models engineered to do so by the saltwaters meant that, over time, and despite the vigorous protests of the freshwaters, the 'New Keynesian' approach became the dominant one within neoclassical macroeconomics. It appeared to neoclassicals that macroeconomics was converging on a 'New Keynesian consensus,' and Blanchard claimed so in 2008:

> there has been enormous progress and substantial convergence. For a while – too long a while – the field looked like a battlefield. Researchers split in different directions, mostly ignoring each other, or else engaging in bitter fights and controversies. Over time however, largely because facts have a way of not going away, a largely shared vision both of fluctuations and of methodology has emerged. Not everything is fine. Like all revolutions, this one has come with the destruction of some knowledge, and it suffers from extremism, herding, and fashion. But none of this is deadly. The state of macro is good […]
>
> Facts have a way of eventually forcing irrelevant theory out (one wishes it happened faster), and good theory also has a way of eventually forcing bad theory out. The new tools developed by the new-classicals came to dominate. The facts emphasized by the new-Keynesians forced imperfections back in the benchmark model. A largely common vision has emerged. (Blanchard 2009: 210)

Given the time lags involved in academic publishing, this unfortunate paper, which was first completed in August 2008 (Blanchard 2008) (eight months *after* the start of the Great Recession, according to the National

Bureau of Economic Research), was published in an academic journal in May 2009, by which time the world as neoclassical economists thought they knew it had come to an end. Forces their models completely ignored finally overwhelmed the economy, and took their vision of the economy with it.

Conclusion

Though I can argue about logical fallacies till the cows come home, this is no substitute for an empirical proof that neoclassical economics is wrong. This was provided in spectacular fashion by the Great Recession. Not only was this not predicted by neoclassical models – according to them, such an event could not even happen.

Box 10.1 The Taylor Rule

The Taylor Rule was first devised by John Taylor as a reasonable empirical approximation to the way the Federal Reserve had in fact set nominal interest rates (Taylor 1993: 202). He noted that the Fed had increased the cash rate by 1.5 percent for every percent that inflation exceeded the Fed's target inflation rate, and reduced the cash rate by 0.5 percent for every percent that real GDP was below the average for the previous decade. When New Keynesian economists incorporated this in their model, they introduced the neoclassical concept of an 'equilibrium' real rate of interest (which is unobservable), so that if actual inflation and the rate of growth were equal to their target levels, the cash rate should be equal to the inflation rate plus this unobservable 'equilibrium' rate.

After the crisis hit, Taylor himself blamed it on the Fed deviating from his rule:

> Why did the Great Moderation end? In my view, the answer is simple. The Great Moderation ended because of a 'Great Deviation,' in which economic policy deviated from what was working well during the Great Moderation. Compared with the Great Moderation, policy became more interventionist, less rules-based, and less predictable. When policy deviated from what was

> working well, economic performance deteriorated. And lo and behold, we had the Great Recession. (Taylor 2007: 166)

There is some merit in Taylor's argument – certainly the low rates in that period encouraged the growth of Ponzi behavior in the finance sector. But his neoclassical analysis ignores the dynamics of private debt, which, as I show in Chapters 12 and 13, explain both the 'Great Moderation' and the 'Great Recession.' Taylor's Rule was more of a statistical coincidence in this period than a reason for the stability prior to the recession.

The Rule also evidently gave Taylor no inkling that a crisis was imminent, since as late as 10 September 2007, he concluded a speech on his Rule with the following statement:

> Of course, we live in a fluid economic world, and we do not know how long these explanations or predictions will last. I have no doubt that in the future – and maybe the not so distant future – a bright economist – maybe one right in this room – will show that some of the explanations discussed here are misleading, or simply wrong. *But in the meantime, this is really a lot of fun*. (Ibid.: 15; emphasis added)

The economic crash of the Great Recession was accompanied by the crash of both the stock market and the housing market, and predictably the neoclassical theory of finance – known as the Efficient Markets Hypothesis – also argued that asset market crashes couldn't happen. In the next chapter, we'll take a diversion to the world of asset markets before returning to the key empirical fact that neoclassical economists were the last people on the planet to see the Great Recession coming.

Postscript: intellectual miasma or corporate corruption?

The extent to which economic theory ignored crucial issues like the distribution of wealth and the role of power in society leads many to extend a conspiracy theory explanation of how economics got into this state. Surely, they argue, economic theory says what the wealthy want to hear?

I instead lay the focus upon the teleological vision to which economists have been committed ever since Adam Smith first coined the phrase 'an invisible hand' as an analogy to the workings of a market economy. The vision of a world so perfectly coordinated that no superior power is needed to direct it, and no individual power sufficient to corrupt it, has seduced the minds of many young students of economics. I should know, because I was one; had the Internet been around when I was a student, someone somewhere would have posted an essay I wrote while in my first year as an undergraduate, calling for the abolition of both unions and monopolies. No corporation paid me a cent to write that paper (though now, if it could be found, I would happily pay a corporation to hide it!).

What enabled me to break away from that delusional analysis was what Australians call 'a good bullshit detector.' At a certain point, the fact that the assumptions needed to sustain the vision of the Invisible Hand were simply absurd led me to break away, and to become the critic I am today.

However, the corporate largesse interpretation of why neoclassical economics has prospered does come into play in explaining why neoclassical economics became so dominant. Many of the leading lights of US academic economics have lived in the revolving door between academia, government and big business, and in particular big finance. The fact that their theories, while effectively orthogonal to the real world, nonetheless provided a smokescreen behind which an unprecedented concentration of wealth and economic power took place, make these theories useful to wealthy financiers, even though they are useless – and in fact outright harmful – to capitalism itself.

The fact that both government and corporate funding has helped the development of these theories, while non-orthodox economists like me have had to labor without research grants to assist them, is one reason why the nonsense that is neoclassical economics is so well developed, while its potential rivals are so grossly underdeveloped.

The corporate dollar may also have played a role in enabling neoclassical economists to continue believing arrant nonsense as they developed their theories. So while I don't explain neoclassical theory on the basis of it serving the interests of the elite, the fact that it does – even though it is counterproductive for the economy itself – and that the

corporate and particularly financial elite fund those who develop it surely has played a role.

On this note, the website LittleSis (http://littlesis.org/) is well worth consulting. It documents the links between business and government figures in the USA, and leading neoclassical economists like Larry Summers feature prominently (see http://blog.littlesis.org/2011/01/10/evidence-of-an-american-plutocracy-the-larry-summers-story/).

11 | THE PRICE IS NOT RIGHT

Why finance markets can get the price of assets so badly wrong

In the first edition of this book, this chapter began with the following paragraphs:

> The Internet stock market boom was[1] the biggest speculative bubble in world history.
>
> Other manias have involved more ridiculously overvalued assets, or more preposterous objects of speculation – such as the tulip craze in 17th century Holland, the South Sea Bubble and Mississippi Bubble of the 18th century, or Japan's 1980s Bubble Economy speculation over Tokyo real estate. But no other bubble – not even the 'Roaring Twenties' boom prior to the Great Depression – has involved so many people, speculating so much money, in so short a time, to such ridiculous valuations.
>
> But of course, an economist wouldn't tell you that. Instead, economists have assured the world that the stock market's valuations reflect the true future prospects of companies. The most famous – and fatuous – such assurance is given in Dow 36,000, which its authors were defending even when the Dow had officially entered a correction from its all-time high of March 2000, and the Nasdaq was firmly in bear market territory (*Time*, 22 May 2000: 92–93). The mammoth valuations, argued Hassett and Glassman, were simply the product of investors reassessing the risk premiums attached to stocks, having realized that over the long term, stocks were no riskier than bonds.

> Economists were similarly reassuring back in 1929, with the most famous such utterance being Irving Fisher's comment that:
>
> Stock prices have reached what looks like a permanently high plateau. I do not feel that there will soon, if ever, be a fifty or sixty point break below present levels, such as Mr. Babson has predicted. I expect to see the stock market a good deal higher than it is today within a few months. (Irving Fisher, *New York Times*, 15 October 1929)
>
> This was published less than two weeks before 'Black Monday,' 28 October 1929, when the Dow Jones Industrial Average closed 12.8% below its previous level, and fell another 11.7% the following day. In just 15 days of wild gyrations from the day of Fisher's comments, the market fell over 120 points (from a level of about 350): twice as far as even Fisher's bearish rival Babson had predicted, and twice as much as Fisher had believed would ever be possible. Three years later, the stock market indices had fallen 90%, and many a once-rich speculator was bankrupt. Investors who trusted economists back then lost their shirts. Trusting souls who accept economic assurances that markets are efficient are unlikely to fare any better this time when the Bull gives way to the Bear.

At the time, I thought that the DotCom Bubble would be the last of the big asset bubbles. I couldn't envisage then that any other asset market could ever be more overvalued. I couldn't imagine that any more preposterous object of speculation could emerge than the plethora of 'DotCom' companies with negative cash flows and over-the-top valuations that lit up the Super Bowl in 2000, and had burnt their investors' money into oblivion months later.

Silly me: I had obviously underestimated the inventiveness of Wall Street. Even as the Nasdaq crashed and burnt, Wall Street had found an even more ridiculous way to entice the public into debt: the fantasy that money could be made by lending money to people with a history of not repaying debt. The Subprime Bubble was born. By the time it burst, the academic sub-discipline of Finance was finally starting to concede that its

model of how asset markets operate was seriously wrong. But by then, it was too late.

The kernel

> 'There's glory for you!'
>
> 'I don't know what you mean by "glory,"' Alice said.
>
> Humpty Dumpty smiled contemptuously. 'Of course you don't – till I tell you. I meant "there's a nice knock-down argument for you!"'
>
> 'But "glory" doesn't mean "a nice knock-down argument,"' Alice objected.
>
> 'When I use a word,' Humpty Dumpty said in rather a scornful tone, 'it means just what I choose it to mean – neither more nor less.'

All sciences invent their own language, just as Lewis Carroll's famous egghead invented his own meanings for words. Many sciences harness words which are in are common usage, but give them a quite different technical meaning. But no other science plays so fast and loose with the English language as economics.

Physics, for example, calls the fundamental constituents of matter 'strings.' This isn't particularly confusing, since it's obvious that physicists don't believe that a length of yarn is the basic unit of matter.

However, when economists call stock markets 'efficient,' the usage is nowhere near as clear cut. A colloquial meaning of efficient is 'does things quickly with a minimum of waste,' and it's clear that this meaning can apply to modern, computerized, Internet-accessible bourses. Thus it often seems reasonable to the public that economists describe finance markets as 'efficient.'

However, when economists say that the stock market is efficient, they mean that they believe that stock markets accurately price stocks on the basis of their unknown future earnings. That meaning shifts 'efficient' from something which is obvious to something which is a debatable proposition.

But that's not the end of the story, because to 'prove' that markets are efficient in this sense, economists make three bizarre assumptions:

- that all investors have identical expectations about the future prospects of all companies;
- that these identical expectations are correct; and
- that all investors have equal access to unlimited credit.

Clearly, the only way these assumptions could hold would be if each and every stock market investor were God. Since in reality the stock market is inhabited by mere mortals, there is no way that the stock market can be efficient in the way that economists define the term. Yet economists assert that stock markets are 'efficient,' and dismiss criticism of these assumptions with the proposition that you can't judge a theory by its assumptions. As Chapter 7 showed, this defense is bunk.

In a way, it's fitting that Lewis Carroll put those words in Humpty Dumpty's mouth, rather than equally appropriate vessels such as the Mad Hatter, or the Red Queen. Humpty Dumpty, after all, had a great fall …

The roadmap

The chapter begins by considering the development over time of the prevailing attitude to finance, starting with the medieval prohibition against the lending of money at interest, and culminating in economists treating the lending of money as no different from any other commodity exchange. The main economist responsible for the economic theory of lending was Irving Fisher, who, as just mentioned, effectively went bankrupt during the depression by following his own theories. However, he subsequently developed a quite different theory, which argued that excessive debt and falling prices could cause depressions. After outlining this theory, I consider the modern theory of finance known as the 'efficient markets hypothesis.' The validity of the assumptions needed to buttress this theory is assessed in the light of the logic outlined in Chapter 7. Since these are domain assumptions, the theory is inapplicable in the real world, so that markets cannot possibly be 'efficient' as economists define the term.

Fisher on finance: from reassuring oracle to ignored Cassandra

Irving Fisher was one of the many victims of the Great Crash of 1929, losing a fortune worth over $100 million in today's dollars, and being reduced to penury.[2] But his greater loss, in many ways, was one of prestige. Before this infamous utterance, he was easily America's most respected and famous economist, renowned for developing a theory of money that explained the valuation of financial assets. After it, he was a pariah.

This was a great pity, because in the depths of the Great Depression, he developed an explanation of how financial speculation could lead to economic collapse. However, this new theory – which rejected many of the assumptions of his previous model of finance – was ignored. Instead, Fisher's pre-Great Depression theory of finance continued as the economic theory of how asset prices are determined.

Decades later, Fisher's 'Debt Deflation Theory of Great Depressions' was rediscovered by the non-orthodox economist Hyman Minsky, while at much the same time Fisher's pre-Great Depression theory was formalized into the efficient markets hypothesis. Fisher thus has the dubious distinction of fathering both the conventional theory of finance – which, like his 1929 self, reassures finance markets that they are rational – and an unconventional theory which argues that speculative bubbles can cause economic depressions.

Pre-Depression Fisher: the time value of money In 1930 Fisher published *The Theory of Interest*, which asserted that the interest rate 'expresses a price in the exchange between present and future goods' (Fisher 1930).[3] This argument was a simple extension of the economic theory of prices to the puzzle of how interest rates are set, but it has an even older genealogy: it was first argued by Jeremy Bentham, the true father of modern neoclassical economics, when in 1787 he wrote 'In defence of usury.'

Never a lender nor a borrower be … Today, usury means lending money at an exorbitant rate of interest; in antiquity, it meant lending money at any rate of interest at all. However, the medieval objection was not to the rate of interest itself, but to the lender's desire to profit from a venture without sharing in its risks. A usurious contract was one in which the lender was guaranteed a positive return, regardless of whether the borrower's venture

succeeded or failed: 'The primary test for usury was whether or not the lender had contracted to lend at interest without assuming a share of the risk inherent to the transaction. If the lender could collect interest regardless of the debtor's fortunes he was a usurer' (Jones 1989).

As trade came to play a larger role in society, the prohibitions against usury were weakened, and the legal definition was modified to match today's colloquial meaning.[4] By Bentham's time, the legal definition referred to a rate of interest greater than 5 percent.

Adam Smith supported this legal limit. Smith argued that the complete prohibition, 'like all others of the same kind, is said to have produced no effect, and probably rather increased than diminished the evil of usury' (Smith 1838 [1776]). However, Smith supported the concept of a legal limit to the rate of interest set just above the going market rate,[5] because such a limit actually improved the allocation of the country's credit. The advantage of a legal limit, according to Smith, was that when set properly it excluded only loans to 'prodigals and projectors,' thus making more of the country's capital available for loan to industrious people:

> The legal rate […] ought not to be much above the lowest market rate. If the legal rate of interest in Great Britain, for example, was fixed so high as eight or ten per cent, the greater part of the money which was to be lent would be lent to prodigals and projectors, who alone would be willing to give this high interest […] A great part of the capital of the country would thus be kept out of the hands which were most likely to make a profitable and advantageous use of it, and thrown into those which were most likely to waste and destroy it. Where the legal rate of interest, on the contrary, is fixed but a very little above the lowest market rate, sober people are universally preferred, as borrowers, to prodigals and projectors. The person who lends money gets nearly as much interest from the former as he dares to take from the latter, and his money is much safer in the hands of the one set of people than in those of the other. A great part of the capital of the country is thus thrown into the hands in which it is most likely to be employed with advantage. (Ibid.)

In defence of usury Bentham's rejoinder to Smith's arguments may well have set the standard for fanciful and specious reasoning to which modern economics has since aspired.

Smith referred to two classes of borrowers who could be expected to accede to rates of interest substantially above the lowest market rate: 'prodigals and projectors.' The former are individuals who would waste the money on conspicuous consumption; the latter are those who promote ill-conceived schemes to the public, which result in inappropriate investment. Smith's case in favor of a legal ceiling to the rate of interest thus had both a 'microeconomic' and a 'macroeconomic' aspect.

Macroeconomics was Smith's key concern: encouraging 'prodigals and projectors' would result in 'a great part of the capital of the country' being thrown into the hands of 'those which were most likely to waste and destroy it.' The ceiling, by removing the incentive to lend to such borrowers, would result in a higher overall quality of investment, and thus higher growth.

Bentham's riposte ignored macroeconomics. Instead, it began from the microeconomic and libertarian presumption that 'no man of ripe years and of sound mind, acting freely, and with his eyes open, ought to be hindered, with a view to his advantage, from making such bargain, in the way of obtaining money, as he thinks fit' (Bentham 1787).

He initially conceded that the restraint of prodigal behavior may give grounds for setting a ceiling to the rate of interest, only to then argue that in practice a prodigal would not be charged an exorbitant rate of interest. He began with the proposition that 'no man […] ever thinks of borrowing money to spend, so long as he has ready money of his own, or effects which he can turn into ready money without loss.' Secondly, the exceptions to the above rule who have the requisite collateral can get a loan at the usual rate. Thirdly, those who do not have security will only be lent to by those who like them, and these friendly persons will naturally offer them the standard rate: 'Persons who either feel, or find reasons for pretending to feel, a friendship for the borrower, can not take of him more than the ordinary rate of interest: persons who have no such motive for lending him, will not lend him at all' (ibid.).

If Bentham were to be believed, the friendly bank manager of the 1950s had many a precursor in eighteenth-century Britain, while the rapacious Shylock perished with Shakespeare in the seventeenth.

A bit of empirical research would have revealed that, though rates of interest had fallen dramatically as finance became institutionalized, there was no shortage of lenders willing to hand prodigals ready cash at high rates of interest, in return for ownership of their assets should they go bankrupt. But Bentham's more important sleight of mind was to ignore the macroeconomic argument that the legislative ceiling to the rate of interest improved the overall quality of investment by favoring 'sober people' over 'prodigals and projectors.'

The historical record favored Smith. The seventeenth, eighteenth and nineteenth centuries are awash with examples of projectors promoting fantastic schemes to a gullible public. The most famous have entered the folklore of society: the Tulip Mania, the South Sea Bubble, the Mississippi Land Scheme (Mackay 1841). What has not sunk in so deeply is that the financial panics that occurred when these bubbles burst frequently ruined whole countries.[6]

However, the tide of social change and the development of economic theory favored Bentham. The statutes setting maximum rates were eventually repealed, the concept of usury itself came to be regarded as one of those quaint preoccupations of a more religious age, and modern economics extended Bentham's concept that 'putting money out at interest, is exchanging present money for future' (Bentham 1787). Of course, the magnificent edifice economists built upon Bentham's morsel assumed that everything happened in equilibrium.

The time value of goods In keeping with the economic belief that the economy is fundamentally a barter system, in which money is merely a lubricant, Fisher restated Bentham's concept in terms of goods, rather than money: the rate of interest 'expresses a price in the exchange between present and future goods' (Fisher 1930).

Fisher's model had three components: the subjective preferences of different individuals between consuming more now by borrowing, or consuming more in the future by forgoing consumption now and lending instead; the objective possibilities for investment; and a market which reconciled the two.

From the subjective perspective, a lender of money is someone who, compared to the prevailing rate of interest, has a low time preference for

present over future goods. Someone who would be willing to forgo $100 worth of consumption today in return for $103 worth of consumption next year has a rate of time preference of 3 percent. If the prevailing interest rate is in fact 6 percent, then by lending out $100 today, this person enables himself to consume $106 worth of commodities next year, and has clearly made a personal gain. This person will therefore be a lender when the interest rate is 6 percent.

Conversely, a borrower is someone who has a high time preference for present goods over future goods. Someone who would require $110 next year in order to be tempted to forgo consuming $100 today would decide that, at a rate of interest of 6 percent, it was worth his while to borrow. That way, he can finance $100 worth of consumption today, at a cost of only $106 worth of consumption next year. This person will be a borrower at an interest rate of 6 percent.

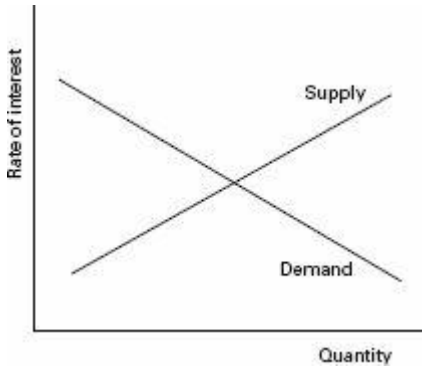
The act of borrowing is thus a means by which those with a high preference for present goods acquire the funds they need now, at the expense of some of their later income.

Individual preferences themselves depend in part upon the income flow that an individual anticipates, so that a wealthy individual, or someone who expects income to fall in the future, is likely to be a lender, whereas a poor individual, or one who expects income to rise in the future, is likely to be a borrower.

At a very low rate of interest, even people who have a very low time preference are unlikely to lend money, since the return from lending would be below their rate of time preference. At a very high rate of interest, even those who have a high time preference are likely to be lenders instead, since the high rate of interest would exceed their rate of time preference. This relationship between the rate of interest and the supply of funds gives us an upward-sloping supply curve for money.

The objective perspective reflects the possibilities for profitable investment. At a high rate of interest, only a small number of investment projects will be expected to turn a profit, and therefore investment will be low. At a low rate of interest, almost all projects are likely to turn a profit over financing costs, so the demand for money will be very high. This relationship between the interest rate and the demand for money gives us a downward-sloping demand curve for money.

The market mechanism brings these two forces into harmony by yielding the equilibrium rate of interest.



11.1 Supply and demand in the market for money

Economics, it appears, is back in familiar territory. But there are some special, time-based nuances to the credit market. In the goods market, transactions occur immediately: one bundle of goods today is exchanged for another bundle of goods today. However, in the credit market, the 'purchaser' (the company offering an investment opportunity) takes immediate delivery of the loan, but repays principal and interest in installments over time. Ancillary assumptions were therefore required to stretch the standard static vision of the market to the time-based creature that credit really is. These additional assumptions, in Fisher's words, were: '(A) The market must be cleared – and cleared with respect to every interval of time. (B) The debts must be paid' (ibid.).

Fisher saw nothing wrong with these ancillary assumptions, until he and countless others personally violated them during the Great Depression.

Fisher during the Crash: 'don't panic'

To his credit, Fisher's response to the Great Depression was worthy of Keynes's apocryphal statement that 'when the facts prove me wrong, I change my mind.' But at first Fisher clung to his pre-Crash optimism that the American economy was fundamentally sound, that a wave of invention had introduced a new era of higher productivity, that the new medium of radio would revolutionize business. It all sounds so familiar today …

A new era … Fisher's comments to a bankers' forum on 'Black Wednesday' – 23 October, when stocks fell by an unprecedented 6.3 percent in one day – confirm the old adage that 'the more things change, the more they remain the same.' Every factor that Fisher then thought justified the stock market's bull run has its counterpart today: it was 'a new era,' a wave of invention (read 'the Internet') justified high valuations, stable prices reduced the uncertainty of share ownership, stocks were better long-term investments than bonds, investment trusts (read 'mutual funds') enabled much more intelligent stock selection, a debt-financed consumer boom was natural when a great increase in income was rationally anticipated.

Fisher first recounted the ways in which the 1929 stock market boom was remarkable. Shares had doubled in value since 1926, and any investor who had 'followed the herd' and bought and sold shares simply on the basis of their popularity would have increased his wealth tenfold in those three years. Stock prices had risen so much that dividend yields were below bond yields. Brokers' loans – effectively margin call lending – were at their highest level in history. All these observations supported the notion that the market 'seems too high and to warrant a major bear movement' (Fisher 1929).

However, he then gave four reasons why the 1929 valuations were sensible: changed expectations of future earnings, reinvestment of dividends, a change in risk premiums, and a change in the way in which future income is discounted.

He supported the first argument with the statement that

> We are now applying science and invention to industry as we never applied it before. *We are living in a new era*, and it is of the utmost importance for every businessman and every banker to understand this new era and its implications […] All the resources of modern scientific chemistry, metallurgy, electricity, are being utilized – for what? To make big incomes for the people of the United States in the future, to add to the dividends of corporations which are handling these new inventions, and necessarily, therefore, to raise the prices of stocks which represent shares in these new inventions. (Ibid.; emphasis added)

This wave of invention, with its return in years yet to come, meant that it was quite natural for the ratio of share price to historic earnings to rise. In fact, these new firms should be expected to make losses as they established their new inventions: 'In the airline industry very little attention is paid to the earnings today, because the price of the stock is purely a speculation on the far bigger returns that are expected in the future. Any new invention […] at first does not give any profit […]' (ibid.).

Low inflation also played a role in high stock valuations, since a stable price level gives 'an immense impulse towards prosperity' (ibid.).

The second factor, the reinvestment of dividends, was a positive force since firms that did this – rather than handing dividends back to investors – were able to grow more rapidly. Hence 'many of the stocks which sell the highest on the stock exchange and which have had the most spectacular rise are not paying any dividends' (ibid.).

The third reason, a change in the way the public estimates risk, occurred because Edgar Smith's influential book *Common Stocks as Long Term Investments* had shown that over the longer term stocks outperformed bonds. As a result, '[t]here has been almost a stampede towards stocks, and away from bonds' (ibid.; in the late 1990s, Hassett and Glassman's *Dow 36,000* and its ilk spread the same delusion).

This movement had led to the establishment of the new profession of investment counseling, and then the new institution of investment trusts, which 'can afford to make studies of stocks which the individual investor could not study' (ibid.). As well as diversifying and spreading risk, these institutions enabled stocks to be scientifically selected. This explained why 'our stock market is highly selective today,' and as a result Fisher wasn't troubled by the fact that: 'Half of the stocks during the last year have fallen in spite of the fact that the average as shown by the index numbers had risen. The leaders are becoming fewer and fewer, and those stocks that are leading have a greater and greater scarcity value' (ibid.).

Fisher conceded that rank speculation played some role in the market, but he blamed this 'lunatic fringe' more for the crash in stock prices than for its run-up over the preceding four years: 'There is a certain lunatic fringe in the stock market, and there always will be whenever there is any successful bear movement going on […] they will put the stocks up above

what they should be and, when frightened, […] will immediately want to sell out' (ibid.).

This speculative fringe ranked fifteenth out of Fisher's fifteen determinants of the level of stock prices, though he was not so confident of his ranking after the market's 6 percent fall on Black Wednesday. Nonetheless, he still argued that 'the other fourteen causes are far more important than this one cause itself.' He acknowledged that most speculation took place with borrowed money – a theme that would later become his *bête noire*. But he argued that most of this money had been borrowed to finance consumption today – rather than just rank speculation – because consumers were simply cashing in on rationally anticipated future increases in income:

To a certain extent it is normal that during an era such as we are now passing through, where the income of the people of the United States is bound to increase faster perhaps than ever before in its history, and it has during the last few years increased amazingly, that we should try to cash in on future income in advance of its occurring, exactly on the principle that when a young man knows he has been given unexpectedly a large bequest, and that it will be in his hands inside a year, he will borrow against it in advance. In other words, there ought to be a big demand for loans at a high rate of interest during a period of great increase in income. (Ibid.)

He concluded with an expectation that the market's 12 percent fall in the preceding eight days was an aberration:

Great prosperity at present and greater prosperity in view in the future […] rather than speculation […] explain the high stock markets, and when it is finally rid of the lunatic fringe, the stock market will never go back to 50 per cent of its present level […] We shall not see very much further, if any, recession in the stock market, but rather […] a resumption of the bull market, not as rapidly as it has been in the past, but still a bull rather than a bear movement. (Ibid.)

Fisher after the Crash: the debt-deflation hypothesis Fisher was, of course, profoundly wrong, and at great personal cost to himself. The market receded 90 percent from its peak, and the index did not regain its 1929 level for a quarter of a century.[7] As the Crash persisted, the slump deepened into the Great Depression, with, at its nadir, over 25 percent of America's

workers unemployed. Fisher's personal fortune evaporated, and his perspective on the American financial system shifted from one of confidence to one of alarm.

He eventually developed a radically different analysis of finance, one in which his ancillary assumptions in *The Theory of Interest* – that 'the market must be cleared, and cleared with respect to every interval of time' and that 'The debts must be paid' – were systematically violated. Now he acknowledged that the market was never in equilibrium, and that debts could fail to be repaid, not just individually but en masse. Static reasoning gave way to an analysis of the dynamic forces which could have caused the Great Depression.

Whereas he had previously assumed that the economy was always in equilibrium, now he appreciated that even if the real economy actually momentarily reached equilibrium, this state would be short lived since 'new disturbances are, humanly speaking, sure to occur, so that, in actual fact, any variable is almost always above or below the ideal equilibrium' (Fisher 1933).

Equilibrium was also likely to be precarious. Whereas beforehand he had simply taken it for granted that equilibrium was stable, now he realized that equilibrium, 'though stable, is so delicately poised that, after departure from it beyond certain limits, instability ensues.' A slight movement away from equilibrium could set in train forces that would drive the economy even farther away, rather than returning it to balance.

While any of a multitude of factors could, according to Fisher, push the system away from equilibrium, the crucial ingredient needed to turn this limited instability into a catastrophic collapse was an excessive level of debt, where 'the breaking of many debtors constitutes a "crash," after which there is no coming back to the original equilibrium.'

He ventured the opinion that the 'two dominant factors' that cause depressions are 'over-indebtedness to start with and deflation following soon after.' Though other factors are important, debt – the entry into a contractual obligation to repay principal with interest – and a falling price level are crucial:

Thus over-investment and over-speculation are often important; but they would have far less serious results were they not conducted with borrowed money. That is, over-indebtedness may lend importance to over-investment

or to over-speculation. The same is true as to over-confidence. I fancy that over-confidence seldom does any great harm except when, as, and if, it beguiles its victims into debt. (Ibid.)

The final sentence in this quote is rather poignant, since Fisher himself was a classic instance of someone whom overconfidence had beguiled into debt.[8]

Overconfidence leads investors to overestimate the prospective gain from investment, or to underestimate the risks, and thus commit themselves to an unsustainable level of debt. In either case, the investor commits funds well beyond the level which returns an optimum gain. Such overconfidence is an inevitability in the real world because, as noted above, all real-world variables are bound to be either above or below their ideal equilibrium values.

A chain reaction then ensues that can tip the economy into depression. It begins with distress selling, at severely reduced prices, driven by the need to cover debt repayments. Falling prices means that the real burden of debt actually rises, even as nominal debt is reduced, and the repayment of debts also reduces the money supply. These effects cause further bankruptcies, reducing profits, investment, output and employment. Pessimism rises, causing those with money to hoard it, which further reduces business activity. The falling price level also has the perverse effect that the real rate of interest rises even though nominal rates have fallen, and this drastically reduces investment.

Fisher's theory was thus an alternative explanation of the Great Depression to both Keynes's rejection of Say's Law and Hicks's 'liquidity trap' (discussed in Chapter 9). But though the chain reaction argument is plausible, Fisher provided no formal proof for it – in contrast to his previous emphasis upon formal mathematical reasoning. Partly for this reason, his thesis was received poorly by the economics profession, and his insights were swamped by the rapid adoption of Hicks's IS-LM analysis after the publication of Keynes's *General Theory*.[9]

After the Great Depression, economists continued to cite his pre-Crash work on finance, while his debt-deflation theory was largely ignored.[10] As a result, the antipathy he saw between the formal concept of equilibrium and the actual performance of asset markets was also ignored. Equilibrium once

again became the defining feature of the economic analysis of finance. This process reached its zenith with the development of what is known as the 'efficient markets hypothesis.'

The efficient markets hypothesis

Non-economists often surmise that the term 'efficient' refers to the speed at which operations take place on the stock market, and/or the cost per transaction. Since the former has risen and the latter fallen dramatically with computers, the proposition that the stock market is efficient appears sensible. Market efficiency is often alleged to mean that 'investors are assumed to make efficient use of all available information,' which also seems quite reasonable.

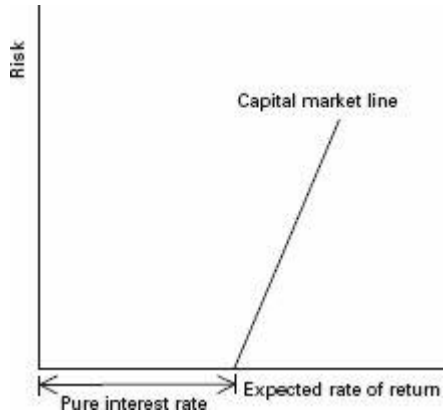
However, the economic concept of efficiency means something quite different from the normal parlance. In the case of the stock market, it means at least four things:

- that the collective expectations of stock market investors are accurate predictions of the future prospects of companies;
- that share prices fully reflect all information pertinent to the future prospects of traded companies;
- that changes in share prices are entirely due to changes in information relevant to future prospects, where that information arrives in an unpredictable and random fashion; and
- that therefore stock prices 'follow a random walk,' so that past movements in prices give no information about what future movements will be – just as past rolls of dice can't be used to predict what the next roll will be.

These propositions are a collage of the assumptions and conclusions of the 'efficient markets hypothesis' (EMH) and the 'capital assets pricing model' (CAPM), which were formal extensions to Fisher's (pre-Depression) time value of money theories. Like the Fisher theories of old, these new theories were microeconomic in nature, and presumed that finance markets are continually in equilibrium. There were several

economists who developed this sophisticated equilibrium analysis of finance. In what follows I s on the work of W. F. Sharpe.

Risk and return It seems reasonable, a priori, to argue that an asset that gives a high return is likely to be riskier than one that gives a lower return. If an investor wants complete safety, then he can invest in government bonds. If a higher rate of return is desired, then he can invest in corporate bonds, or shares. The former hold the risk of default, while the latter can rise or fall unpredictably in price, and do not have a guaranteed income flow. Therefore there is a 'trade-off' between return and risk: a higher return can be earned, but only at the cost of a higher level of risk.



11.2 The capital market line

Sharpe provided an explanation for this in terms of the theory of individual behavior discussed in Chapter 3. Once again, we find ourselves tobogganing up and down indifference curves.

The individual rational investor Sharpe began by assuming that 'an individual views the outcome of any investment in probabilistic terms; he is willing to act on the basis of […] expected value and standard deviation' (Sharpe 1964: 427–8). An investor gets greater utility from a higher return than a lower one, and lower utility from an asset with a high standard deviation than a lower one. This assumption enabled Sharpe to plot an investor's preferences in terms of indifference curves, with the two 'goods' being risk and return.

However, there was one twist compared to standard indifference curve analysis as outlined in Chapter 3. Risk is a 'bad,' not a 'good' – and a

consumer maximizes his utility by experiencing as little risk as possible. So the most desirable investment is one that gives a very high return with very little risk. Consequently, rather than being drawn to show that more of both goods is better, these indifference curves are drawn to show that more return and less risk is better.

With standard goods, the consumer prefers more of both, so the desirable direction to move in on the indifference map is up and to the right – which means you feel better as you get more of both commodities. But with return and risk as the 'goods,' the desirable direction is more return and less risk. Sharpe drew expected return on the horizontal axis and risk on the vertical, so the most desirable direction was to the right – which gave you more return – and down – which gave you less risk. The highest utility comes from the highest return and the *lowest* risk.

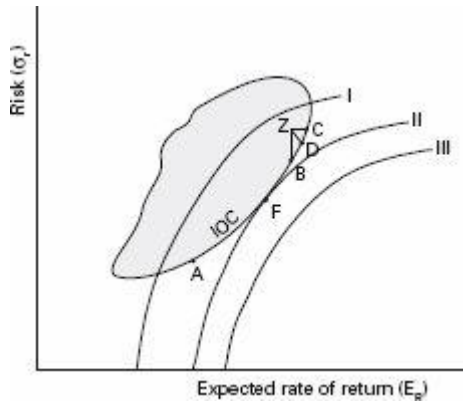
That takes care of the consumer's preferences. To complete the analysis, a budget line is needed as well – and here again there was a twist compared to the analysis of consumption. Rather than the budget line being the investor's income, the budget 'line' was the spectrum of investments that an investor could make. Each individual investment was a share in some company,[11] and all the information about them was reduced to their expected returns and the standard deviation of their expected returns. These could have any pattern at all – some investments would have a very high expected return and low variability, others a low expected return and high variability, and so on. Each company could then be described by a point on the graph of return versus risk, where the horizontal position was the return and the vertical position was the risk.

This resulted in a 'cloud' of possible investments that were potentially available to investors, where the most desirable investments were those with high return – the farther out along the horizontal axis, the better – and low risk – the lower down on the vertical axis, the better.

With this picture of investor behavior, Sharpe showed that the only investments that are rational for this investor are those that fall on the edge of the cloud of possible investments, which he labels the 'investment opportunity curve' or IOC (ibid.: 429). These investments give the highest return and the lowest risk possible. Any other combination that is not on the

edge of the cloud can be topped by one farther out that has both a higher return and a lower risk.[12]

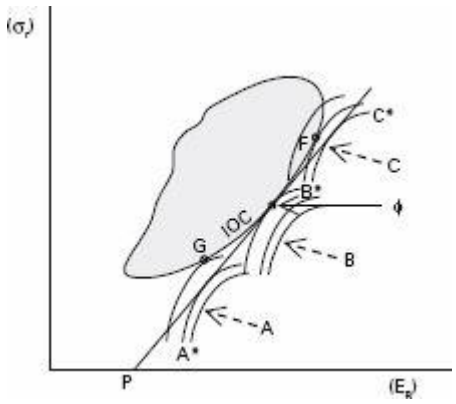
If this were the end of the matter, then the investor would choose the particular combination that coincided with their preferred risk–return trade-off, and that would be that.



11.3 Investor preferences and the investment opportunity cloud

However, it's possible to combine share-market investments with a bond that has much lower volatility, and Sharpe assumed the existence of a bond that paid a very low return, but had no risk. Sharpe linked bond and share investments with one further assumption: *that the investor could borrow as much as he wanted at the riskless rate of interest*. This assumption meant that, in Sharpe's model, an investor could invest some money in the riskless (but low-return) bond, and some money in risky (but higher-return) shares to create an investment portfolio.

This portfolio was represented by a straight line linking the riskless bond with a selection of shares (where the only selection that made sense was one that was on the Investment Opportunity Curve, and tangential to a line drawn through the riskless bond). Sharpe called this line the 'capital market line' or CML (ibid.: 425).



11.4 Multiple investors (with identical expectations)

With borrowing, the investor's risk–return preferences no longer determined which shares he bought; instead, they determined where he sat on the CML.

An ultra-conservative investor would just buy the riskless bond and nothing else: that would put him on the horizontal axis (where risk is zero) but only a short distance out along the horizontal axis – which means only a very low return. Someone who was happy with the market return – the return on an investment in shares alone – would buy only shares. Someone who wanted a *higher* return than shares provided could do so by borrowing money at the riskless rate and buying shares with this borrowed money as well as their own (in the real world, this is called buying shares on margin).

All together now? At this stage, Sharpe encountered a problem. As well as every investor having a different set of indifference curves between risk and return, each would also have a different opinion about the return and risk that would be associated with each possible investment. Thus investor C might think that investment F – say the Internet company Yahoo – was likely to yield a low return at high risk, while investor A might expect that Yahoo will give high returns with little variation.

In other words, each investor would perceive a different 'cloud' of investment opportunities. The edge of the cloud of investment opportunities, the IOC, would be different for every investor, in terms of both location and the investments in it.

Equally, lenders may charge a different rate of interest to every borrower, so that the location P would differ between individuals. They

might also restrict credit to some (or all) investors, so that the length of the line between each investor's P would differ – rather than being infinitely long, as Sharpe assumed. It might not even be a line, but could well be a curve, with lenders charging a higher rate of interest as borrowers committed themselves to more and more debt.

In other words, as with every neoclassical theorem, Sharpe encountered an aggregation problem in going from the isolated individual to the level of society. And, like every neoclassical economist, he took the time-honored approach of assuming the problem away. He assumed (a) that all investors could borrow or lend as much as they liked at the same rate, and (b) that investors all agreed on the expected prospects for each and every investment.

Sharpe admitted that these were extreme assumptions, but he justified them by an appeal to the methodological authority of 'assumptions don't matter' Milton Friedman. In Sharpe's words:

> In order to derive conditions for equilibrium in the capital market we invoke two assumptions. First, we assume a common pure rate of interest, with all investors able to borrow or lend funds on equal terms. Second, we assume homogeneity of investor expectations: investors are assumed to agree on the prospects of various investments – the expected values, standard deviations and correlation coefficients described in Part II. *Needless to say, these are highly restrictive and undoubtedly unrealistic assumptions*. However, since the proper test of a theory is not the realism of its assumptions but the acceptability of its implications, and since these assumptions imply equilibrium conditions which form a major part of classical financial doctrine, it is far from clear that this formulation should be rejected – especially in view of the dearth of alternative models leading to similar results. (Ibid.; emphasis added)

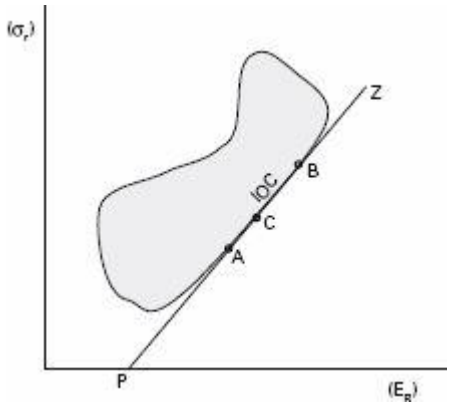
Though Sharpe doesn't explicitly say this, he also assumes that investor expectations are accurate: that the returns investors expect firms to achieve will actually happen.

With these handy assumptions under his belt, the problem was greatly simplified. The riskless asset was the same for all investors. The IOC was

the same for all investors. Therefore all investors would want to invest in some combination of the riskless asset and the same share portfolio. All that differed were investor risk–return preferences.

Some would borrow money to move farther 'northeast' (towards a higher return with higher risk) than the point at which their indifference map was tangential to the IOC. Others would lend money to move 'southwest' from the point of tangency between their indifference map and the IOC, thus getting a lower return and a lower risk.

Since all investors will attempt to buy the same portfolio, and no investors will attempt to buy any other investment, the market mechanism kicks in. This one portfolio rises in price, while all other investments fall in price. This process of repricing investments alters their returns, and flattens the edge of the IOC.



11.5 Flattening the IOC

The final step in Sharpe's argument relates the return on any single share to the overall market return, with a relation known these days as the share's 'beta.'[13] What this means in practice is that the efficient markets hypothesis asserts that the more volatile a share's returns are, the higher will be its expected yield. There is a trade-off between risk and return.

Sharpe's paper formed the core of the EMH. Others added ancillary elements – such as the argument that how a firm is internally financed has no impact on its value, that dividends are irrelevant to a share's value, and so on. If this set of theories were correct, then the propositions cited earlier would be true: the collective expectations of investors will be an accurate prediction of the future prospects of companies; share prices will fully reflect all information pertinent to the future prospects of traded

companies.[14] Changes in share prices will be entirely due to changes in information relevant to future prospects; and prices will 'follow a random walk,' so that past movements in prices give no information about what future movements will be.

Reservations The outline above covers the theory as it is usually presented to undergraduates (and victims of MBA programs), and as it was believed by its adherents among stockbrokers and speculators (of whom there are now almost none). But Sharpe was aware that it was unsatisfactory, mainly because of side effects from the assumptions that investors are in complete agreement about the future prospects of traded companies, and that all investors can borrow or lend as much as they want at the riskless rate of interest.

One obvious side effect of the first assumption is that, once equilibrium is reached, trade on the stock exchange should cease. Thereafter, any trading should merely be the result of the random arrival of new information, or the temporary disturbance of equilibrium via the floating of some new security. The trading profile of the stock market should therefore be like that of an almost extinct volcano.

Instead, even back in the 1960s when this paper was written, the stock market behaved like a very active volcano in terms of both price volatility and the volume of trades. It has become even more so since, and in 1987 it did a reasonable, though short-lived, impression of Krakatoa.

The second assumption implies that anyone could borrow sufficient money to purchase all the shares in, say, Microsoft, and pay no more than the riskless rate of interest to do it. This implies a degree of liquidity that is simply impossible in the real world.

Sharpe very honestly discussed both the reality of these assumptions, and the implications of dropping them. He readily conceded that 'even the most casual empiricism' suggests that the assumption of complete agreement is false: 'People often hold passionately to beliefs that are far from universal. The seller of a share of IBM stock may be convinced that it is worth considerably less than the sales price. The buyer may be convinced that it is worth considerably more' (Sharpe 1970). If this assumption is dropped, then in place of the single 'security market line,' and a spectrum of efficient investments which is the same for all investors, there is a

different security market line for each investor. The clean simplicity of the EMH collapses.

The assumption that we can all borrow (or lend) as much as we like at the riskless rate of interest is just as unrealistic as the assumption that all investors agree. Sharpe concedes that the theory collapses once one accepts the reality that the borrowing rate normally exceeds the lending rate, that investors are credit rationed, and that the borrowing rate tends to rise as the amount being borrowed increases:

> The consequence of accommodating such aspects of reality are likely to be disastrous in terms of the usefulness of the resulting theory […] The capital market line no longer exists. Instead, there is a capital market curve – linear over some ranges, perhaps, but becoming flatter as [risk] increases over other ranges. Moreover, there is no single optimal combination of risky securities; the preferred combination depends upon the investors' preferences […] The demise of the capital market line is followed immediately by that of the security market line. The theory is in a shambles. (Ibid.)

But in the end, faced with a choice between an unrealistic theory and no theory at all, Sharpe opts for theory. His comfort in this choice continues to be Milton Friedman's methodological escape route that the unrealism of assumptions 'is not important in itself. More relevant, the implications are not wildly inconsistent with observed behavior' (ibid.).

But as discussed in Chapter 9, this argument that assumptions don't matter is valid only if they are negligibility assumptions (which dismiss features of the real world which are irrelevant or immaterial to the system being modeled) or heuristic assumptions (which are used to simplify argument en route to a more general theory, where the assumptions are dropped).

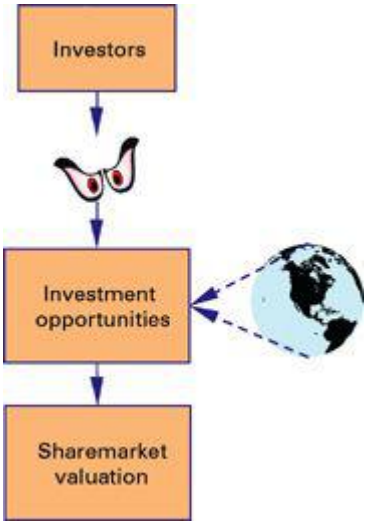
Do Sharpe's assumptions qualify under either of those headings? Clearly not. They are not negligibility assumptions – if they were, then dropping them would not leave the theory 'in a shambles.' They are not heuristic assumptions since, as Sharpe concedes, once they are dropped the theory collapses, and he had no alternative to offer.

Instead, they are domain assumptions (factors that are required to make the theory valid, and in the absence of which the theory is invalid), and therefore the theory is valid only in a world in which those assumptions apply.

That is clearly not our world. The EMH cannot apply in a world in which investors differ in their expectations, in which the future is uncertain, and in which borrowing is rationed. It should have been taken seriously only had Sharpe or its other developers succeeded in using it as a stepping stone to a theory which took account of uncertainty, diverse expectations, and credit rationing. Since they did not do so, the EMH should never have been given any credibility – yet instead it became an article of faith for academics in finance, and a common belief in the commercial world of finance.

Sharpe deserves commendation for honestly discussing the impact on his theory of relaxing his assumptions – unfortunately, the same can't be said for the textbook writers who promulgated his views. However, the problems he saw with his theory are just the tip of the iceberg. There are so many others that it is difficult to think of a theory that could less accurately describe how stock markets behave.

Efficient or prophetic market? Figure 11.6 illustrates the process which the EMH alleges investors use to determine the value of capital assets. Investors objectively consider information about the investment opportunities offered by different companies, and data about world economic prospects. Information that affects the future prospects of investments arrives randomly, generating random movements in the expected future prospects of firms. Investors' rational appraisal of this information leads to an efficient valuation of shares on the basis of expected return and risk, with price variations being caused by the random arrival of new information pertinent to share prices.



11.6 How the EMH imagines that investors behave

This is a one-way process: there is no feedback from share market valuations to investor perceptions, and most importantly, investors are uninterested in what other investors are doing. This, of course, follows naturally from the assumption that all investors agree about the valuations of all companies: why bother checking what your neighbor thinks when you know he thinks exactly what you do (and any difference between his behavior and yours simply reflects his different risk–return preferences)?

To put it mildly, there are serious problems with this theory of stock market behavior. For starters, the EMH makes no distinction between investors' expectations of the future and the future which actually occurs. In essence, the EMH presumes that investors' expectations will be fulfilled: that returns will actually turn out to be what investors expected them to be. In effect, every stock market investor is assumed to be Nostradamus. What economists describe as 'efficient' actually requires that investors be prophetic.

As soon as you allow that investors can disagree, then this economic notion of 'efficient expectations' also collapses. If investors disagree about the future prospects of companies, then inevitably the future is not going to turn out as most – or perhaps even any – investors expect.

This divergence between expectations and outcomes will set up disequilibrium dynamics in the stock market – precisely the sort of behavior that the EMH cannot model, because it is above all a theory of market equilibrium. If investors influence each other's expectations, this is likely to

lead to periods when the market is dominated by pessimistic and optimistic sentiment, and there will be cycles in the market as it shifts from one dominant sentiment to the other.

The efficient markets hypothesis was used to berate market participants for believing that such phenomena as 'bull' and 'bear' markets actually existed: there was always only the efficient market. But even a slight concession to reality indicates that bull and bear phases will be part and parcel of a real-world stock market.

Risks ain't risks Sharpe's measure of risk was standard deviation, a statistical measure of how much the values thrown up by some process vary. If values are fairly evenly distributed around an average, then roughly two-thirds of all outcomes will be one standard deviation either side of the average.

For example, tests of IQ often have an average of 100 and a standard deviation of 16. This means that two-thirds of the population will score between 84 and 116 on an IQ test.

There are at least two problems with applying this concept to investment:

- Is variability really what an investor means by risk?
- To actually work out a standard deviation, you need some process that has thrown up lots of historical data with a pattern which can be expected to continue recurring in the future.

Consider two investments: Steady has an average return of 3 percent, and a standard deviation of 3 percent; Shaky has an average return of 9 percent, and a standard deviation of 6 percent. Which one is 'riskier'?

According to Sharpe's criterion, Shaky is riskier: its standard deviation is twice as big as Steady's. However, according to any sane investor, Steady would be riskier – since there's a much higher chance of getting a negative return from Steady than there is from Shaky. In other words, what an investor really worries about is not so much variability as downside risk. Standard deviation is a very poor proxy for this, even if a standard deviation can be meaningfully calculated in the first place.

This brings us to the second problem. Standard deviation can be used as a measure of variability for things such as the expected outcome of a dice roll, the age at which someone will die, even a golfer's possible scores. However, even here there are differences in how reliable a guide historical averages and standard deviations can be to future outcomes. So long as the dice are well designed, a roll is going to have a one in six chance of turning up a 2 for a considerable time – until, for example, repeated rolls erode its edges. The historical averages for death, however, have changed dramatically in the West even during one lifetime, and major changes (for better or worse, depending on whether genetic engineering or global ecological problems come out on top during the twenty-first century) can be expected in the future. And if an eighteen-year-old golfer had an average of 70 and a standard deviation of 5 now, would you rely on those numbers as a guide to his performance in thirty years' time?

In other words, for measures like standard deviation to be reliable, past outcomes must remain a reliable guide to future outcomes. This is not going to be the case for an investment, because the future performance of a company depends upon future economic circumstances, future inventions, the actions of future competitors, all things to which the past provides no reliable guide beyond a very short time horizon. Investment, and stock market speculation, are, in other words, subject not to risk, but to uncertainty.

We have already discussed the implications of uncertainty for economic analysis. For stock market investors, uncertainty means that the expected yield of an investment over the medium- to long-term future simply can't be known:

Our knowledge of the factors which will govern the yield of an investment some years hence is usually very slight and often negligible. If we speak frankly, we have to admit that our basis of knowledge for estimating the yield ten years hence of a railway, a copper mine, a textile factory, the goodwill of a patent medicine, an Atlantic liner, a building in the City of London, amounts to little and sometimes to nothing; or even five years hence. In fact, those who seriously attempt to make any such estimate are often so much in the minority that their behavior does not govern the market. (Keynes 1936)

Uncertainty, not risk, is the main factor standing between investors and an accurate knowledge of the future prospects of companies. As a result, the expected yield of an investment, the other variable in the EMH model of investor behavior, simply can't be known.

'The dark forces of time and ignorance …' The efficient markets hypothesis argues that investors try to maximize their utility, where the only determinants of that utility are expected returns on the one hand, and risk on the other.

This kind of analysis has been soundly applied to interpret gambling. A gambler playing a game of blackjack faces known payoffs, and the known probabilities of drawing any given card. A good gambler is someone who intelligently applies these well-known regularities to decide how much to bet, when to hold, and when to risk another flip of the card.

This is an invalid concept to apply to an investor's behavior, since the game played in the casino of the stock market is subject to uncertainty, not risk.

Nonetheless, investors still need to form some expectations of the future if they are going to act at all. These will be based partly on factors they currently know – such as prevailing economic conditions – and partly on factors they can't know. In practice, they rely mainly upon the knowable factors simply because they are knowable: investors therefore extrapolate current trends into the indefinite future. As Keynes puts it:

It would be foolish, in forming our expectations, to attach great weight to matters which are very uncertain. It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and scanty. For this reason the facts of the existing situation enter, in a sense disproportionately, into the formation of our long-term expectations; our usual practice being to take the existing situation and to project it into the future, modified only to the extent that we have more or less definite reasons for expecting a change. (Ibid.)

This is clearly an unreliable practice, but in an uncertain world there is simply no other way to act. It is something that we must do in order not to be paralyzed into inaction, but it is something that, at a deep level, we are

aware is untrustworthy. As a result, our forecasts of the future are tempered by an additional factor, the degree of confidence we have that these forecasts will be at least approximately correct. The more significant the degree of change expected, the more fragile that confidence will be.

The share market's valuations therefore reflect both collective forecasts, and the confidence with which these forecasts are made. In tranquil times these valuations will be relatively stable, but

[i]n abnormal times in particular, when the hypothesis of an indefinite continuance of the existing state of affairs is less plausible than usual even though there are no express grounds to anticipate a definite change, the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for a reasonable calculation. (Ibid.)

Therefore, in this uncertain world, the stock market will be ruled not by dispassionate analysis, but by euphoria, fear, uncertainty and doubt. It will be a place, not of analytic rationality, but of emotion.

The madness of the third degree Keynes once described himself as a speculator who had lost two fortunes and made three. His assessment of the behavior of stock market speculators was thus that of a well-informed insider. Keynes described the stock market as a game of 'Musical Chairs […] a pastime in which he is victor […] who secures a chair for himself when the music stops. These games can be played with zest and enjoyment, though all the players know that […] when the music stops some of the players will find themselves unseated' (ibid.).

The essence of this game is not to work out what particular shares are likely to be worth, but to work out what the majority of other players are likely to think the market will think they are worth, since 'it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30, if you also believe that the market will value it at 20 three months hence' (ibid.). In one of the most evocative analogies ever used by an economist, Keynes compared investing in shares to

> those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly

> corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. (Ibid.)

Though this may seem to be a description of the behavior of amateur investors in Internet chat rooms, Keynes insists that it is also the modus operandi of professional stock managers. First, because the future is uncertain, the kind of long-term forecasting which the EMH assumes is the norm is effectively impossible. It is far easier to anticipate 'changes in the conventional basis of valuation a short time ahead of the general public' (ibid.).

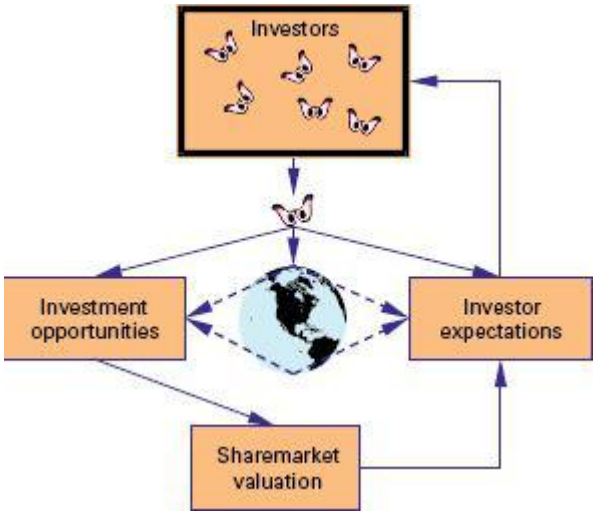
Secondly, the boards that employ such professional stock managers discipline their behavior to make them conform to the norm. Any manager who is truly trying to anticipate future economic trends is bound to make recommendations that are wildly at variance with what is popular in the market, and this behavior will appear eccentric and ill informed in comparison to the current market favorites. Imagine, for example, what would have happened to a funds manager who in mid-2000 advised the fund to sell all its shares in Yahoo, or Amazon, and spend the proceeds buying, for example, bonds.

As Keynes eloquently put it, 'Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally.' Unconventional managers are thus weeded out, leaving behind only those who swim with the crowd.

Thirdly, the long-term investor has to ignore the prospect of quick short-term capital gains, and this runs counter to human nature's desire for quick results.

Finally, a long-term investor can't afford to be highly geared, since the results of being wrong will be expensive and the accumulated financing cost over the long run will be great. Speculators, on the other hand, are attracted

to gearing by the allure of large immediate gains now, at a cost of only minor short-term interest charges (especially when the prevailing aura of confidence during a bull market leads them to discount the possibility of large losses).



11.7 How speculators actually behave

Thus, according to Keynes, rather than looking dispassionately at investment prospects and world economic conditions, the main thing share market investors do is look furtively and emotionally at each other, to attempt to predict how the majority will value particular companies in the immediate future.

This behavior is pictured in Figure 11.7. Though investors do still keep an eye on individual investments and world conditions, and the world does throw in surprising events from time to time, in the main investors analyze the investment community itself.

As a result, there is a feedback from current share valuations to investors' behavior via the impact that present valuations have on investor expectations. A rising market will tend to encourage investors to believe that the market will continue rising; a falling market will maintain the sentiment of the bears. Such a market can find itself a long way from equilibrium as self-reinforcing waves of sentiment sweep through investors. These waves can just as easily break – though long after any rational calculation might suggest that they should – when it becomes clear that the

wave has carried valuations far past a level which is sustainable by corporate earnings.

Addendum: Fama overboard

Eugene Fama and his collaborator Kenneth French played a key role in promoting the efficient markets hypothesis, right from Fama's first major paper while still a PhD student, in which he stated that: 'For the purposes of most investors the efficient markets model seems a good first (and second) approximation to reality. In short, the evidence in support of the efficient markets model is extensive, and (somewhat uniquely in economics) contradictory evidence is sparse' (Fama 1970: 416).

Since then, Fama has become almost synonymous with the efficient markets hypothesis – he, rather than Sharpe, is the author referred to as the originator of the hypothesis in most textbooks on finance. So it's rather significant that, in a major survey article published in 2004, he and French effectively disowned the theory:

> The attraction of the CAPM is that it offers powerful and intuitively pleasing predictions about how to measure risk and the relation between expected return and risk. Unfortunately, the empirical record of the model is poor – poor enough to invalidate the way it is used in applications. The CAPM's empirical problems may reflect theoretical failings, the result of many simplifying assumptions […]
>
> In the end, we argue that whether the model's problems reflect weaknesses in the theory or in its empirical implementation, *the failure of the CAPM in empirical tests implies that most applications of the model are invalid*. (Fama and French 2004: 25; emphasis added)

Their reasons for reaching this conclusion mirror many of the points covered in Chapter 15 on the alternative 'Fractal Markets Hypothesis' and 'Inefficient Markets Hypothesis' (which I wrote in 2000, four years before Fama and French's paper was published): empirical research shows that the

actual behavior of the market strongly contradicts the predictions of the EMH. Specifically:

- share market returns are not at all related to the so-called 'betas';
- much higher returns and lower volatility can be gained by selecting undervalued stocks (ones whose share market value is substantially below their book value); and
- far from there being a trade-off between risk and return, it is possible to select a portfolio that has both high return and low volatility, by avoiding the so-called 'growth stocks' that are popular with market participants.

In considering why the data so strongly contradicted the theory, Fama admitted two points that I labored to make in this chapter: that the theory assumes that all agents have the same expectations about the future *and that those expectations are correct*. Though they put this in a very awkward way, this is unmistakably what they said in this paragraph:

> Sharpe (1964) and Lintner […] add two key assumptions to the Markowitz model to identify a portfolio that must be mean-variance-efficient. The first assumption is *complete agreement*: given market clearing asset prices at *t-1*, investors agree on the joint distribution of asset returns from *t-1* to *t*. *And this distribution is the true one* – that is, it is the distribution from which the returns we use to test the model are drawn. (Ibid.: 26; emphasis added)

A whole generation of economists has thus been taught a theory about finance that assumes that people can predict the future – without that being admitted in the textbook treatments to which they have been exposed, where instead euphemisms such as 'investors make use of all available information' hide the absurd assumptions at the core of the theory.

So wrong it's almost right

The critiques above raise one curious question: how could a theory which was so obviously wrong nonetheless generate predictions about stock

market behavior that, at a superficial level, looked roughly right?

One of the key predictions of the EMH is that 'you can't beat the market': in a perfect capital market, price fluctuations simply reflect the random arrival of new information, and yesterday's price trends are as relevant to tomorrow's as the last roll of the dice is to the next.

On the other hand, if the market is as 'imperfect' as argued above, and trends therefore exist, surely it should be possible for the intelligent investor to profit from these trends? If so, wouldn't this eventually lead to all opportunities for profit being sought out, thus removing the trends and, hey presto, making the market efficient? Not necessarily, for two reasons: a factor discussed briefly in Chapter 8: 'chaos,' and the institutional structure of the market, which Keynes detailed in the *General Theory*.

We'll consider these issues in detail in Chapters 13–14, when I finally leave behind the surreal world of neoclassical economics and consider alternative theories that actually try to be realistic about how a complex monetary economy operates. But first, we have to consider the ultimate denouement of neoclassical economics: its utter failure to anticipate the biggest economic event since the Great Depression.

As this and the previous chapter have pointed out, neoclassical economists of the 1920s also failed to see the Great Depression coming, so their failure to anticipate this crisis was par for the course. Then, their failure led to the temporary overthrow of neoclassical economics by Keynes, but as detailed in Chapter 10, neoclassical economists led a successful counter-revolution that not only eliminated Keynes's ideas from economics, but also set Keynes up to be blamed for this crisis – since the most prominent neoclassical economists of the early twenty-first century called themselves 'New Keynesians.'

In the 1920s, the most prominent neoclassical economist was Irving Fisher, and his failure to see the crisis coming destroyed his public reputation.[15] But though Fisher could be criticized for not foreseeing the Great Depression, he could not be blamed for causing it. He was, after all, merely an observer.

This time round, the most prominent neoclassical was Milton Friedman's acolyte Ben Bernanke. Whereas Fisher had merely been an observer, when the Great Recession hit, Bernanke was chairman of the

organization charged with ensuring that such calamities don't happen: the Federal Reserve. And he had gotten the job because neoclassical economists believed that, out of all of them, he knew best why the Great Depression occurred, and he was therefore the best man to make sure that 'It' could never happen again.

How wrong they were.

12 | MISUNDERSTANDING THE GREAT DEPRESSION AND THE GREAT RECESSION

Bernanke's *Essays on the Great Depression* (Bernanke 2000) is near the top of my stack of books that indicate how poorly neoclassical economists understand capitalism. Most of the others are books of pure theory, such as Debreu's *Theory of Value* (Debreu 1959), or textbooks like Varian's *Microeconomic Analysis* (Varian 1992). Bernanke's distinguished itself by being empirical: he was, he claimed, searching the data to locate the causes of the Great Depression, since:

> To understand the Great Depression is the Holy Grail of macroeconomics. Not only did the Depression give birth to macroeconomics as a distinct field of study, but also – to an extent that is not always fully appreciated – the experience of the 1930s continues to influence macroeconomists' beliefs, policy recommendations, and research agendas. And, practicalities aside, finding an explanation for the worldwide economic collapse of the 1930s remains a fascinating intellectual challenge. (Bernanke 2000: 5)

However, what Bernanke was actually doing was searching for an explanation *that was consistent with neoclassical theory*. Statements to this effect abound throughout the *Essays*, and they highlight the profound difficulty he faced – since according to neoclassical theory, events like the Great Depression should not occur. This disconnection between reality and neoclassical theory had at least the following manifestations that Bernanke admitted to in his *Essays*:

- Monetary variables affect inflation, but are not supposed to affect real variables – money is supposed to be 'neutral':

> Of course, the conclusion that monetary shocks were an important source of the Depression raises a central question in macroeconomics, which is *why nominal shocks should have real effects* (p. 7)
>
> the gold standard theory leaves unsolved the corresponding 'aggregate supply puzzle,' namely, why were the observed worldwide declines in nominal aggregate demand associated with such deep and persistent contractions in real output and employment? Or, in the language of contemporary macroeconomics, *how can we explain what appears to be a massive and very long-lived instance of monetary nonneutrality*? (p. 277)

- A prolonged macro downturn is inconsistent with rational micro behavior:

> my theory […] does have the virtues that, first, it seems capable of explaining the unusual length and depth of the Depression; and, second, it can do this without assuming markedly irrational behavior by private economic agents. *Since the reconciliation of the obvious inefficiency of the Depression with the postulate of rational private behavior remains a leading unsolved puzzle of macroeconomics, these two virtues alone provide motivation for serious consideration of this theory* (p. 42; emphasis added)

- Rational behavior by agents should lead to all prices – including money wages – adjusting rapidly to a monetary shock, so that its impact should be transient:

> slow nominal-wage adjustment (in the face of massive unemployment) is *especially difficult to reconcile with the postulate of economic rationality*. We cannot claim to understand the Depression until we can provide a rationale for this paradoxical behavior of wages (p. 7)

- Rapid adjustment of prices should bring the economy back to equilibrium:

> the failure of nominal wages (and, similarly, prices) to adjust seems *inconsistent with the postulate of economic rationality* (p. 32; emphasis added)

Bernanke began well when he stated that the causes of the Great Depression had to lie in a collapse in aggregate demand – though even here he manifested a neoclassical bias of expecting capitalism to rapidly return to equilibrium after any disturbance:

> Because the Depression was characterized by sharp declines in both output and prices, the premise of this essay is that declines in aggregate demand were the dominant factor in the onset of the Depression.
>
> This starting point leads naturally to two questions: First, what caused the worldwide collapse in aggregate demand in the late 1920s and early 1930s (the 'aggregate demand puzzle')? Second, why did the Depression last so long? In particular, why didn't the 'normal' stabilizing mechanisms of the economy, such as the adjustment of wages and prices to changes in demand, limit the real economic impact of the fall in aggregate demand (the 'aggregate supply puzzle'). (Ibid.: ix)

However, from this point on, his neoclassical priors excluded both salient data and rival intellectual perspectives on the data. His treatment of Hyman Minsky's 'Financial Instability Hypothesis' – which is outlined in

Chapter 13 – is particularly reprehensible. In the entire volume, there is a single, utterly dismissive reference to Minsky:

> Hyman Minsky (1977) and Charles Kindleberger […] have in several places argued for the inherent instability of the financial system but in doing so have had to depart from the assumption of rational economic behavior. [A footnote adds:] I do not deny the possible importance of irrationality in economic life; however it seems that the best research strategy is to push the rationality postulate as far as it will go. (Ibid.: 43)

As we shall see, this is a parody of Minsky's hypothesis. He devoted slightly more space to Irving Fisher and his debt-deflation theory, but what he presented was likewise a parody of Fisher's views, rather than a serious consideration of them:

> The idea of debt-deflation goes back to Irving Fisher (1933). Fisher envisioned a dynamic process in which falling asset and commodity prices created pressure on nominal debtors, forcing them into distress sales of assets, which in turn led to further price declines and financial difficulties. His diagnosis led him to urge President Roosevelt to subordinate exchange-rate considerations to the need for reflation, advice that (ultimately) FDR followed.
>
> Fisher's idea was less influential in academic circles, though, because of the counterargument that debt-deflation represented no more than a redistribution from one group (debtors) to another (creditors). Absent implausibly large differences in marginal spending propensities among the groups, it was suggested, pure redistributions should have no significant macro-economic effects […] (Ibid.: 24)[1]

There are many grounds on which this is a misrepresentation of Fisher,[2] but the key fallacy is the proposition that debt has no macroeconomic effects. From Bernanke's neoclassical perspective, debt merely involves the transfer of spending power from the saver to the borrower, while deflation merely increases the amount transferred, in debt servicing and repayment,

from the borrower back to the saver. Therefore, unless borrowers and savers have very different propensities to consume, this transfer should have no impact on aggregate demand.

The contrast with the theoretical case that Marx, Schumpeter, Keynes and Minsky made about debt and aggregate demand could not be more stark – and in the next chapter I'll make the empirical case that a collapse in debt-financed demand was *the* cause of both the Great Depression and the Great Recession. Bernanke's neoclassical goggles rendered him incapable of comprehending the best explanations of the Great Depression, and led him to ignore the one data set that overwhelmingly explained the fall in aggregate demand and the collapse in employment.

The three reasons he ultimately provided for the Great Depression were (a) that it was caused by the then Federal Reserve's mismanagement of the money supply between 1928 and 1931; (b) that the slow adjustment of money wages to the fall in aggregate demand is what made it last so long; and (c) that the gold standard transmitted the collapse internationally. His conclusion on the first point was emphatic: 'there is now overwhelming evidence that the main factor depressing aggregate demand was a worldwide contraction in world money supplies. This monetary collapse was itself the result of a poorly managed and technically flawed international monetary system (the gold standard, as reconstituted after World War I)' (ibid.: ix).

He was also emphatic about his 'smoking gun': the Great Depression was triggered by the Federal Reserve's reduction of the US base money supply between June 1928 and June 1931:

> The monetary data for the United States are quite remarkable, and tend to underscore the stinging critique of the Fed's policy choices by Friedman and Schwartz […] the United States is the only country in which the discretionary component of policy was arguably significantly destabilizing […] the ratio of monetary base to international reserves […] fell consistently in the United States from […] 1928:II […] through the second quarter of 1931. As a result, U.S. nominal money growth was precisely zero between 1928:IV and 1929:IV, despite both gold inflows and an increase in the money multiplier.

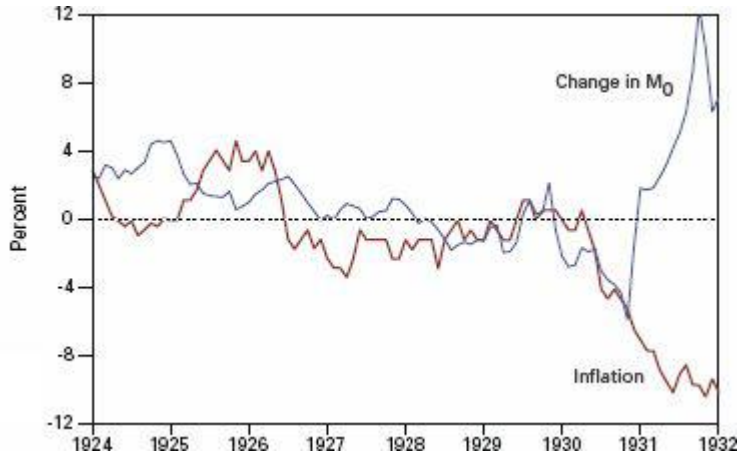
> The year 1930 was even worse in this respect: between 1929:IV and 1930:IV, nominal money in the United States fell by almost 6 [percent], even as the U.S. gold stock increased by 8 [percent] over the same period. The proximate cause of this decline in M_1 was continued contraction in the ratio of base to reserves, which reinforced rather than offset declines in the money multiplier. This tightening seems clearly inconsistent with the gold standard's 'rules of the game,' and locates much of the blame for the early (pre-1931) slowdown in world monetary aggregates with the Federal Reserve. (Ibid.: 153)

There are four problems with Bernanke's argument, in addition to the fundamental one of ignoring the role of debt in macroeconomics. First, as far as smoking guns go, this is a pop-gun, not a Colt .45. Secondly, it has fired at other times since World War II (once in nominal terms, and many times when adjusted for inflation) without causing anything remotely like the Great Depression. Thirdly, a close look at the data shows that the correlations between changes in the rate of growth of the money supply[3] and unemployment conflict with Bernanke's argument that mismanagement of the monetary base was the *causa causans* of the Great Depression. Fourthly, the only other time that it has led to a Great Depression-like event was when Bernanke himself was chairman of the Federal Reserve.

Between March 1928 and May 1929, base money fell at an average rate of just over 1 percent per annum in nominal terms, and a maximum rate of minus 1.8 percent.[4] It fell at the same rate between 1948 and 1950, and coincided with a garden-variety recession, rather than a prolonged slump: unemployment peaked at 7.9 percent and rapidly returned to boom levels of under 3 percent. So the pop-gun has fired twice in nominal terms, and only once did it 'cause' a Great Depression.

It could also be argued, from a neoclassical perspective, that the Fed's reduction in base money in the lead-up to the Great Depression was merely a response to the rate of inflation, which had turned negative in mid-1924. Neoclassical theory emphasizes money's role as a means to facilitate transactions, and a falling price level implies a need for less money. On this point Milton Friedman, whom Bernanke cited as a critic of the Federal

Reserve for letting base money fall by 1 percent per annum, argued elsewhere that social welfare would be maximized if the money supply actually fell by 10 percent per year.[5]

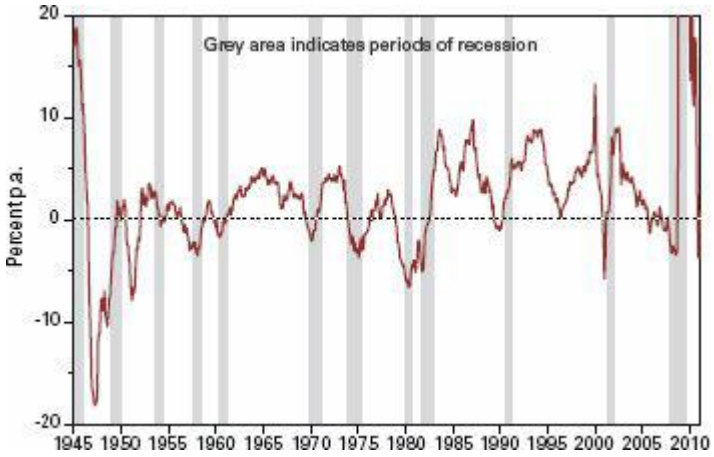


12.1 Inflation and base money in the 1920s

When the inflation-adjusted rate of change of base money is considered, there were numerous other periods when base money fell as fast as in 1928/29, without leading to a depression-scale event. The average inflation-adjusted rate of growth of M_0 in mid-1928 to mid-1929 was minus 0.5 percent, and even in 1930 M_0 fell by a maximum of 2.2 percent per annum in real terms. There were six occasions in the post-World War II period when the real rate of decline of M_0 was greater than this without causing a depression-like event[6] (though there were recessions on all but one occasion). Why did the pop-gun fire then, but emit no smoke?

The reason is, of course, that the pop-gun wasn't really the guilty culprit in the crime of the Great Depression, and Friedman and Bernanke's focus upon it merely diverted attention from the real culprit in this investigation: the economy itself. Capitalism was on trial because of the Great Depression, and the verdict could well have been attempted suicide – which is the last verdict that neoclassical economists could stomach, because they are wedded to the belief that capitalism is inherently stable. They cannot bring themselves to consider the alternative perspective that capitalism is inherently unstable, and that the financial sector causes its most severe breakdowns.

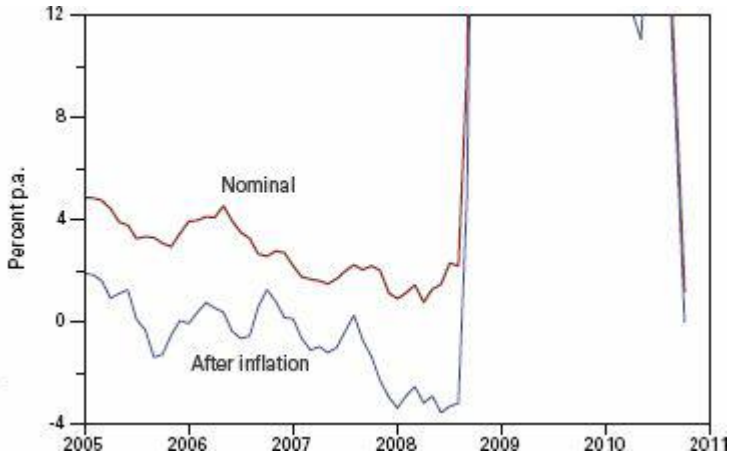
To neoclassicals like Friedman and Bernanke, it was better to blame one of the nurses for incompetence, than to admit that capitalism is a manic-depressive social system that periodically attempts to take its own life. It was better to blame the Fed for not administering its M_0 medicine properly, than to admit that the financial system's proclivity to create too much debt causes capitalism's periodic breakdowns.



12.2 Inflation and base money in the post-war period

It is therefore a delicious if socially painful irony that the only other time that the pop-gun fired and a depression-like event *did* follow was when the chairman of the Federal Reserve was one Ben S. Bernanke.

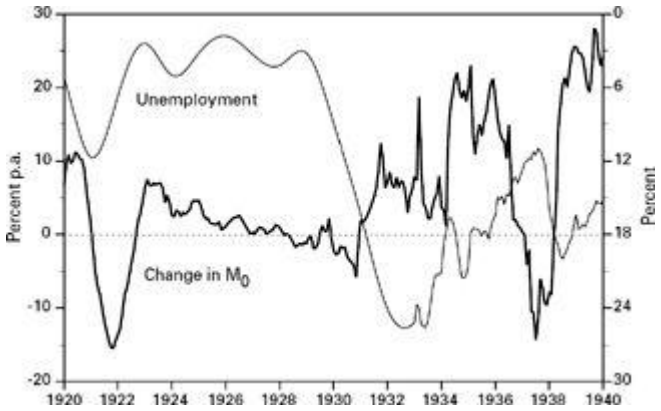
Bernanke began as chairman on 1 February 2006, and between October 2007 and July 2008, the change in M_0 was an inflation-adjusted *minus* 3 percent – one percent lower than its steepest rate of decline in 1930–33. The rate of change of M_0 had trended down in nominal terms ever since 2002, when the Greenspan Fed had embarked on some quantitative easing to stimulate the economy during the recession of 2001. Then, M_0 growth had turned from minus 2 percent nominal (and minus 6 percent real) at the end of 2000 to plus 11 percent nominal (and 8 percent real) by July 2001. From there it fell steadily to 1 percent nominal – and minus 3 percent real – by the start of 2008.



12.3 Bernanke's massive injection of base money in QE1

Whatever way you look at it, this makes a mockery of the conclusion to Bernanke's fawning speech at Milton Friedman's ninetieth birthday party in November 2002: 'Let me end my talk by abusing slightly my status as an official representative of the Federal Reserve. I would like to say to Milton and Anna: Regarding the Great Depression. You're right, we did it. We're very sorry. But thanks to you, we won't do it again' (Bernanke 2002b).

Either Bernanke forgot what he learnt from Friedman and his own research once in office – since Friedman *and Bernanke's* criticism of the 1920s Fed was that it let the growth rate of M_0 drop too low *before* the crisis – or the advice itself was irrelevant. The latter is of course the case. As I argue in the next chapter, the key to preventing depressions is to prevent an explosion in the ratio of private debt to GDP, so that debt-financed demand cannot reach a level from which its collapse will trigger a depression. Far from explaining what caused the Great Depression, Friedman and Bernanke's simplistic perspective diverted attention from the real culprit – the expansion of private debt by the banking sector – and ignored the enormous growth of debt that occurred while the central bank was under the thrall of neoclassical economics.



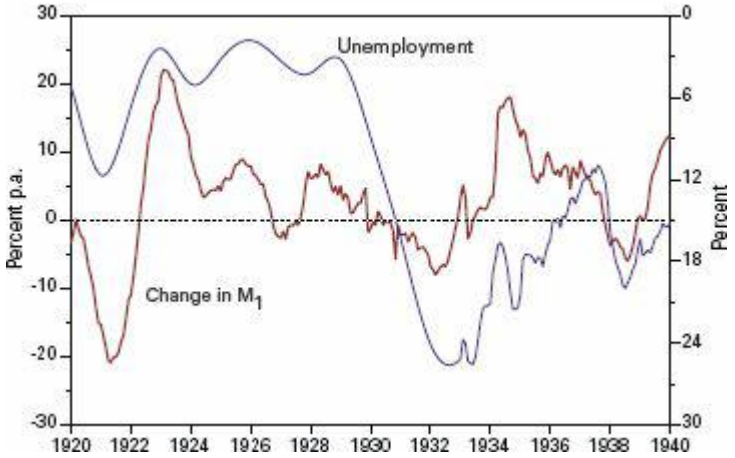
12.4 Change in M_0 and unemployment, 1920–40

The relative irrelevance of changes in base money as a cause of changes in unemployment, let alone a cause of serious economic breakdown, can be gauged by looking at the correlation between the growth of M_0 and the rate of unemployment over the period from 1920 till 1940 – across both the boom of the Roaring Twenties and the collapse of the Great Depression (see Figure 12.4). If too slow a rate of growth of M_0 can trigger a depression, as Bernanke asserts, then surely there should be a *negative* correlation between the change in M_0 and the rate of unemployment: unemployment should fall when the rate of change of M_0 is high, and rise when it is low.

The correlation has the right sign for the period from 1920 till 1930 (minus 0.22 for changes in nominal M_0 and minus 0.19 after inflation) but the wrong one for the period from 1930 till 1940 (plus 0.28 for nominal M_0 and 0.54 after inflation), and it is positive for the entire period 1920–40 (plus 0.44 for nominal change to M_0, and 0.61 for the inflation-adjusted rate of change). Therefore unemployment *increased* when the rate of growth of M_0 increased, and fell when it fell. Lagging the data on the basis that changes in M_0 should precede changes in unemployment doesn't help either – the correlation remains positive.

On the other hand, the correlation of changes in M_1 to unemployment is negative as expected over both the whole period (minus 0.47 for nominal change and minus 0.21 for inflation-adjusted change) and the sub-periods of the Roaring Twenties (minus .31 for nominal M_1 and 0.79 for inflation-adjusted) and the Great Depression (minus 0.62 for nominal and 0.31 for

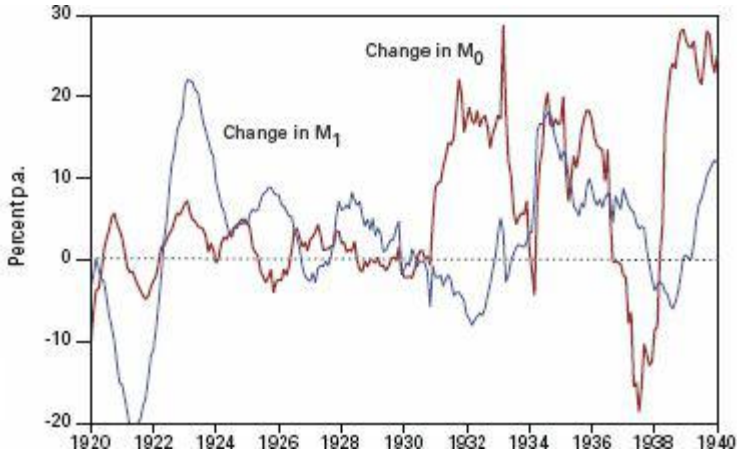
real). So any causal link relates more to private-bank-driven changes in M_1 than to central-bank-driven changes in M_0.



12.5 Change in M1 and unemployment, 1920–40

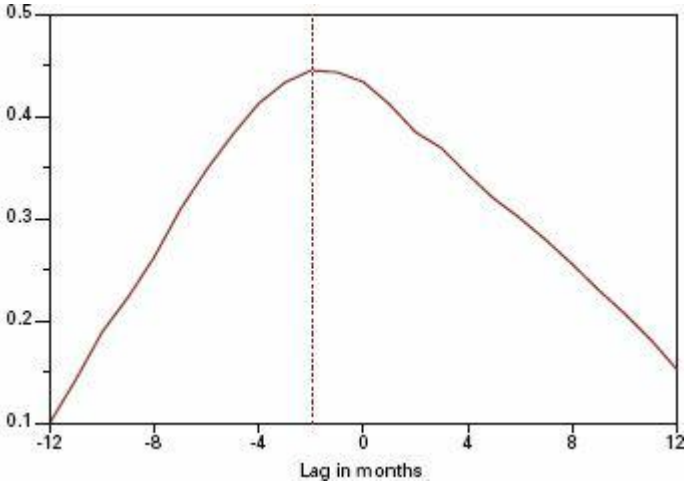
There are only two interpretations of this, neither of which supports the case that Bernanke made against the 1920s Fed.

The first is that, far from changes in M_0 driving unemployment, *the unemployment rate drives changes in M0*. The Fed largely ignored the level of unemployment when it was low (during the 1920s), but went into panic policy mode when it exploded during the Great Depression. It therefore increased the level of M_0 when unemployment rose, and decreased it when unemployment seemed to be falling. The causation between changes in M_0 and unemployment is therefore the reverse of the one Bernanke sought to prove.



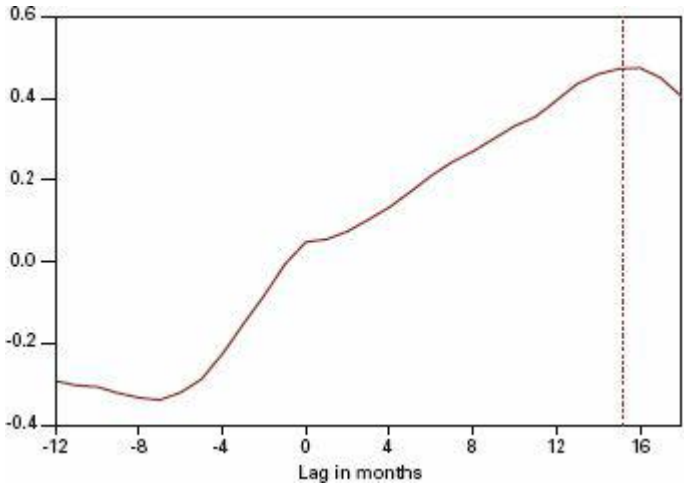
12.6 Change in M0 and M1, 1920–40

The second is that other factors are far more important in determining the rate of unemployment – and by extension, causing Great Depressions as well – than the Fed's quantitative monetary policy. Two hints that the private financial system was the culprit are given by the negative relationship between changes in M_1 and unemployment, and by the fact that the relationship of M_0 to M_1 shifted dramatically when the Great Depression hit.



12.7 M0–M1 correlation during the Roaring Twenties

Before the Great Depression, there was a positive relationship between changes in M_0 and changes in M_1, and changes in M_0 appeared to lead changes in M_1 by about one to two months. This is the direction of causation expected by the conventional model of money creation – the 'Money Multiplier' – which argues that commercial banks need reserves in order to be able to lend (though the magnitude is lower than might be expected).



12.8 M0–M1 correlation during the Great Depression

After the Great Depression, this relationship broke down completely, and changes in M_1 appeared to lead changes in M_0 by up to fifteen months. This contradicts the conventional theory – a point I elaborate upon shortly.

So Bernanke's analysis of what caused the Great Depression is erroneous, and to make matters worse, he didn't even follow his own advice prior to the Great Recession when chairman of the Federal Reserve. But he certainly took his own analysis seriously *after* the Great Recession began – increasing M_0 as never before in an attempt to turn deflation into inflation.

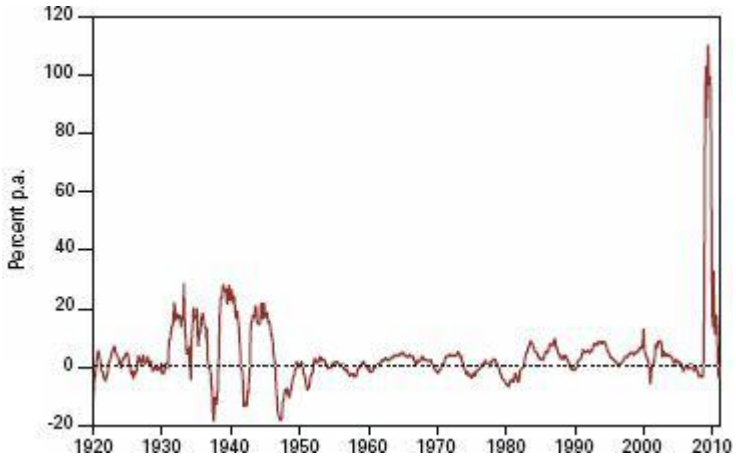
After the Great Recession: Bernanke to the rescue?

Bernanke foreshadowed that he might do this in a speech for which he gained the nickname 'Helicopter Ben' in 2002. With the unfortunate title of 'Deflation: making sure "It" doesn't happen here,' it proved to be remarkably unprescient in terms of the economic future, since the US did slip into deflation. But the speech accurately signaled what he did do, once what he had hoped to avoid actually occurred:

> Like gold, U.S. dollars have value only to the extent that they are strictly limited in supply. But the U.S. government has a technology, called a printing press (or, today, its electronic equivalent), that allows it to produce as many U.S. dollars as it wishes at essentially no cost. By increasing the number of U.S. dollars in circulation […] the U.S. government can also reduce

> the value of a dollar in terms of goods and services, which is equivalent to raising the prices in dollars of those goods and services. We conclude that, under a paper-money system, a determined government can always generate higher spending and hence positive inflation […]
>
> Normally, money is injected into the economy through asset purchases by the Federal Reserve. To stimulate aggregate spending when short-term interest rates have reached zero, the Fed must expand the scale of its asset purchases or, possibly, expand the menu of assets that it buys. Alternatively, the Fed could find other ways of injecting money into the system – for example, by making low-interest-rate loans to banks or cooperating with the fiscal authorities. Each method of adding money to the economy has advantages and drawbacks, both technical and economic. One important concern in practice is that calibrating the economic effects of nonstandard means of injecting money may be difficult, given our relative lack of experience with such policies. Thus, as I have stressed already, prevention of deflation remains preferable to having to cure it. If we do fall into deflation, however, we can take comfort that the logic of the printing press example must assert itself, and sufficient injections of money will ultimately always reverse a deflation. (Bernanke 2002a)

In late 2008, Bernanke turned on the printing presses as never before, doubling base money in a mere five months, when the previous doubling had taken thirteen years.

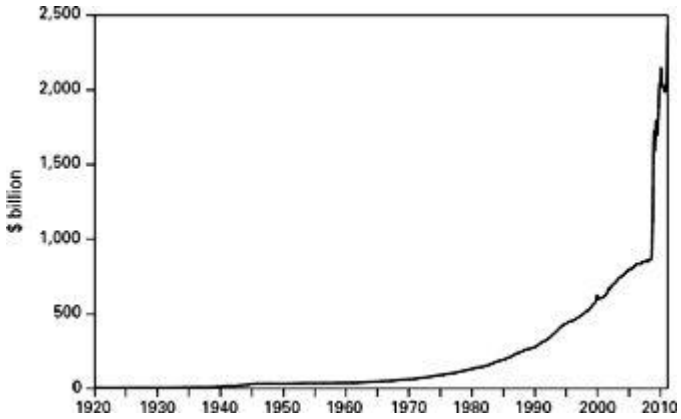


12.9 Bernanke's 'quantitative easing' in historical perspective

In inflation-adjusted terms, he expanded M_0 at a rate of over 100 percent a year, when its average annual rate of growth for the preceding five decades was 2.3 percent. By the time Bernanke finally took his foot off the M_0 accelerator one and a half years later, base money had jumped from $850 billion to $2.15 trillion (see Figure 12.10).

There is little doubt that this massive, unprecedented injection of base money did help reverse the deflation that commenced very suddenly in 2008, when inflation fell from plus 5.6 percent in mid-2008 to minus 2.1 percent a year later – the sharpest fall in inflation in post-World War II history. But I expect Bernanke was underwhelmed by the magnitude of the change: inflation rose from minus 2.1 percent to a peak of 2.7 percent, and it rapidly fell back to a rate of just 1 percent. That is very little inflationary bang for a large amount of bucks.

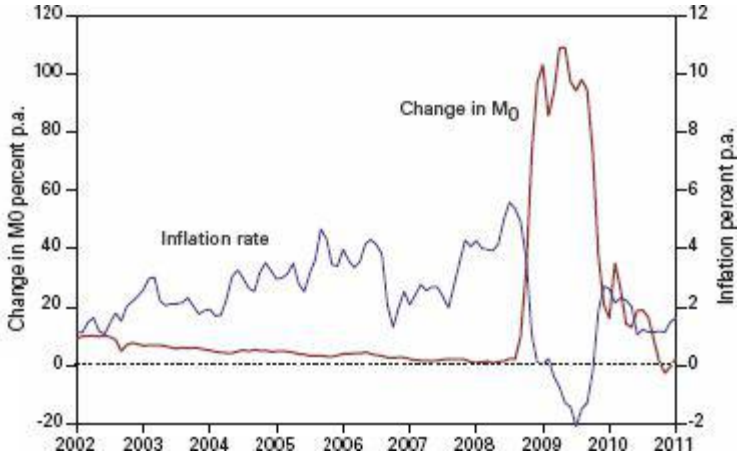
According to the conventional model of money creation – known as the 'Money Multiplier' – this large an injection of government money into the reserve accounts of private banks should have resulted in a far larger sum of bank-created money being added to the economy – as much as $10 trillion. This amplification of Bernanke's $1.3 trillion injection should have rapidly revived the economy – according to neoclassical theory. This is precisely what President Obama, speaking no doubt on the advice of his economists, predicted when he explained the strategy they had advised him to follow, twelve weeks after he took office:



12.10 The volume of base money in Bernanke's 'quantitative easing' in historical perspective

> And although there are a lot of Americans who understandably think that government money would be better spent going directly to families and businesses instead of banks – 'where's our bailout?' they ask – the truth is that a dollar of capital in a bank can actually result in eight or ten dollars of loans to families and businesses, *a multiplier effect* that can ultimately lead to a faster pace of economic growth. (Obama 2009: 3; emphasis added)

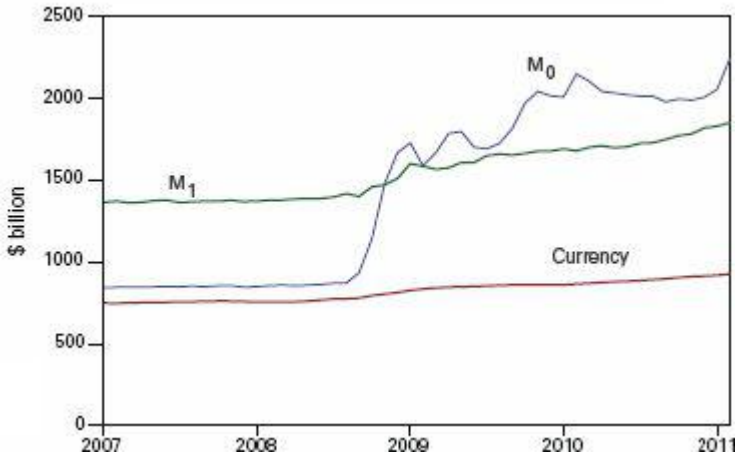
Only that isn't what happened. The dramatic increase in bank reserves spurred only a tiny increase in money in circulation: the 110 percent growth rate of M_0 resulted in only a 20 percent rate of growth of M_1.



12.11 Change in M1 and inflation before and during the Great Recession

The difference in growth rates was so great that there is now *less* money in check accounts and currency in circulation than there is money in the

reserve accounts of the commercial banks.



12.12 The money supply goes haywire

The 'eight or ten dollars of loans to families and businesses' from each extra 'dollar of capital in a bank' simply didn't happen. What went wrong?

The mythical Money Multiplier

Few concepts are more deserving than the 'Money Multiplier' of Henry Mencken's aphorism that 'Explanations exist; they have existed for all time; there is always a well-known solution to every human problem – neat, plausible, and wrong.'[7]

In this model, money is created in a two-stage process. First, the government creates 'fiat' money, say by printing dollar bills and giving them to an individual. The individual then deposits the dollar bills in his bank account. Secondly, the bank keeps a fraction of the deposit as a reserve, and lends out the rest to a borrower. That borrower then deposits this loaned money in another bank account, and the process repeats.

Let's say that the amount created by the government is $100, the fraction the banks keep as a reserve (known as the 'Reserve Requirement' and set by the government or central bank) is 10 percent, and it takes banks a week to go from getting a new deposit to making a loan. The process starts with the $100 created by the government. One week later, the first bank has created another $90 by lending 90 percent of that money to a borrower. A week later, a second bank creates another $81 – by keeping $9 of the new deposit in reserve and lending out the other $81. The process

keeps on going so that, after many weeks, there will be $1,000 created, consisting of the initial printing of $100 by the government, and $900 in credit money created by the banking system – which is matched by $900 in additional debt. There will be $900 of credit money in circulation, facilitating trade, while another $100 of cash will be held by the banks in reserve (see Table 12.1).

TABLE 12.1 The alleged Money Multiplier process ($)

Week	Loans	Deposits	Cash kept by bank	Sum of loans	Sum of cash
0	0	100	10	0	10
1	90	90	9	90	19
2	81	81	8	171	27
3	73	73	7	244	34
4	66	66	7	310	41
5	59	59	6	369	47
6	53	53	5	422	52
7	48	48	5	470	57
8	43	43	4	513	61
9	39	39	4	551	65
10	35	35	3	586	69
Total after 10 weeks	686.19	586.19	68.62	586.19	68.62
Final totals	1,000	900	100	900	100

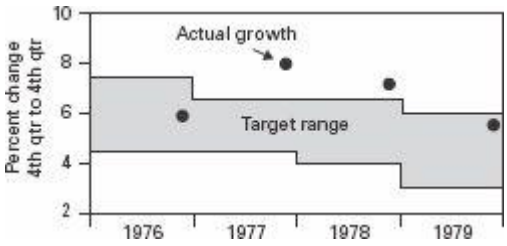
In this simple illustration, all the notes remain in the banks' vaults, while all commerce is undertaken by people electronically transferring the sums in their deposit accounts. Of course, we all keep some notes in our pockets as well for small transactions, so there's less credit created than the example implies, but the model can be modified to take account of this.

This process is also known as 'Fractional Reserve Banking,' and it's the process that Obama, on the advice of his economists, relied upon to rapidly bring the Great Recession to an end. Its failure to work was superficially due to some issues that Bernanke was well aware of,[8] but the fundamental reason why it failed is that, as a model of how money is actually created, it is 'neat, plausible, and wrong.'

The fallacies in the model were first identified by practical experience, and then empirical research.

In the late 1970s, when Friedman's monetarism dominated economic debate and the Federal Reserve Board under Volcker attempted to control

inflation by controlling the rate of growth of the money supply, the actual rate normally exceeded the maximum target that the Board set (Lindsey, Orphanides et al. 2005: 213). Falling below the target range could be explained by the model, but consistently exceeding it was hard to reconcile with the model itself.



12.13 Lindsey, Orphanides, Rasche 2005, p. 213

Empirical research initiated by Basil Moore (Moore 1979, 1983, 1988a, 1997, 2001) and later independently corroborated by numerous researchers, including Kydland and Prescott (1990), confirmed a simple operational observation about how banks actually operate made in the very early days of the monetarist controversy, by the then senior vice-president of the New York Federal Reserve, Alan Holmes.

The 'Money Multiplier' model assumes that banks need excess reserves before they can make loans. The model process is that first deposits are made, creating excess reserves, and then these excess reserves allow loans to be made, which create more deposits. Each new loan reduces the level of excess reserves, and the process stops when this excess has fallen to zero.

But in reality, Holmes pointed out, banks create loans first, which simultaneously creates deposits. If the level of loans and deposits then means that banks have insufficient reserves, then they get them afterwards – and they have a two-week period in which to do so.[9] In contrast to the Money Multiplier fantasy of bank managers who are unable to lend until they receive more deposits, the real-world practicality of banking was that the time delay between deposits and reserves meant that the direction of causation flowed, not from reserves to loans, but from loans to reserves.

Banks, which have the reserves needed to back the loans they have previously made, extend new loans, *which create new deposits simultaneously*. If this then generates a need for new reserves, and the

Federal Reserve refuses to supply them, then it would force banks to recall old or newly issued loans, and cause a 'credit crunch.'

The Federal Reserve is therefore under great pressure to provide those reserves. It has some discretion about how to provide them, but unless it is willing to cause serious financial ructions to commerce on an almost weekly basis, it has no discretion about whether those reserves should be provided.

Holmes summed up the monetarist objective of controlling inflation by controlling the growth of base money – and by inference the Money Multiplier model itself – as suffering from 'a naive assumption':

> that the banking system only expands loans after the [Federal Reserve] System (or market factors) have put reserves in the banking system. *In the real world, banks extend credit, creating deposits in the process, and look for the reserves later.* The question then becomes one of whether and how the Federal Reserve will accommodate the demand for reserves. In the very short run, the Federal Reserve has little or no choice about accommodating that demand; over time, its influence can obviously be felt. (Holmes 1969: 73; emphasis added)

With causation actually running from bank lending and the deposits it creates to reserve creation, the changes in credit money should therefore precede changes in fiat money. This is the opposite of what is implied by the 'Money Multiplier' model (since in it government money – base money or M_0 – has to be created before credit money – M_1, M_2 and M_3 – can be created), and it is precisely what Kydland and Prescott found in their empirical analysis of the timing of economic variables:

> There is no evidence that either the monetary base or M_1 leads the cycle, although some economists still believe this monetary myth. Both the monetary base and M_1 series are generally procyclical and, if anything, the monetary base lags the cycle slightly […] The difference in the behavior of M_1 and M_2 suggests that the difference of these aggregates (M_2 minus M_1) should be considered […] The difference of M_2–M_1 leads the cycle by even more than M_2, with the lead being about three quarters […] (Kydland and Prescott 1990: 4)

Well before Kydland and Prescott reached this statistical conclusion, the post-Keynesian economist Basil Moore pointed out the implication of the actual money creation process for macroeconomic theory. When macroeconomic models actually considered the role of money, they treated the money supply as an exogenous variable under the direct control of the government – this is an essential feature of Hicks's IS-LM model, for instance. But since credit money is created before and causes changes in government money, the money supply must instead be endogenous. The 'Money Multiplier' model of money creation was therefore a fallacy:

> This traditional view of the bank money creation process relies on the bank reserves–multiplier relation. The Fed is posited to be able to affect the quantity of bank deposits, and thereby the money stock, by determining the nominal amount of the reserve base or by changing the reserve multiplier […]
>
> There is now mounting evidence that the traditional characterization of the money supply process, which views changes in an exogenously controlled reserve aggregate as 'causing' changes in some money stock aggregate, is fundamentally mistaken. Although there is a reasonably stable relationship between the high-powered base and the money stock, and between the money stock and aggregate money income, the causal relationship implied is exactly the reverse of the traditional view. (Moore 1983: 538)

It is possible to interpret this reverse causation as representing 'a lack of moral fiber' by central bankers – accommodating banks' loan-creation rather than regulating it in the interests of the economy – but Moore pointed out that the provision of reserves by central banks to match loan-creation by banks merely mirrored the standard behavior of banks with respect to their business clients. Businesses need credit in order to be able to meet their costs of production prior to receiving sales receipts, and this is the fundamental beneficial role of banks in a capitalist economy:

> In modern economies production costs are normally incurred and paid prior to the receipt of sales proceeds. Such costs represent a

> working capital investment by the firm, for which it must necessarily obtain finance. Whenever wage or raw materials price increases raise current production costs, unchanged production flows will require additional working capital finance. In the absence of instantaneous replacement cost pricing, firms must finance their increased working capital needs by increasing their borrowings from their banks or by running down their liquid assets. (Ibid.: 545)

Banks therefore accommodate the need that businesses have for credit via additional lending – and if they did not, ordinary commerce would be subject to Lehman Brothers-style credit crunches on a daily basis. The Federal Reserve then accommodates the need for reserves that the additional lending implies – otherwise the Fed would cause a credit crunch: 'Once deposits have been created by an act of lending, the central bank must somehow ensure that the required reserves are available at the settlement date. Otherwise the banks, no matter how hard they scramble for funds, could not in the aggregate meet their reserve requirements' (ibid.: 544).

Consequently, attempts to use the 'Money Multiplier' as a control mechanism – either to restrict credit growth as during the monetarist period of the late 1970s, or to cause a boom in lending during the Great Recession – are bound to fail. It is not a control mechanism at all, but a simple measure of the ratio between the private banking system's creation of credit money and the government's creation of fiat money. This can vary dramatically over time: growing when the private banks are expanding credit rapidly and the government tries – largely vainly – to restrain the growth in money; collapsing when private banks and borrowers retreat from debt in a financial crisis, and the government tries – again, largely vainly – to drive the rate of growth of money up.

12.14 The empirical 'Money Multiplier', 1920–40

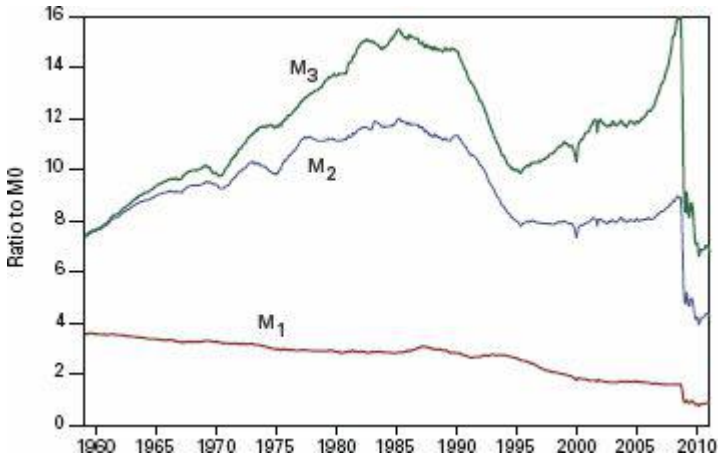
This is something that Bernanke should have known from his own research on the Great Depression. Then, the 'Money Multiplier' rose from under 6 in the early 1920s to over 9 in 1930, only to plunge to below 4.5 by 1940 (see Figure 12.14).

Perhaps he did remember this lesson of history, since his increase in base money was far greater than that of his predecessors. He may well have put such a massive influx of money into the system simply because he feared that little or no additional credit money would be forthcoming as a result. Better then to flood the economy with fiat money and hope that that alone would cause the desired boost to aggregate demand.

We will have to await his memoirs to know, but even if so, he (and Obama's other neoclassical economic advisors) made the wrong choice by putting this injection of fiat money into the reserve accounts of the banks, rather than giving it to the public – as Obama considered in his 'where's our bailout?' counterpoint in his April 2009 speech.

The money drove up the unused reserves of the banking sector as never before (from $20 billion before the crisis to over $1 trillion after it) and the 'Money Multipliers' – which in reality are no more than the ratios of the three measures of the broad money supply, M_3, M_2 and M_1, to base money – collapsed as never before. The M_3 ratio fell from over 16 to under 8, and has continued to fall to below 7 since then; the M_2 ratio – the one most comparable to the M_1 ratio back in the 1920s–1940s – fell from 9 to below 4, while most embarrassingly of all, the M_1 ratio fell below 1, hit as low as 0.78, and is still below 0.9 two years after Bernanke's fiat money injection.

Some ‘multiplier effect.’ Obama was sold a pup by his neoclassical advisors. The huge injection of fiat money would have been far more effective had it been given to the public, who at least would have spent it into circulation.



12.15 The empirical ‘Money Multiplier’, 1960–2012

Don’t mention the data

As this book details, neoclassical economics is awash with examples of its internal contradictions being ignored by its believers, so in one sense their practice of pretending that the Money Multiplier determines the amount of money in the economy is just another example of neoclassical economists believing in something that doesn’t exist. However, the Money Multiplier is different in at least two ways. First, many neoclassical economists *know* that it doesn’t exist, and secondly, its non-existence is empirically obvious. So rather than ignoring the problem because they are unaware of it, or of its ramifications – as with the Sonnenschein-Mantel-Debreu conditions – they ignore it simply because it is inconvenient to acknowledge it.

Admitting that the Money Multiplier doesn’t exist is inconvenient because, if s0, then the supply of money is not exogenous – set by the government – but endogenous – determined by the workings of a market economy. This in turn means that this endogenous process affects real economic variables such as the level of investment, the level of employment and the level of output, when it has always been a tenet of

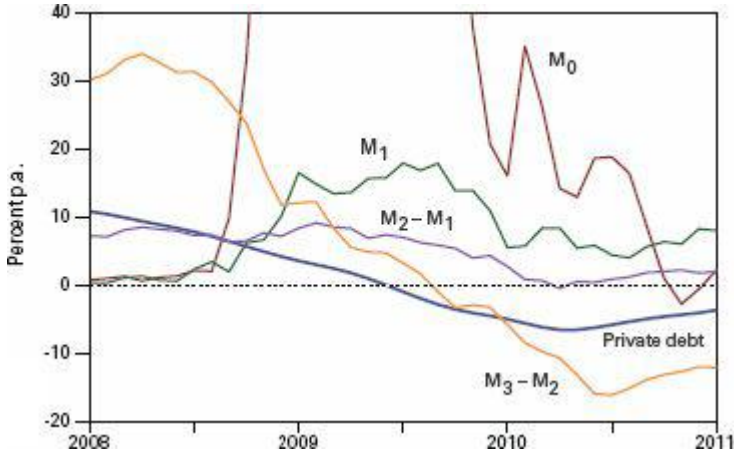
neoclassical theory that 'money doesn't matter.' So acknowledging the empirically bleedingly obvious fact that the Money Multiplier is a myth also means letting go of another favorite neoclassical myth, that the dynamics of money can safely be ignored in economic analysis. Consequently, clear evidence that the Money Multiplier is a myth has been ignored even by the neoclassical economists who know otherwise.

One of the clearest instances of this is the difference between the very emphatic conclusion that Kydland and Prescott reached about the importance of credit, and their subsequent theoretical work. In their conclusion to their empirical paper, they made a clear case for the need to develop a theory of endogenous credit:

> The fact that the transaction component of real cash balances (M_1) moves contemporaneously with the cycle while the much larger nontransaction component (M_2) leads the cycle suggests that credit arrangements could play a significant role in future business cycle theory. *Introducing money and credit into growth theory in a way that accounts for the cyclical behavior of monetary as well as real aggregates is an important open problem in economics*. (Kydland and Prescott 1990: 15; emphasis added)

However, they have done nothing since to develop such a theory. Instead, they have continued to champion the 'Real Business Cycle Theory' that they developed prior to this empirical research, and Carpenter and Demiralp note that Kydland continues 'to refer to the very narrow money multiplier and accord it a principal role in the transmission of monetary policy' (Carpenter and Demiralp 2010: 2, commenting on Freeman and Kydland 2000).

This charade of continuing to believe in a concept whose non-existence was an empirical fact could be maintained for as long as the Money Multiplier didn't have any real-world significance. Unfortunately, the 'bailout the banks' strategy that Obama was advised to follow by Bernanke depended crucially on the Money Multiplier working to turn the huge increase in reserves into an even larger increase in private sector lending. It was an abject failure: excess reserves increased by a factor of 50, but private sector lending fell, as did credit money.



12.16 The disconnect between private and fiat money during the Great Recession

A recent paper by Federal Reserve associate director Seth Carpenter entitled 'Money, reserves, and the transmission of monetary policy: does the Money Multiplier exist?' (Carpenter and Demiralp 2010) finally acknowledges this:

> Since 2008, the Federal Reserve has supplied an enormous quantity of reserve balances relative to historical levels as a result of a set of nontraditional policy actions. These actions were taken to stabilize short-term funding markets and to provide additional monetary policy stimulus at a time when the federal funds rate was at its effective lower bound.
>
> The question arises whether or not this unprecedented rise in reserve balances ought to lead to a sharp rise in money and lending. *The results in this paper suggest that the quantity of reserve balances itself is not likely to trigger a rapid increase in lending […] the narrow, textbook money multiplier does not appear to be a useful means of assessing the implications of monetary policy for future money growth or bank lending*. (Ibid.: 29; emphasis added)

This acknowledgment of reality is good to see, but – compared both to the data and the empirically oriented work of the rival 'post-Keynesian' school of thought – it is thirty years and one economic crisis too late. It also

post-dates the effective abolition of the Reserve Requirement – an essential component of the 'Money Multiplier' model – by about two decades.

Since 1991, the publicly reported Reserve Requirement has been effectively applicable only to household bank accounts, which are a tiny fraction of the aggregate deposits of the banking system (see Table 12 in O'Brien 2007: 52). As Carpenter and Demiralp note, today reserve requirements 'are assessed on only about one-tenth of M_2':

> Casual empirical evidence points away from a standard money multiplier and away from a story in which monetary policy has a direct effect on broader monetary aggregates. The explanation lies in the institutional structure in the United States, especially after 1990.
>
> First, there is no direct link between reserves and money – as defined as M_2. Following a change in required reserves ratios in the early 1990s, reserve requirements are assessed on only about one-tenth of M_2.
>
> Second, there is no direct link between money – defined as M_2 – and bank lending. Banks have access to non-deposit funding (and such liabilities would also not be reservable), so the narrow bank lending channel breaks down in theory. Notably, large time deposits, a liability that banks are able to manage more directly to fund loans, are not reservable and not included in M_2. Banks' ability to issue managed liabilities increased substantially in the period after 1990, following the developments and increased liquidity in the markets for bank liabilities.
>
> Furthermore, the removal of interest rate ceilings through Regulation Q significantly improved the ability of banks to generate non-reservable liabilities by offering competitive rates on large time deposits. Additionally, money market mutual funds account for about one-fifth of M_2, but are not on bank balance sheets, and thus they cannot be used to fund lending. These facts imply that the tight link suggested by the multiplier between

> reserves and money and bank lending does not exist. (Carpenter and Demiralp 2010: 4–5)

The effective freedom of banks to decide how much money they will keep in reserve – and thus not use as a source of income – versus the amount they will lend, effectively leaves the private banks free to create as much credit as they wish. This is a freedom they have exploited with gusto, as I detail in the next chapter.

After the Great Recession II: neoclassical responses

One would hope that the complete failure of neoclassical models to anticipate the Great Recession might lead to some soul-searching by neoclassical economists: was there not something fundamentally wrong in their modeling that they could be blindsided by such a huge event?

Unfortunately, they are so wedded to their vision of the economy that even an event like the Great Recession can't shake them. Their near-universal reaction has been that it was simply an extreme event – like a sequence of a dozen coin-tosses that all resulted in 'heads,' which is a feasible though very rare outcome.[10] Though such a thing is possible, when it will happen can't be predicted.

In saying this, they of course ignored the public warnings from myself and others, as documented by Bezemer (Bezemer 2009, 2010, 2011), despite the fact that those warnings were made, not merely in non-mainstream academic publications, but in the media as well. Here I can't resist quoting the governor of my own country's central bank, Glenn Stevens: 'I do not know anyone who predicted this course of events. This should give us cause to reflect on how hard a job it is to make genuinely useful forecasts. What we have seen is truly a "tail" outcome – the kind of outcome that the routine forecasting process never predicts. But it has occurred, it has implications, and so we must act on it' (Stevens 2008).

That speech, made in Sydney in December 2008, ignored not only the well-known warnings in the USA by Peter Schiff and Nouriel Roubini, but my own in Australia since December 2005. These had included appearances on the leading current affairs programs *60 Minutes* (60 Minutes 2008) and *The 7.30 Report* (7.30 Report 2007).

Central bankers like Stevens and Bernanke had to live in a cocoon not to know of such warnings, and neoclassical economics provides the silk of this cocoon, because it refuses to consider any analysis of economics that does not make neoclassical assumptions. Since those who predicted the crisis did so – as they had to – using non-neoclassical tools, to Bernanke and his brethren around the world, those warnings did not exist.

Unfortunately, the Great Recession does exist, and neoclassical economists have been forced to consider it. Their responses have taken two forms: tweaking the 'exogenous shocks' to their models until the models generate results that look like the Great Recession; and adding additional tweaks to the core neoclassical model that at least to some degree incorporate the effects of debt. Both approaches completely miss the real causes of this crisis.

It's just a jolt to the left …

As of February 2011, there were two neoclassical papers that attempted to comprehend the Great Recession using New Keynesian models which, of course, had completely failed to anticipate it (McKibbin and Stoeckel 2009; Ireland 2011). Since the underlying theory generates tranquil equilibrium growth rather than crises, the authors instead looked for a plausible set of exogenous shocks that, if simulated in their models, generate something that resembled the Great Recession. These shocks remain unspecified, however, beyond stating that they emanate from 'households,' or 'technology.' *Neither even considered modifying their models to include the role of private debt*.[11]

Ireland started promisingly, with the thought that perhaps the underlying theory itself should be challenged: 'Indeed, the Great Recession's extreme severity makes it tempting to argue that new theories are required to fully explain it' (Ireland 2011: 31).

However, the apostate road was quickly abandoned, with the assertion that 'it would be premature to abandon existing models just yet.' One ground given for persevering with neoclassical models displayed the standard neoclassical ignorance of dynamic modeling, by asserting that: 'Attempts to explain movements in one set of endogenous variables, like GDP and employment, by direct appeal to movements in another, like asset

market valuations or interest rates, sometimes make for decent journalism but rarely produce satisfactory economic insights' (ibid.: 32).

Having dismissed the need for a change of approach, he went in search of 'shocks' that might explain why the economy so suddenly and for so long diverged from its equilibrium, with the objective of showing that the Great Recession was really no different to 'the two previous downturns in 1990–91 and 2001': 'this paper asks whether, in terms of its macroeconomics, the Great Recession of 2007–09 really stands apart from what came before […]' (ibid.).

Using his small-scale 'New Keynesian' model, Ireland concluded that unspecified 'adverse shocks' to the household's consumption preferences and the firm's technology caused all three recessions: 'the Great Recession began in late 2007 and early 2008 with a series of adverse preference and technology shocks in roughly the same mix and of roughly the same magnitude as those that hit the United States at the onset of the previous two recessions […]'. What made this recession different, however, was that the shocks went on for longer, and got bigger over time: 'The string of adverse preference and technology shocks continued, however, throughout 2008 and into 2009. Moreover, these shocks grew larger in magnitude, adding substantially not just to the length but also to the severity of the great recession […]' (ibid.: 48).

Ireland stated his positive conclusions for the New Keynesian approach halfway through the paper, claiming that his results: 'speak to the continued relevance of the New Keynesian model, perhaps not as providing the very last word on but certainly for offering up useful insights into both macroeconomic analysis and monetary policy evaluation' (ibid.: 33).

This is laughable, given both the author's methodology, and manifest ignorance of the fallacies in neoclassical thought – as evidenced by the manner in which he measured the gap between output during the recessions and the ideal level of output. He envisages a 'benevolent social planner,' who can derive a 'social welfare function' that reconciles all social conflict over the distribution of income, reproducing – I am sure without knowing the source – Samuelson's bizarre vision of capitalism as one big happy family:

> it is helpful to define a welfare-theoretic measure of the output gap, based on a comparison between the level of output that prevails in equilibrium and *the level of output chosen by a benevolent social planner* who can overcome the frictions associated with monetary trade and sluggish nominal price adjustment. *Such a planner chooses* the efficient level of output and the efficient amounts of labor to allocate to […] production […] *to maximize a social welfare function reflecting the same preference orderings over consumption and leisure embedded into the representative household's utility function*. (Ibid.: 38; emphases added)

McKibbin and Stoekel use a larger scale with six household-firm agents – one for each of six economic sectors (energy, mining, agriculture, manufacturing durables, manufacturing non-durables, and services) – and fifteen countries as well. As a New Keynesian model it allows for various 'imperfections,' and tellingly they remark that without 'short-run nominal wage rigidity' and a stylized but trivial role for money ('Money is introduced into the model through a restriction that households require money to purchase goods'), the model would simply predict that full-employment equilibrium would apply at all times:

> The model also allows for short-run nominal wage rigidity (by different degrees in different countries) and therefore allows for significant periods of unemployment depending on the labor-market institutions in each country. *This assumption, when taken together with the explicit role for money, is what gives the model its 'macroeconomic' characteristics*. (Here again the model's assumptions differ from the standard market-clearing assumption in most CGE models.) […]
>
> Although *it is assumed that market forces eventually drive the world economy to neoclassical steady-state growth equilibrium*, unemployment does emerge for long periods owing to wage stickiness, to an extent that differs between countries owing to differences in labor-market institutions. (McKibbin and Stoeckel 2009: 584; emphases added)

As with Ireland, they manipulate the shocks applied to their model until its short-run deviations from the steady state mimic what occurred during the Great Recession, and as with Ireland, one shock is not enough – three have to be used:

1. the bursting of the housing bubble, causing a reallocation of capital and a loss of household wealth and drop in consumption;
2. a sharp rise in the equity risk premium (the risk premium of equities over bonds), causing the cost of capital to rise, private investment to fall, and demand for durable goods to collapse;
3. a reappraisal of risk by households, causing them to discount their future labor income and increase savings and decrease consumption. (Ibid.: 587)

Not even this was enough to replicate the data: they also needed to assume that two of these 'shocks' – the risk tolerances of business and households – changed their magnitudes over the course of the crisis. A previous paper had found that 'a temporary shock to risk premia, as seems to have happened in hindsight, does not generate the large observed real effects,' so they instead considered an extreme shock, followed by an attenuation of it later: 'The question is then, what would happen if business and households initially assumed the worst – that is, a long lasting permanent rise in risk premia – but unexpectedly revised their views on risk to that of a temporary scenario 1 year later whereby things are expected to return to "normal"?' (ibid.: 582).

The procedure adopted in both these papers amplifies Solow's acerbic observation that 'New Keynesian' models fit the data better than 'New Classical' ones do, simply because the modelers add 'imperfections […] chosen by intelligent economists to make the models work better […]' (Solow 2001: 26). Now, to cope with the Great Recession – whose characteristics cannot be fitted even by the base New Keynesian model – the modeler also adds shocks that make the imperfections fit the data better, and even manipulates the shocks themselves until the model's output finally appears to match reality.

This is not science, but evasion. Adding tweaks to a deficient model – now including adding variable shocks – to avoid confronting the reality that

the model itself has failed, is the behavior of a 'degenerative scientific research program,' to use Lakatos's phrase.

Krugman's paper should have been better than these, in that at least he admits that one key component of reality that has been omitted in neoclassical economics – the role of private debt – needs to be incorporated to explain the Great Recession.

'Like a dog walking on its hind legs': Krugman's Minsky model

While Krugman's 'Debt, deleveraging, and the liquidity trap: a Fisher-Minsky-Koo approach' (Krugman and Eggertsson 2010) deserves some praise as the first neoclassical attempt to model Minsky after decades of ignoring him, the paper itself embodies everything that is bad in neoclassical economics.

This reflects poorly, not so much on Krugman – who has done the best he can with the neoclassical toolset to model what he thinks Minsky said – but on the toolset itself, which is so inappropriate for understanding the economy in which we actually live.

Attempts to increase the realism of the neoclassical model follow a mold that is as predictable as sunrise – but nowhere near as beautiful. The author takes the core model – which cannot generate the real-world phenomenon under discussion – and then adds some twist to the basic assumptions which, hey presto, generate the phenomenon in some highly stylized way. The mathematics (or geometry) of the twist is explicated, policy conclusions (if any) are then drawn, and the paper ends.

The flaw with this game is the very starting point, and since Minsky put it best, I'll use his words to explain it:

> Can 'It' – a Great Depression – happen again? And if 'It' can happen, why didn't 'It' occur in the years since World War II? These are questions that naturally follow from both the historical record and the comparative success of the past thirty-five years. *To answer these questions it is necessary to have an economic theory which makes great depressions one of the possible states in which our type of capitalist economy can find itself*. (Minsky 1982: xii; emphasis added)

The flaw in the neoclassical game is that it never achieves Minsky's final objective, because the 'twist' that the author adds to the basic assumptions of the neoclassical model are never incorporated into its core. The basic theory therefore remains one in which the key phenomenon under investigation – in this case, the crucial one Minsky highlights of how depressions come about – cannot happen. With the core theory unaltered, the performance is rather like that of a dog that learns how to walk on its hind legs on command, but which will revert to four-legged locomotion when the performance is over.[12]

Krugman himself is unlikely to stop walking on two legs – he enjoys standing out in the crowd of neoclassical quadrupeds – but the pack will return to form once this crisis ultimately gives way to tranquility.

However, one way in which Krugman doesn't stand out from the pack is how he treats rival schools of thought in economics: he ignores them.

The scholarship of ignorance and the ignorance of scholarship Krugman's paper cites nineteen works,[13] three of which are non-neoclassical – Fisher's classic 1933 'debt deflation' paper, Minsky's last book, *Stabilizing an Unstable Economy* (Minsky 1986), and Richard Koo's *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession* (Koo 2009). The other sixteen include one empirical study (McKinsey Global Institute 2010) and fifteen neoclassical papers written between 1989 (Bernanke and Gertler 1989) and 2010 (Woodford 2010) – five of which are papers by Krugman or his co-author.

Was this the best he could have done? Hardly! For starters, the one Minsky reference he used was, in my opinion, Minsky's worst book – and I'm speaking as someone in a position to know. Anyone wanting to get a handle on the Financial Instability Hypothesis from Minsky himself would be far better advised to read the essays in *Can 'It' Happen Again?* (Minsky 1982 [1963]), or his original book *John Maynard Keynes* (Minsky 1975) – which, despite its title, is not a biography, but the first full statement of his hypothesis.[14]

Krugman's ignorance of Minsky prior to the crisis was par for the course among neoclassical authors, since they only read papers published in what they call the leading journals – such as the *American Economic Review* – which routinely reject non-neoclassical papers without even

refereeing them.[15] Almost all academic papers on or by Minsky have been published in non-mainstream journals – the *American Economic Review* (*AER*), for example, has published a grand total of two papers on or by Minsky, one in 1957 (Minsky 1957) and the other in 1971 (Minsky 1971). If the *AER* and the other so-called leading journals were all you consulted as you walked up and down the library aisles, you wouldn't even know that Minsky existed – and most neoclassicals didn't know of him until after 2007.

Before the 'Great Recession' too, you might have been justified in ignoring the other journals – such as the *Journal of Post Keynesian Economics*, the *Journal of Economic Issues*, the *Review of Political Economy* (let alone the *Nebraska Journal of Economics and Business*, where several of Hyman's key papers were published) because these were 'obviously' inferior journals, where papers not good enough to make it into the *AER*, the *Economic Journal*, *Econometrica* and so on were finally published.

But after the Great Recession, when the authors who foresaw the crisis came almost exclusively from the non-neoclassical world (Bezemer 2009, 2010, 2011), and whose papers were published almost exclusively in the non-mainstream journals, neoclassical economists like Krugman should have eaten humble pie and consulted the journals they once ignored.

That might have been difficult once: which journals would you look in, if all you knew was that the good stuff – the models that actually predicted what happened – *hadn't* been published in the journals you normally consulted? But today, with the Internet, that's not a problem. Academic economists have as their bibliographic version of Google the online service Econlit (www.aeaweb.org/econlit/index.php), and there it's impossible to do even a cursory search on Minsky and not find literally hundreds of papers on or by him. For example, a search on the keywords 'Minsky' and 'model' turned up 106 references (including three by yours truly – Keen 1995, 1996, 2001b).

Twenty-seven of these are available in linked full text (one of which is also by yours truly; Keen 1995), so that you can download them direct to your computer from within Econlit, while others can be located by searching through other online sources, without having to trundle off to a

physical library to get them. To not have any references at all from this rich literature is simply poor scholarship. Were Krugman a student of mine, he'd have failed this part of his essay.

So in attempting to model a debt crisis in a capitalist economy, Krugman has used as his guide Fisher's pivotal paper, Minsky's worst book, and about ten neoclassical references written by someone other than himself and his co-author. How did he fare?

Mishandling an 'omitted variable' One thing I can compliment Krugman for is honesty about the state of neoclassical macroeconomic modeling before the Great Recession. His paper opens with the observation that 'If there is a single word that appears most frequently in discussions of the economic problems now afflicting both the United States and Europe, that word is surely "debt"' (Krugman and Eggertsson 2010: 1), and then admits that private debt played no role in neoclassical macroeconomic models before the crisis:

> Given both the prominence of debt in popular discussion of our current economic difficulties and the long tradition of invoking debt as a key factor in major economic contractions, one might have expected debt to be at the heart of most mainstream macroeconomic models – especially the analysis of monetary and fiscal policy. Perhaps somewhat surprisingly, however, it is quite common to abstract altogether from this feature of the economy. Even economists trying to analyze the problems of monetary and fiscal policy at the zero lower bound – and yes, that includes the authors – have often adopted representative-agent models in which everyone is alike, and in which the shock that pushes the economy into a situation in which even a zero interest rate isn't low enough takes the form of a shift in everyone's preferences. (Ibid.: 2)

This, along with the unnecessary insistence on equilibrium modeling, is the key weakness in neoclassical economics: if you omit so crucial a variable as debt from your analysis of a market economy, there is precious little else you will get right. So Krugman has taken at least one step in the right direction.

However, from this mea culpa, it's all downhill, because he made no fundamental shift from a neoclassical approach; all he did was modify his base 'New Keynesian' model to incorporate debt as he perceived it. On this front, he fell into the same trap that ensnared Bernanke, of being incapable of conceiving that aggregate debt can have a macroeconomic impact: 'Ignoring the foreign component, or looking at the world as a whole, the overall level of debt makes no difference to aggregate net worth – one person's liability is another person's asset' (ibid.: 3).

This one sentence established that Krugman failed to comprehend Minsky, who realized – as did Schumpeter and Marx before him – that growing debt in fact boosts aggregate demand:

> If income is to grow, the financial markets […] must generate an aggregate demand that, aside from brief intervals, is ever rising. For real aggregate demand to be increasing […] it is necessary that current spending plans, summed over all sectors, be greater than current received income […] It follows that over a period during which economic growth takes place, at least some sectors finance a part of their spending by emitting debt or selling assets. (Minsky 1982: 6)

Krugman also has no understanding of the endogeneity of credit money – that banks create an increase in spending power by simultaneously creating money and debt. Lacking any appreciation of how money is created in a credit-based economy, Krugman instead sees lending as simply a transfer of spending power from one agent to another: *neither banks nor money exist in the model he built*.

Instead, rather than modeling the economy as a single representative agent, he modeled it as consisting of two agents, one of whom was impatient while the other was patient. Debt was simply a transfer of spending power from the patient agent to the impatient one, and therefore the debt itself had no macroeconomic impact – it simply transferred spending power from the patient agent to the impatient one. The only way this could have a macroeconomic impact was if the 'impatient' agent was somehow constrained in ways that the patient agent was not, and that's exactly how Krugman concocted a macroeconomic story out of this neoclassical microeconomic fantasy: 'In what follows, we begin by setting

out a flexible-price endowment model in which "impatient" agents borrow from "patient" agents [where what is borrowed is not money, but "risk-free bonds denominated in the consumption good" (p. 5)], but are subject to a debt limit.'

To then generate a crisis, Krugman had to introduce an ad hoc and unexplained change to this debt limit: 'If this debt limit is, *for some reason*, suddenly reduced, the impatient agents are forced to cut spending; if the required deleveraging is large enough, the result can easily be to push the economy up against the zero lower bound. If debt takes the form of nominal obligations, Fisherian debt deflation magnifies the effect of the initial shock' (Krugman and Eggertsson 2010: 3; emphasis added)

He then generalized this with 'a sticky-price model in which the deleveraging shock affects output instead of, or as well as, prices' (ibid.), brought in nominal prices *without money* by imagining 'that there is a nominal government debt traded in zero supply […] We need not explicitly introduce the money supply' (ibid.: 9), modeled production – yes, the preceding analysis was of a no-production economy in which agents simply trade existing 'endowments' of goods distributed like manna from heaven – under imperfect competition (ibid.: 11), added a central bank that sets the interest rate (in an economy without money) by following a Taylor Rule, and on it went.

The mathematics was complicated, and real brain power was exerted to develop the argument – just as, obviously, it takes real brain power for a poodle to learn how to walk on its hind legs. But it was the wrong mathematics: it compared two equilibria separated by time, whereas truly dynamic analysis considers change over time regardless of whether equilibrium applies or not. And it was wasted brain power, because the initial premise – that aggregate debt has no macroeconomic effects – was false.

Krugman at least acknowledged the former problem – that the dynamics are crude: 'The major limitation of this analysis, as we see it, is its reliance on strategically crude dynamics. To simplify the analysis, we think of all the action as taking place within a single, aggregated short run, with debt paid down to sustainable levels and prices returned to full ex ante flexibility by the time the next period begins' (ibid.: 23).

But even here, I doubt that he would consider genuine dynamic modeling without the clumsy neoclassical device of assuming that all economic processes involve movements from one equilibrium to another. Certainly this paper remained true to the perspective he gave in 1996 when speaking to the European Association for Evolutionary Political Economy: 'I like to think that I am more open-minded about alternative approaches to economics than most, but *I am basically a maximization-and-equilibrium kind of guy. Indeed, I am quite fanatical about defending the relevance of standard economic models* in many situations […].' He described himself as an 'evolution groupie' to this audience, but then made the telling observation that:

> Most economists who try to apply evolutionary concepts start from some deep dissatisfaction with economics as it is. I won't say that I am entirely happy with the state of economics. But let us be honest: I have done very well within the world of conventional economics. I have pushed the envelope, but not broken it, and have received very widespread acceptance for my ideas. What this means is that I may have more sympathy for standard economics than most of you. My criticisms are those of someone who loves the field and has seen that affection repaid.

Krugman's observations on methodology in this speech also highlight why he was incapable of truly comprehending Minsky – because he starts from the premise that neoclassical economics itself has proved to be false, that macroeconomics must be based on individual behavior: 'Economics is about what individuals do: not classes, not "correlations of forces," but individual actors. This is not to deny the relevance of higher levels of analysis, but they must be grounded in individual behavior. *Methodological individualism is of the essence*' (Krugman 1996; emphases added)

No it's not: methodological individualism is one of the key flaws in neoclassical macroeconomics, as the SMD conditions establish. Economic processes have to be modeled at a higher level of aggregation, as Kirman argued (Kirman 1989: 138) and Minsky, in practice, did.

So while Krugman reached some policy conclusions with which I concur – such as arguing against government austerity programs during a

debt-deflationary crisis – his analysis is proof for the prosecution that even 'cutting-edge' neoclassical economics, by continuing to ignore the role of aggregate debt in macroeconomic dynamics, is part of the problem of the Great Recession, not part of its solution.

Conclusion: neat, plausible, and wrong

Mencken's aphorism suits not merely the Money Multiplier, but the whole of neoclassical economics: 'neat, plausible, and wrong.' If we are to avoid another Great Depression – more bleakly, if we are to get out of the one we are still in – then neoclassical economics has to be consigned to the dustbin of intellectual history. But that by itself is not enough: we need a replacement theory that does not make the many methodological mistakes that have made neoclassical economics such a singularly misleading and dangerous guide to the management of a capitalist economy.

The manner in which neoclassical economists have dealt with the crisis also makes a mockery of the basis on which neoclassical macroeconomics was based: its criticism of the preceding IS-LM 'Keynesian' models that they were based on many 'ad hoc' parameters – as Solow observed, 'the main argument for this modeling strategy has been a more aesthetic one: its virtue is said to be that it is compatible with general equilibrium theory, and thus it is superior to ad hoc descriptive models that are not related to "deep" structural parameters' (Solow 2007: 8). However, to cope with the Great Recession, neoclassical economists are now introducing ad hoc changes to these '"deep" structural parameters' – in order to explain why risk is suddenly re-evaluated and so on – and even introducing 'ad hoc' shocks. Neoclassical attempts to reproduce the crisis therefore fail the Lucas Critique which gave birth to this approach in the first place.

A complete, ready-made replacement does not exist. But there are alternative ways of thinking about economics that provide a good foundation on which an empirically grounded, non-ideological theory of economics can be built. I now turn to these alternatives, starting with the perspective that enabled me to be one of the very few economists who saw the Great Recession coming.

PART 3 | ALTERNATIVES

DIFFERENT WAYS TO THINK ABOUT ECONOMICS

13 | WHY I DID SEE 'IT' COMING

I was certainly not the only economist to expect that a serious economic crisis was imminent before the Great Recession began.

The post-Keynesian and Austrian schools of thought explicitly consider credit and money in their models of the economy, and many economists in these schools expected a crisis – the former group because of their familiarity with Hyman Minsky's Financial Instability Hypothesis, and the latter because of their familiarity with Hayek's argument about the impact of interest rates being held too low by government policy. However, the vast majority of these did not go public with their warnings.

Bezemer identified twelve individuals including myself who did publicly warn of the approaching crisis (Bezemer 2009, 2010, 2011), and a poll conducted by the *Real-World Economics Review* to decide who should win the inaugural Revere Award for Economics[1] resulted in an additional eighty-four individuals being nominated (Fullbrook 2010).

What distinguished me (and the late Wynne Godley) from the rest of these prescient and voluble few is that I had developed a mathematical model of how this crisis might come about. That model put into dynamic, disequilibrium form the economic vision of the late Hyman Minsky, which was in turn built on the insights of the great non-neoclassical thinkers Marx, Schumpeter, Fisher and Keynes. Minsky's strength was to weave these individually powerful and cohesive but incomplete analyses into one coherent tapestry that explained capitalism's greatest weakness: its proclivity to experience not merely economic cycles, but also occasional depressions that challenged the viability of capitalism itself.

The Financial Instability Hypothesis

Minsky's starting point was that, since the Great Depression had occurred, and since similar if smaller crises were a recurrent feature of the

nineteenth century, before 'Big Government' became the norm in market economies, an economic model had to be able to generate a depression as one of its possible outcomes: 'Can "It" – a Great Depression – happen again? And if "It" can happen, why didn't "It" occur in the years since World War II? These are questions that naturally follow from both the historical record and the comparative success of the past thirty-five years. *To answer these questions it is necessary to have an economic theory which makes great depressions one of the possible states in which our type of capitalist economy can find itself*' (Minsky 1982: 5; emphasis added).

For this reason, Minsky explicitly rejected neoclassical economics:

> The abstract model of the neoclassical synthesis cannot generate instability. When the neoclassical synthesis is constructed, capital assets, financing arrangements that center around banks and money creation, constraints imposed by liabilities, and the problems associated with knowledge about uncertain futures are all assumed away. For economists and policy-makers to do better we have to abandon the neoclassical synthesis. (Ibid.: xiii)

In place of the non-monetary, equilibrium-fixated, uncertainty-free, institutionally barren and hyper-rational individual-based reductionist neoclassical model, Minsky's vision of capitalism was strictly monetary, inherently cyclical, embedded in time with a fundamentally unknowable future, institution-rich and holistic, and considered the interactions of its four defining social entities: industrial capitalists, bankers, workers and the government.

I published my first paper on Minsky's hypothesis in 1995 (Keen 1995), and the following summary of Minsky's verbal model of a financially driven business cycle is reproduced from that paper.[2] I provide it verbatim here since its conclusion – written in 1993, long before neoclassical economists began to congratulate themselves about the 'Great Moderation' – shows that the calamity the world economy fell into in 2007/08 was not an unpredictable 'Black Swan' event, but something that was entirely foreseeable *with the right economic theory*:

Minsky's analysis of a financial cycle begins at a time when the economy is doing well (the rate of economic growth equals or exceeds that needed to reduce unemployment), but firms are conservative in their portfolio management (debt to equity ratios are low and profit to interest cover is high), and this conservatism is shared by banks, who are only willing to fund cash-flow shortfalls or low-risk investments. The cause of this high and universally practiced risk aversion is the memory of a not too distant system-wide financial failure, when many investment projects foundered, many firms could not finance their borrowings, and many banks had to write off bad debts. Because of this recent experience, both sides of the borrowing relationship prefer extremely conservative estimates of prospective cash flows: their risk premiums are very high.

However, the combination of a growing economy and conservatively financed investment means that most projects succeed. Two things gradually become evident to managers and bankers: 'Existing debts are easily validated and units that were heavily in debt prospered: it pays to lever' (Minsky 1982, p. 65). As a result, both managers and bankers come to regard the previously accepted risk premium as excessive. Investment projects are evaluated using less conservative estimates of prospective cash flows, so that with these rising expectations go rising investment and asset prices. The general decline in risk aversion thus sets off both growth in investment and exponential growth in the price level of assets, which is the foundation of both the boom and its eventual collapse.

More external finance is needed to fund the increased level of investment and the speculative purchase of assets, and these external funds are forthcoming because the banking sector shares the increased optimism of investors (Minsky 1980, p. 121). The accepted debt to equity ratio rises, liquidity decreases, and the growth of credit accelerates.

This marks the beginning of what Minsky calls 'the euphoric economy' (Minsky 1982, pp. 120–124), where both lenders and borrowers believe that the future is assured, and therefore that most investments will succeed. Asset prices are revalued upward as previous valuations are perceived to be based on mistakenly conservative grounds. Highly liquid, low-yielding financial instruments are devalued, leading to a rise in the interest rates offered by them as their purveyors fight to retain market share.

Financial institutions now accept liability structures for both themselves and their customers 'that, in a more sober expectational climate, they would have rejected' (Minsky 1980, p. 123). The liquidity of firms is simultaneously reduced by the rise in debt to equity ratios, making firms more susceptible to increased interest rates. The general decrease in liquidity and the rise in interest paid on highly liquid instruments triggers a market-based increase in the interest rate, even without any attempt by monetary authorities to control the boom. However, the increased cost of credit does little to temper the boom, since anticipated yields from speculative investments normally far exceed prevailing interest rates, leading to a decline in the elasticity of demand for credit with respect to interest rates.

The condition of euphoria also permits the development of an important actor in Minsky's drama, the Ponzi financier (Minsky 1982, pp. 70, 115 […]). These capitalists profit by trading assets on a rising market, and incur significant debt in the process. The servicing costs for Ponzi debtors exceed the cash flows of the businesses they own, but the capital appreciation they anticipate far exceeds the interest bill. They therefore play an important role in pushing up the market interest rate, and an equally important role in increasing the fragility of the system to a reversal in the growth of asset values.

Rising interest rates and increasing debt to equity ratios eventually affect the viability of many business activities, reducing the interest rate cover, turning projects that were

originally conservatively funded into speculative ones, and making ones that were speculative 'Ponzi.' Such businesses will find themselves having to sell assets to finance their debt servicing – and this entry of new sellers into the market for assets pricks the exponential growth of asset prices. With the price boom checked, Ponzi financiers now find themselves with assets that can no longer be traded at a profit, and levels of debt that cannot be serviced from the cash flows of the businesses they now control. Banks that financed these assets purchases now find that their leading customers can no longer pay their debts – and this realization leads initially to a further bank-driven increase in interest rates. Liquidity is suddenly much more highly prized; holders of illiquid assets attempt to sell them in return for liquidity. The asset market becomes flooded and the euphoria becomes a panic, the boom becomes a slump.

As the boom collapses, the fundamental problem facing the economy is one of excessive divergence between the debts incurred to purchase assets, and the cash flows generated by them – with those cash flows depending upon both the level of investment and the rate of inflation.

The level of investment has collapsed in the aftermath of the boom, leaving only two forces that can bring asset prices and cash flows back into harmony: asset price deflation, or current price inflation. This dilemma is the foundation of Minsky's iconoclastic perception of the role of inflation, and his explanation for the stagflation of the 1970s and early 1980s.

Minsky argues that if the rate of inflation is high at the time of the crisis, then though the collapse of the boom causes investment to slump and economic growth to falter, rising cash flows rapidly enable the repayment of debt incurred during the boom. The economy can thus emerge from the crisis with diminished growth and high inflation, but few bankruptcies and a sustained decrease in liquidity. Thus, though this course involves the twin 'bads' of

inflation and initially low growth, it is a self-correcting mechanism in that a prolonged slump is avoided.

However, the conditions are soon reestablished for the cycle to repeat itself, and the avoidance of a true calamity is likely to lead to a secular decrease in liquidity preference.

If the rate of inflation is low at the time of the crisis, then cash flows will remain inadequate relative to the debt structures in place. Firms whose interest bills exceed their cash flows will be forced to undertake extreme measures: they will have to sell assets, attempt to increase their cash flows (at the expense of their competitors) by cutting their margins, or go bankrupt. In contrast to the inflationary course, all three classes of action tend to further depress the current price level, thus at least partially exacerbating the original imbalance. The asset price deflation route is, therefore, not self-correcting but rather self-reinforcing, and is Minsky's explanation of a depression.

The above sketch basically describes Minsky's perception of an economy in the absence of a government sector. With big government, the picture changes in two ways, because of fiscal deficits and Reserve Bank interventions. With a developed social security system, the collapse in cash flows that occurs when a boom becomes a panic will be at least partly ameliorated by a rise in government spending – the classic 'automatic stabilizers,' though this time seen in a more monetary light. The collapse in credit can also be tempered or even reversed by rapid action by the Reserve Bank to increase liquidity. With both these forces operating in all Western economies since World War II, Minsky expected the conventional cycle to be marked by 'chronic and … accelerating inflation' (Minsky 1982, p. 85). However, by the end of the 1980s, the cost pressures that coincided with the slump of the early 1970s had long since been eliminated, by fifteen years of high unemployment and the diminution of OPEC's cartel power. The crisis of the late 1980s thus occurred in a milieu of low

> inflation, raising the specter of a debt deflation. (Keen 1995: 611–14)

I added the following qualification about the capacity for government action to attenuate the severity of a debt deflation – while not addressing its underlying causes – to my précis of Minsky in the first edition of *Debunking Economics*:

> If a crisis does occur after the Internet Bubble finally bursts, then it could occur in a milieu of low inflation (unless oil price pressures lead to an inflationary spiral). Firms are likely to react to this crisis by dropping their margins in an attempt to move stock, or to hang on to market share at the expense of their competitors. This behavior could well turn low inflation into deflation.
>
> The possibility therefore exists that America could once again be afflicted with a debt deflation – though its severity could be attenuated by the inevitable increase in government spending that such a crisis would trigger. America could well join Japan on the list of the global economy's 'walking wounded' – mired in a debt-induced recession, with static or falling prices and a seemingly intractable burden of private debt. (Keen 2001a: 254)

That a crisis might occur, and even that government action might attenuate it, was something that one could anticipate with Minsky's verbal economic theory. But a market economy is a complex system – the most complex social system that has ever existed – and its very complexity means that feedback effects might occur that are simply impossible to predict with a verbal model alone. For that reason, in my PhD I decided to attempt what Minsky had not succeeded in doing: to provide a mathematical model that did justice to the compelling verbal description he gave of debt deflation.

Modeling Minsky

Minsky did develop a mathematical model of a financially driven business cycle in his PhD, which resulted in the one paper he ever had published in a mainstream economic journal, the *American Economic Review* (Minsky 1957).[3] But the model was unsatisfactory for a number of reasons, and he subsequently abandoned it to stick with predominantly verbal reasoning.

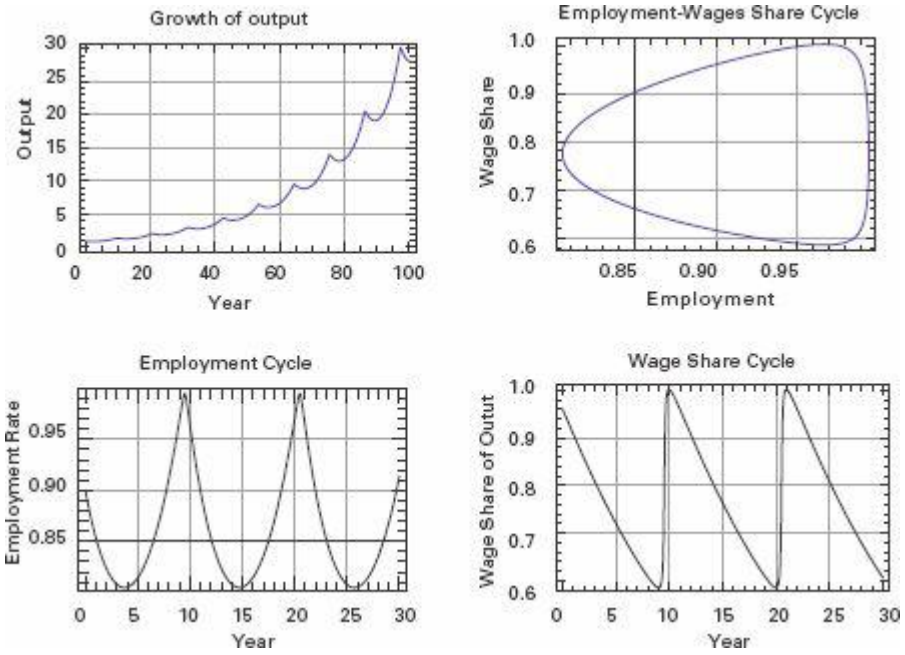
Minsky's failure to develop a satisfactory mathematical model was partly due to bad timing: the 1950s pre-dated the development of complexity theory, which made trying to build a model of his hypothesis virtually impossible. Minsky simply added a financial dimension to the dominant linear trade cycle model of the day, which was a particularly unsuitable foundation for his hypothesis.[4] In 1993, well after complexity theory had developed, I built my initial Minsky model using the far more suitable foundation of the cyclical growth model developed by the non-neoclassical economist Richard Goodwin (Goodwin 1967).

Goodwin's model considered the level of investment and the distribution of income in a simple two-class model of capitalism. A high initial wage and high rate of employment meant that wages absorbed most of output, so that profit was low – and therefore investment was low. The low rate of investment meant that the capital stock grew slowly (or fell because of depreciation), leading to a low rate of growth of output (or even falling output) and hence a growing unemployment rate – since population growth would then exceed the rate of economic growth.

The rising unemployment rate reduced workers' bargaining power, leading to stagnant or falling wages – which increased capitalists' profit share. They then increased investment, leading to a boom that drove the employment rate up, which strengthened the bargaining power of workers. Wages then rose and, because employment was high, wages absorbed most of output – which is where the cycle began.[5]

This was a classic dynamic model of 'circular causation' that is very common in biological modeling, but sadly a rarity in economics because of the neoclassical obsession with equilibrium. It also had a startling characteristic compared to the standard fare in economics: it was inherently cyclical. Given an arbitrary starting point, the model generated regular cycles in both the distribution of income and the employment rate. There

was no tendency toward equilibrium, but no tendency to breakdown either: the same cycle repeated for ever.



13.1 Goodwin's growth cycle model

Economists were falsely of the opinion that this was impossible. As John Hicks (remember him?) put it: 'A mathematically unstable system does not fluctuate; it just breaks down. The unstable position is one in which it will not tend to remain' (Hicks 1949).

As is so often the case, Hicks was right in particular and wrong in general. If they were unstable, then dynamic versions of the linear models that he and most neoclassical economists worked with would indeed break down – by returning impossible values for variables, such as negative prices or infinite levels of output. But Goodwin's model was inherently nonlinear, because two variables in the system – the wage rate and the level of employment – had to be multiplied together to work out wages and hence profits. As I explained in Chapter 9, nonlinear models can have persistent cycles without breaking down.

The professor of applied mathematics turned non-orthodox economist John Blatt observed that Goodwin's model was the best of the many dynamic economic models he had reviewed, and suggested that it would provide an excellent foundation for modeling financial dynamics in

capitalism. In stark contrast to the neoclassical obsession with equilibrium, one of Blatt's criticisms of Goodwin's basic model was that its equilibrium was not unstable:

> Of course, the model is far from perfect. In particular, we feel that the existence of an equilibrium which is not unstable (it is neutral) is a flaw in this model […] The first flaw can be remedied in several ways […] [such as] introduction of a financial sector, including money and credit as well as some index of business confidence. Either or both of these changes is likely to make the equilibrium point locally unstable, as is desirable […] But, while it is obvious that much work remains to be done, we have no doubt that the Goodwin model is the most promising of all the 'schematic models' of the trade cycle and well deserves further investigation. (Blatt 1983: 210–11)

I took up Blatt's suggestion in my PhD, by adding Keynes's model of how capitalists form conventions to cope with uncertainty, and Minsky's emphasis upon the role of debt in financing investment plans during a boom.

Of Keynes's three conventions to cope with uncertainty, the most important was the tendency to project forward current conditions: 'We assume that the present is a much more serviceable guide to the future than a candid examination of past experience would show it to have been hitherto' (Keynes 1937: 214).

A simple way to capture this in a mathematical model was to argue that capitalists would invest very little when the rate of profit today was very low, and invest a lot when the rate of profit was high. This was easily captured by replacing Goodwin's simple but unrealistic assumption that capitalists invested *all* their profits with a nonlinear relationship that meant investment would be less than profits when the rate of profit was low, and more than profits when the rate of profit was high.

Minsky improved upon Keynes by incorporating the insights of Schumpeter and Fisher on the essential role of debt in a capitalist economy: when capitalists' desire to invest exceeded retained earnings – as they would do during a boom – then capitalists would borrow to finance the

additional investment. I introduced this with a simple differential equation that said the rate of change of debt equaled investment minus profits.[6]

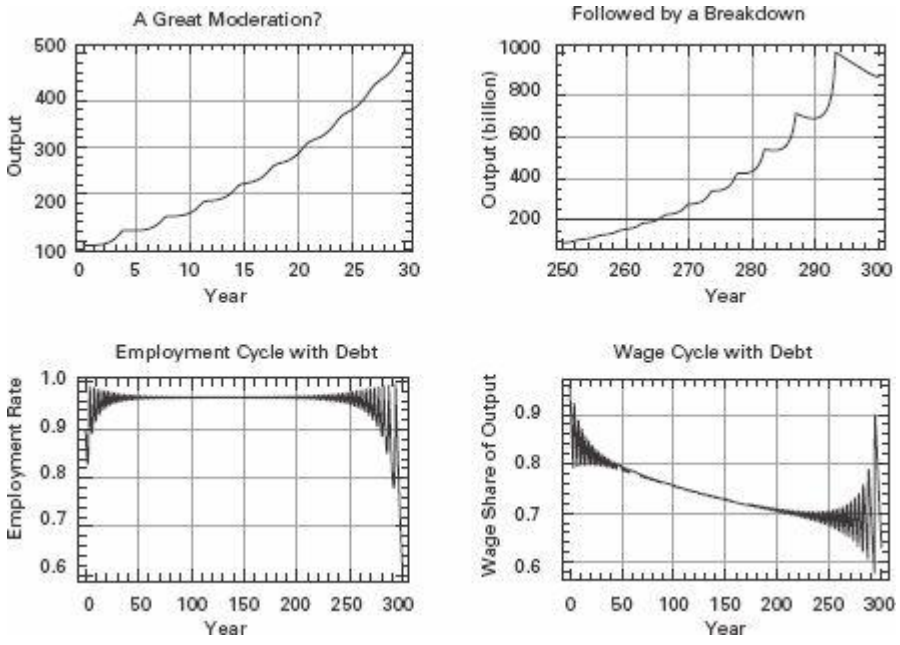
My first Minsky model This added one additional dynamic to Goodwin's model: the rate of change of debt, which rose when investment exceeded profits and fell when profits exceeded investment. During a boom, capitalists borrow to finance investment, and this drives up the debt-to-output ratio. During a slump, capitalists invest less than profits, and this reduces the debt-to-output ratio. The change in the debt ratio then affects the rate of profit, since profits are now equal to output, minus wages, minus interest on outstanding debt.

This simple extension to Goodwin's model dramatically altered its behavior. Goodwin's basic model generated fixed cycles indefinitely; this extended system could generate several different outcomes, ranging from a convergence to equilibrium values for income distribution, the employment rate and the debt-to-output ratio; cycles in all three variables of varying magnitudes over time; or a blowout in the debt-to-GDP ratio: a debt-induced depression.

The model also had three fascinating and, as it turned out, prescient characteristics.

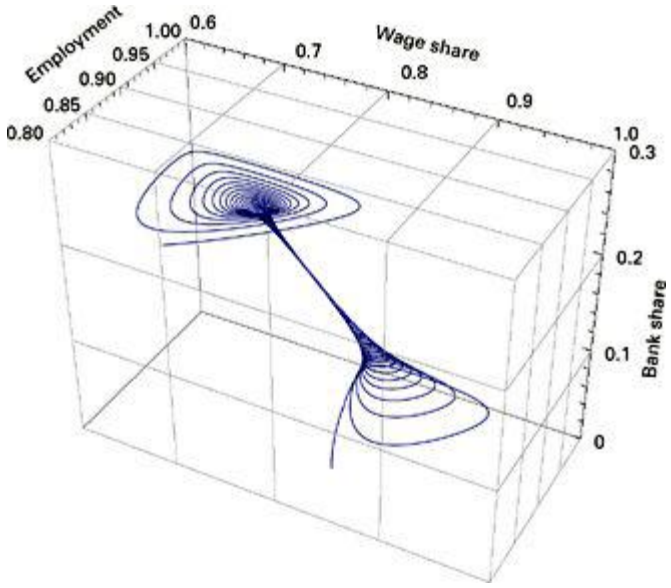
First, even though capitalists were the only borrowers in this simple model, the debt repayment burden actually fell on workers: the wages share of output fell as the debt level rose, while the profit share fluctuated around an equilibrium value.

Secondly, if the model did head toward a debt-induced breakdown, the debt-to-output ratio ratcheted up over time: debt would rise during a boom, reach a peak and then fall during a slump, but a new boom would begin before the debt-to-output ratio had dropped to its original value.



13.2 My 1995 Minsky model

Thirdly, the breakdown was preceded by a period of reduced volatility: fluctuations in employment and output would start off very large and then fall – the model generated a 'Great Moderation' *before* one appeared in the empirical record. But slowly, as the debt ratio rose even higher, the volatility started to rise again, until there was one last extreme cycle in which the debt level went so high that debt repayments overwhelmed the capacity of capitalists to pay.



13.3 The vortex of debt in my 1995 Minsky model

The economy then went into a death spiral as the level of debt overwhelmed the capacity of capitalists to service that debt. A 'Great Moderation' gave way to a 'Great Recession' – see Figure 13.3.

When I first completed this model in April 1992, the 'Great Moderation' had yet to begin, but the peculiar dynamics of the model struck me as remarkable. This led me to finish my first published paper on this model with a flourish that, at the time, seemed grandiose, but which ultimately proved to be prophetic:

> From the perspective of economic theory and policy, this vision of a capitalist economy with finance requires us to go beyond that habit of mind which Keynes described so well, the excessive reliance on the (stable) recent past as a guide to the future. *The chaotic dynamics explored in this paper should warn us against accepting a period of relative tranquility in a capitalist economy as anything other than a lull before the storm*. (Keen 1995: 634; emphasis added)

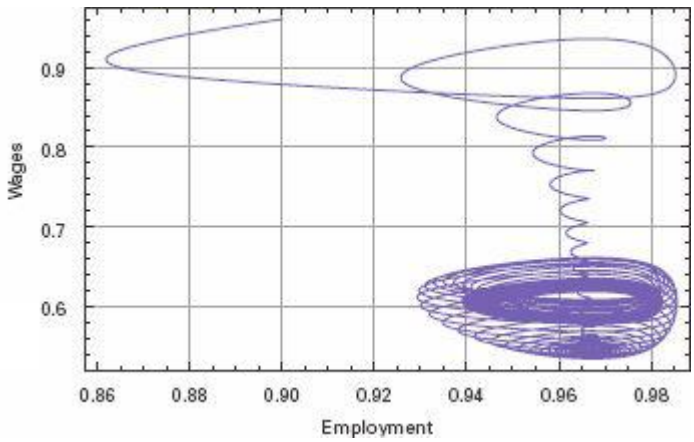
However, Minsky had also noted that government spending could stabilize an unstable economy. In that same paper I modeled this possibility by introducing government spending as an effective subsidy to capitalists that grew as unemployment rose and fell as it subsided – though workers receive unemployment benefits, the unemployed spend everything they get on consumption, so that corporations are the ultimate recipients of government welfare. Similarly, I modeled government taxation of business as rising as profits rose, and falling when profits fell.

As well as adding a fourth 'system state' to the model – the level of net government spending as a proportion of output – this modified the definition of profit. It was now output, minus wages, minus interest payments on debt, minus taxes plus the government subsidy.

In the model, the presence of government spending acted as a counterweight to the private sector's tendency to accumulate debt: a rising subsidy and falling taxes during a slump gave business additional cash flows with which to repay debt during a slump, while rising taxes and a

falling subsidy during a boom attenuated the private sector's tendency to accumulate debt.

The result was a system which was inherently cyclical, but in which the cycles stayed within manageable bounds: there was no systemic breakdown, as there had been in the pure private sector model. It was a pure limit cycle of the kind Blatt thought should be generated by a realistic model (Blatt 1983: 211).



13.4 Cyclical stability with a counter-cyclical government sector

Reality, I expected, lay somewhat between these two extremes of a private sector en route to a debt-induced breakdown, and a cyclical system kept within bounds by the 'automatic stabilizers' of government spending and taxation. The government sector modeled in this paper 'held the line' against rising unemployment, whereas in the real world governments had retreated from trying to restrain rising unemployment. I also knew that Ponzi-style behavior had become more dominant in the real world over time – something that I had not modeled explicitly, since in my model all borrowing led to productive investment. Also, though the models considered the role of private debt, they were only implicitly monetary, and I could not capture the impact of inflation or deflation upon the economy.

So there were ways in which I did not expect the real world to match my models. I resolved to extend them over time – to make them explicitly monetary, to model governments that gradually reduced their role as fiscal stabilizers, to incorporate borrowing for purely speculative reasons and so on – but in the immediate aftermath I was distracted from this agenda by the ferocious reaction that neoclassical economists had to the chapter 'Size

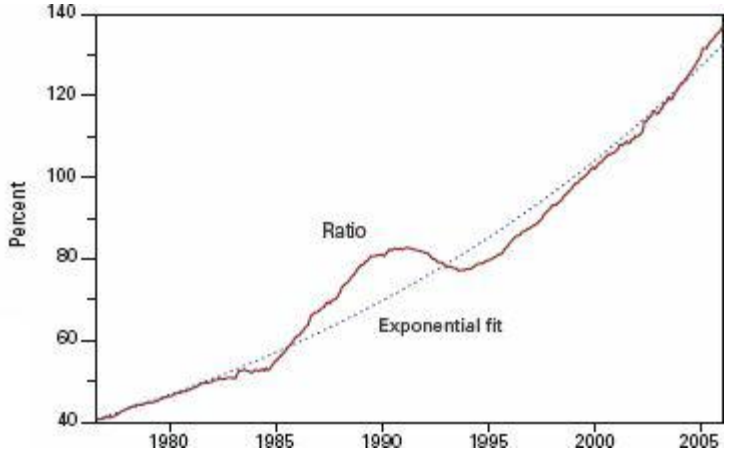
does matter' in the first edition of *Debunking Economics*. That dispute consumed my research energies in the four years from 2001 till 2005.

Finally in December 2005, I attempted to leave this argument behind and at long last write the book-length treatment of Minsky's hypothesis that I had first committed to do in 1998.[7] When I checked the ratio of private debt to GDP for the first time in over a decade, I quickly realized that a crisis would strike long before my technical book on how such crises came about would be ready.

Reality check, December 2005

The last thing I expected was that the real world would be in worse shape than my models implied, but that's what appeared to be the case in December 2005. While drafting an expert witness report on debt in a predatory lending case, I scribbled – before I had checked the data – that 'Debt to GDP ratios have been rising exponentially.' I expected that I'd need to attenuate that statement once I checked the data – the ratio would have been rising, I thought, though not at an exponential rate.

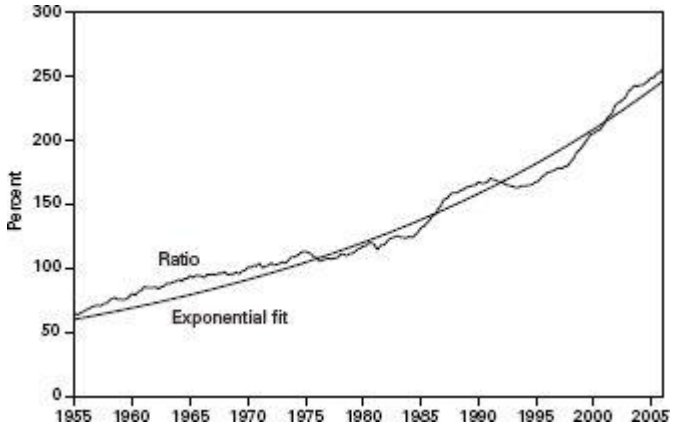
I vividly remember my stunned reaction when I first plotted the data, at about 1 a.m. on 22 December in Perth, Western Australia. Australia's private debt to GDP level had increased more than fivefold since the mid-1960s, and the rate of increase was clearly exponential – and it had a burst super-bubble in the 1980s, similar to the cyclical fluctuations in the debt-to-income ratio generated by my Minsky model.



13.5 Australia's private debt-to-GDP ratio, 1975–2005

I quickly downloaded the US Flow of Funds data to see whether Australia was unique. Obviously, it wasn't – see Figure 13.6. This was, as I expected, a global phenomenon. The US debt ratio was slightly less obviously exponential, but had increased even more than the Australian, and over a longer time period. Similar data could be found for most OECD nations, and especially the Anglo-Saxon countries.

Such an exponential rise in the debt ratio had to break, and when it did the global economy would be thrust into a downturn that would surely be more severe than those of the mid-1970s and early 1990s – the last times that the bursting of speculative bubbles had caused serious recessions. There was even the prospect that this would be an 'It' break: a debt-induced downturn so severe that the outcome would be not merely a recession, but a depression.



13.6 US private debt to GDP, 1955–2005

Someone had to raise the alarm, and I realized that, at least in Australia, I was probably that somebody. I once again put *Finance and Economic Breakdown* on the backburner, and devoted myself to warning the general public and policy-makers of the impending economic crisis. I began with media interviews, progressed to sending out a 'Debtwatch' report on debt coinciding with the Reserve Bank of Australia's monthly meetings from November 2006 (Keen 2006),[8] and in March 2007 I established the Debtwatch blog (www.debtdeflation.com/blogs).

Raising the alarm was not enough. I also had to dramatically improve my empirical understanding of the role of debt in a capitalist economy, and extend my Minsky model to cover the issues that I clearly had not paid

sufficient attention to in 1995: the impact of Ponzi finance, and the active role of the financial sector in financial crises.

The empirical dynamics of debt

The key insight about the role of debt in a capitalist society was provided by Schumpeter: in a growing economy, the increase in debt funds more economic activity than could be funded by the sale of existing goods and services alone: 'in real life total credit must be greater than it could be if there were only fully covered credit. The credit structure projects not only beyond the existing gold basis, but also beyond the existing commodity basis' (Schumpeter 1934: 95, 101).

Aggregate demand in a credit-driven economy is therefore equal to income (GDP) *plus the change in debt*. This makes aggregate demand far more volatile than it would be if income alone was its source, because while GDP (and the level of accumulated debt) changes relatively slowly, the change in debt can be sudden and extreme. In addition, if debt levels are already high relative to GDP, then the change in the level of debt can have a substantial impact on demand.

A numeric example illustrates this process (see Table 13.1). Consider an economy with a GDP of $1,000 billion that is growing at 10 percent per annum, where this is half due to inflation and half due to real growth, and which has a debt level of $1,250 billion that is growing at 20 percent per annum. Aggregate demand will therefore be $1,250 billion: $1,000 billion from GDP, and $250 billion from the increase in debt (which will rise from $1,250 billion to $1,500 billion over the course of the year).

Imagine that the following year, GDP continues to grow at the same 10 percent rate, but debt growth slows down from 20 percent per annum to 10 percent (the debt-to-GDP ratio will therefore stabilize at 150 percent). Demand from income will be $1,100 billion – 10 percent higher than the previous year – while demand from additional debt will be $150 billion (10 percent of the $1,500 billion level at the start of the year).

Aggregate demand in this second year will thus be $1,250 billion – exactly the same as the year before. However, since inflation is running at 5 percent, this will mean a fall in real output of about 5 percent – a serious recession. So just a slowdown in the rate of growth of debt can be enough

to trigger a recession. An absolute fall in debt isn't needed to cause problems, though it certainly will make things worse still.

TABLE 13.1 A hypothetical example of the impact of decelerating debt on aggregate demand

Year	0	1
Real growth	5%	5%
Inflation	5%	5%
Nominal GDP	$1,000	$1,100
Nominal debt	$1,250	$1,500
Debt growth rate	20%	10%
Growth in debt	$250	$150
Nominal aggregate demand	$1,250	$1,250
Change in nominal demand ($)	N/A	0
Change in nominal demand (%)	N/A	0
Real aggregate demand	$1,250	$1,187.5
Change in real demand	N/A	–5.0%

Schumpeter ignored the role of asset markets in the economy, so that in his model the increase in debt financed investment (and the sale of goods financed consumption). Therefore in his model, aggregate demand equals aggregate supply, but part of aggregate demand is debt-financed. In this example, demand financed by the sale of goods and services purchased $1,000 billion of consumer goods, while $250 billion of investment goods were bought on credit. Twenty percent of aggregate demand therefore came from rising debt.

Two consequences follow from this, of which Schumpeter was fully cognizant.

First, the expansion of credit must come, not from someone's savings being transferred to another person via a loan – which is the conventional model of how banks operate – but by the banking sector creating new money and credit 'out of nothing':

> [I]n so far as credit cannot be given out of the results of past enterprise […] it can only consist of credit means of payment created ad hoc, which can be backed neither by money in the strict sense nor by products already in existence […]
>
> It provides us with the connection between lending and credit means of payment, and leads us to what I regard as the nature of

> the credit phenomenon […] credit is essentially the creation of purchasing power for the purpose of transferring it to the entrepreneur, but not simply the transfer of existing purchasing power. (Ibid.: 106–7)

The banking sector therefore must have the capacity to create purchasing power – an issue I return to in the next chapter.

Secondly, the numerical example given here involves an unsustainable rate of growth of debt in the first year, so that there has to be a slowdown in the rate of growth of debt, which will cause a recession. However, the increased debt also helps create productive capacity for the economy, which can later be used to service the debt. There is thus a limit to the severity of cycles that can result: though excessive debt growth will cause a boom, and the inevitable slowdown in the growth of debt will cause a slump, the economy's capacity to produce is expanded by the growth of debt. Serious adjustments might be needed – falling prices, debt write-offs as some firms go bankrupt, and so on – but ultimately the economy will be able to reduce debt to manageable levels again, and growth will resume once more.

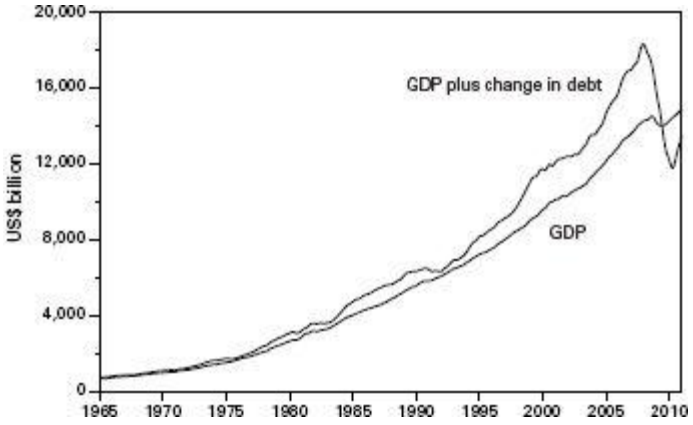
Minsky extended Schumpeter by considering Ponzi finance as well – lending to finance the speculative purchase of existing assets. Now, as well as aggregate demand being both income plus the change in debt, aggregate supply is both the output of new goods and services *and the net turnover of existing assets*. This breaches the virtuous cycle that Schumpeter saw between rising debt and a rising capacity to service that debt, because the money borrowed to buy assets adds to society's debt level without increasing its productive capacity. Thus when a slump follows a debt-fuelled boom, it is possible that debt servicing will exceed the economy's available cash flows – leading to not merely a recession, but a depression.

This Minskian process has been playing out in America ever since the mid-1960s when Minsky first developed his Financial Instability Hypothesis. Minsky himself identified 1966 as the time at which America made the transition from a productive to a Ponzi economy: 'A close examination of experience since World War II shows that the era quite naturally falls into two parts. The first part, which ran for almost twenty years (1948–1966), was an era of largely tranquil progress. This was

followed by an era of increasing turbulence, which has continued until today' (Minsky 1982: xiii).

Minsky's judgment was based largely on his financial interpretation of the US business cycle from that point on:

> The first serious break in the apparently tranquil progress was the credit crunch of 1966. Then, for the first time in the postwar era, the Federal Reserve intervened as a lender of last resort to refinance institutions – in this case banks – which were experiencing losses in an effort to meet liquidity requirements. The credit crunch was followed by a 'growth' recession, but the expansion of the Vietnam War promptly led to a large federal deficit which facilitated a recovery from the growth recession.
>
> The 1966 episode was characterized by four elements: (1) a disturbance in financial markets that led to lender-of-last-resort intervention by the monetary authorities; (2) a recession (a growth recession in 1966); (3) a sizable increase in the federal deficit; and (4) a recovery followed by an acceleration of inflation that set the stage for the next disturbance. The same four elements can be found in the turbulence of 1969–70, 1974–75, 1980, and 1981. (Ibid.: xiv–xv)



13.7 Aggregate demand in the USA, 1965–2015

Empirically, the late 1960s also marked the point at which the accumulated debt of the private sector exceeded 100 percent of GDP. From

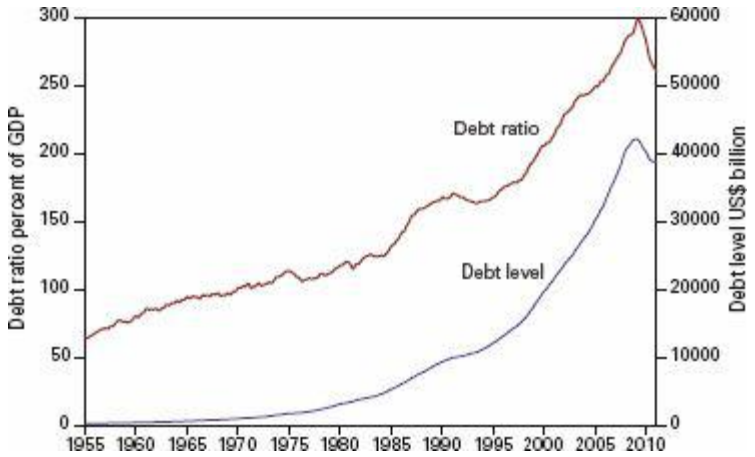
that point on, the dynamics of debt began to dominate macroeconomic performance in the USA – first generating a false prosperity, and then a calamitous collapse when the great debt bubble finally burst (see Figure 13.7).

TABLE 13.2 The actual impact of decelerating debt on aggregate demand

Year	2007/08	2008/09
Real growth	2.3%	–2.7%
Inflation	4.3%	0%
Nominal GDP	$14.29 tn	$14.19 tn
Nominal debt	$40.6 tn	$42.1 tn
Debt growth rate	28.1%	10.7%
Growth in debt	$4 tn	$1.52 tn
Nominal aggregate demand	$18.3 tn	$15.7 tn
Change in nominal demand ($)	N/A	–$2.6 tn
Change in nominal demand (%)	N/A	–14.2%
Real aggregate demand	$18.3 tn	$15.7 tn
Change in real demand[1]		–14.2%

Note: 1 The change in real demand was the same as the change in nominal demand since inflation was effectively zero in 2009

For the first time since the Great Depression, the aggregate level of private debt began to fall in January 2009. But the economic downturn began well before, when the rate of growth of debt slowed from its peak level, just as the numerical example illustrates.

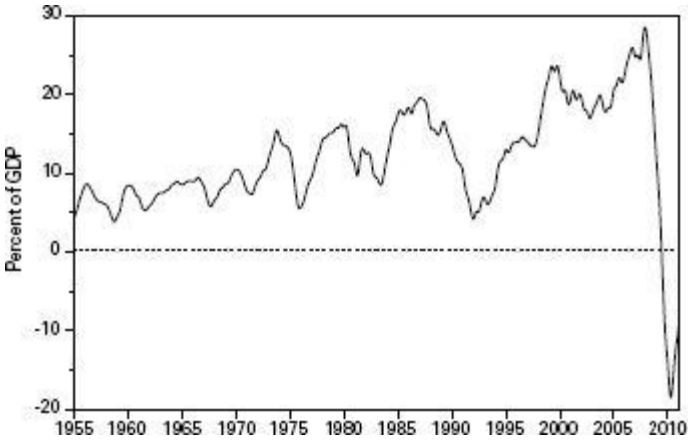


13.8 US private debt

The debt bubble went out with a bang: the increase in private sector debt in 2008, the final year of the bubble, was a truly stupendous $4 trillion, which boosted aggregate demand from GDP alone by over 28 percent. A year later, debt was growing by 'only' $1.5 trillion, with the result that aggregate demand slipped from its peak level of US$18.3 trillion in 2008 to $15.7 trillion at the beginning of 2009. Though GDP had fallen slightly over calendar year 2009 – from $14.3 trillion to $14.2 trillion – by far the biggest hit to the USA's solar plexus came simply from a slowdown in the rate of growth of debt. Though real GDP fell by a mere 2.7 percent, aggregate demand fell by a massive 14.2 percent – see Table 13.2.

The year 2008 thus brought to a close a period of literally half a century in which private debt had always been growing, and thus adding to aggregate demand. This of itself was not inherently a problem: as both Schumpeter and Minsky argued, rising debt is necessary to finance entrepreneurial activity and to enable the economy to grow. The problem for America, and most of the OECD, was that this increase in debt was rising relative to GDP – indicating that what was being funded was not good, Schumpeterian innovation, but bad Ponzi-finance speculation. The annual increase in debt, which had hovered around 5 percent of GDP in the 1950s and 1960s, rose in a series of peaks and troughs to the 28 percent peak of 2008, from where it plunged to a maximum rate of decline of over 18 percent in early 2010 – see Figure 13.9.

The $2.6 trillion drop in aggregate demand hit America's asset markets hard. Though the Dow Jones rallied towards the end of the year, it closed 34 percent down – a bone-crushing decline in the apparent wealth of America's stockholders (see Figure 13.10).

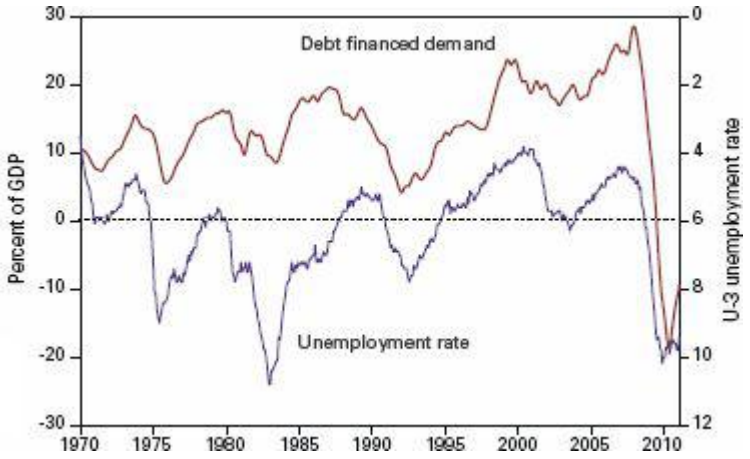


13.9 The change in debt collapses as the Great Recession begins

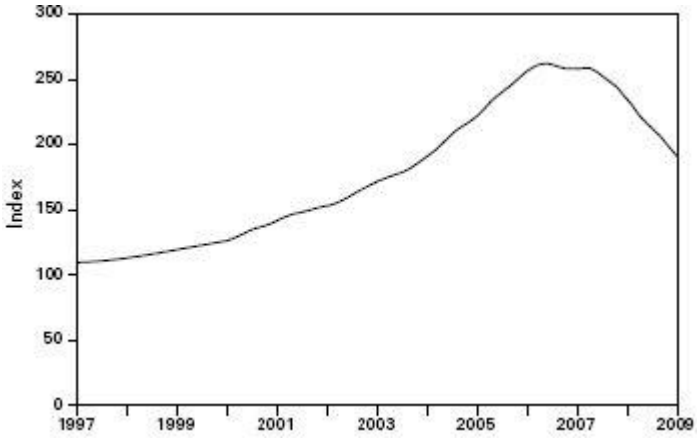
13.10 The Dow Jones nosedives

The long bubble in the housing market – which neoclassical economists like Ben Bernanke had strenuously denied was a bubble – burst under the weight of sheer fraud involved in subprime lending, well before the debt bubble propelling it started to slow.[9] It continued its decline relentlessly in 2008/09, with house prices falling another 19 percent (in real terms) on top of the 10 percent decline from their peak in March 2006 – see Figure 13.12.

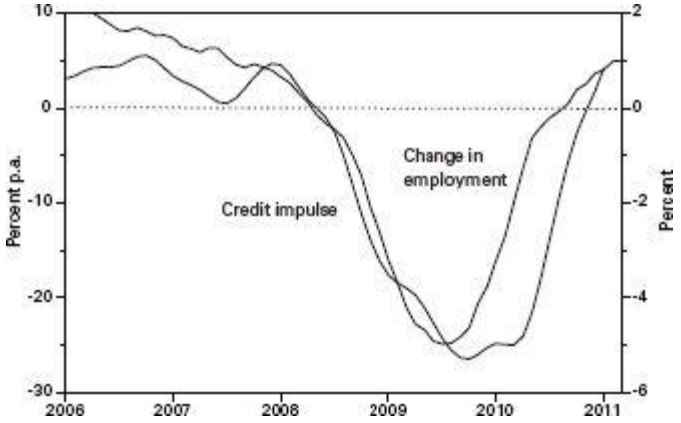
Unemployment rose from 4.4 percent at the beginning of 2007 to 5.5 percent at its end, and then to 7.6 percent as 2009 began. Here the hand of debt was clearly visible, for the simple reason that, since the change in debt is a major component of aggregate demand, and aggregate demand determines employment, *unemployment rises if the rate of change of debt falls* (and vice versa). As the level of debt has risen relative to GDP, the ebb and flow of unemployment has fallen more and more under the sway of changes in the level of private debt.



13.11 The correlation of debt-financed demand and unemployment



13.12 The housing bubble bursts



13.13 The Credit Impulse and change in employment

The dominance of debt has been obvious, not only in the collapse into the Great Recession, but even in the apparent recovery from it in late 2010 and early 2011 (a recovery that I believe will prove temporary, and which is also exaggerated by unreliable government statistics). Here an apparent paradox emerges: because aggregate demand is the sum of GDP plus the change in debt, the rate of change of aggregate demand can be boosted by *a slowdown in the rate at which debt is falling*.

The logic here is a simple extrapolation from the observation that the level of aggregate demand is the sum of GDP[10] plus the change in debt: given this, the *change* in aggregate demand is equal to the change in GDP plus the *acceleration* of debt. Therefore the factor that determines debt's impact upon the rate of economic growth – and hence the change in the rate of unemployment – is not the rate of change of debt, but *the rate of change of its rate of change*.

Biggs, Mayer and Pick, who first made this observation, noted that it had a seemingly counter-intuitive outcome that the economy can receive a boost from credit, even if the aggregate level of debt is falling, so long as the rate of that fall decreases: 'the flow of credit and GDP can increase even while the stock of credit is falling' (Biggs, Mayer et al. 2010: 5). They measured the impact of the acceleration of credit on changes in aggregate demand using the ratio of the acceleration of debt to GDP (which they termed 'the Credit Impulse'; ibid.: 3), and this measure clearly illustrated their apparently bizarre conclusion that the slight recovery in late 2010 was driven in large measure by a slowdown in the rate of deceleration of credit – see Figure 13.13.[11]

There are thus three factors that need to be considered to understand the impact of debt on a capitalist economy: the level of debt, the rate of change of debt, and its rate of acceleration – all measured with respect to the level of GDP.

Box 13.1 Definitions of unemployment

The official definition of unemployment has been reworked numerous times, in ways that reduce the recorded number, so much so that the published levels drastically understate the actual level. The official OECD

definition (see stats.oecd.org/glossary/detail.asp?ID=2791) requires that those recorded as unemployed must be both available for work and actively looking for work in the reference period, which excludes those who have become discouraged by the sheer unavailability of employment opportunities during a major recession, but many OECD countries have further tailored the definition to reduce the recorded numbers.

The Australian government's definition is typical here: in addition to the OECD requirements, it also records as employed people who 'worked *for one hour* for pay, profit, commission or payment in kind in a job or business, or on a farm; or worked *for one hour or more without pay* in a family business or on a farm' (McLennan 1996: 47). To regard someone who has worked only one hour in a week as employed is simply absurd – at least fifteen hours of work at the minimum wage are needed to be paid even the equivalent of unemployment benefits.

Similar distortions apply in other countries. The USA, for example, ceases counting someone as unemployed if they have been out of work for more than a year – a change in definition introduced in 1994 (see en.wikipedia.org/wiki/Unemployment#United_States_Bureau_of_Labor_Statistics and en.wikipedia.org/wiki/Current_Population_Survey#Employment_classification for more details). Abuses of statistics like this have prompted private citizens to record what official statistics ignore. The opinion-polling organization Roy Morgan Research (www.roymorgan.com.au/) now publishes its own survey of Australian unemployment, which it puts at 7.9 percent versus the recorded figure of 5.5 percent (the not-seasonally-adjusted figure as of January 2011).

Shadowstats (www.shadowstats.com/alternate_data/unemploymentcharts) maintains an alternative measure for the USA that includes long-term discouraged workers. This is now more than twice as high as the official US measure: at the time of writing (February 2011), the official U-3 measure was 9.0 percent, while the Shadowstats measure was 22.2 percent.

This, plus changes in the structure of employment, make comparisons with past economic crises like the Great Depression very difficult. John Williams, the founder of Shadowstats, estimates that his measure of unemployment would have shown that 34–35 percent of the workforce was

unemployed during the Great Depression – versus the 25 percent actually recorded back then, since the proportion of the population working on farms was much higher in the 1930s than now (27 percent then versus 2 percent now). The workers who were underemployed on farms – but nonetheless fed – reduced the numbers officially recorded as unemployed back then.

Given these problems, I regard the US's U-6 measure of unemployment today – which includes those who have been unemployed for two years or less – as more comparable to the Great Depression figures than its U-3 measure, which omits those who have been unemployed for a year or more. On that basis, one in six Americans are out of work today, versus the peak rate of one in four during the Great Depression. The current crisis, though it is called the Great Recession, is therefore really a depression too.

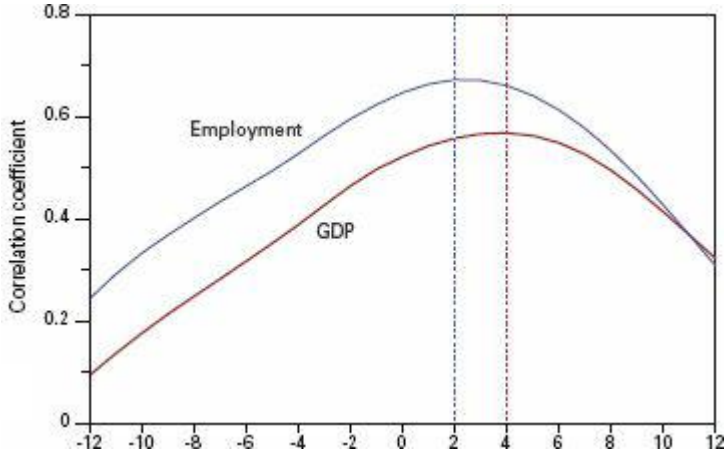
The first factor indicates the aggregate burden that debt imposes upon society. Since the level of debt is a stock, while the level of GDP is a flow (of income per year), the ratio tells us how many years of income it would take to reduce debt to zero. Of course, a target of zero debt is neither feasible nor desirable – as explained earlier, some debt is necessary to support entrepreneurial innovation. But the ratio indicates how debt-encumbered an economy has become, and the larger it is, the longer it will take to get back to any desired lower level.

It also provides the best measure of the burden the financial sector imposes upon the economy, since the net cost of the financial sector is the level of debt (multiplied by the inflation-adjusted gap between the rate of interest on loans and that on deposits – a gap that has been relatively constant, though the nominal and real rates of interest themselves have been very volatile).

The second factor indicates how much aggregate demand is being generated by rising debt – or reduced by falling debt. When the economy is growing, so too will credit, and again this is not a bad thing when that debt finances investment. The danger arises when the rate of growth of debt becomes a substantial determinant of overall demand – as it has in the Ponzi economy the USA has become. A large debt-financed contribution to

aggregate demand will almost certainly have a large component of Ponzi finance behind it, and such an increase necessarily requires a decline in debt-financed spending in the near future, which will usher in a recession.

The third factor is the best leading indicator of whether employment and the economy are likely to grow in the near future. The Credit Impulse leads both changes in GDP and changes in employment, with the lead (in the USA) being about two months to employment and four months to GDP.



13.14 Correlation of credit impulse and change in employment and GDP

The Credit Impulse is also the key financial source of capitalism's inherently cyclical nature. To maintain a stable rate of employment, the rate of growth of aggregate demand has to equal the rate of growth of employment and labor productivity, which are both relatively stable. But since the rate of growth of aggregate demand depends on the rate of growth of GDP and the *acceleration* of debt, a stable rate of growth of aggregate demand requires a constant acceleration of debt.

The only level at which this is possible is zero. Just as maintaining a constant *positive* rate of acceleration while driving a car is impossible – since otherwise the car would ultimately be travelling faster than the speed of light – a constant positive rate of acceleration of debt can't be maintained, because this would mean that debt would ultimately be infinitely larger than GDP. Since in the real world it is impossible for the acceleration of debt to always be zero, the economy will therefore necessarily have cycles driven by the expansion and contraction of credit.

These three factors – the level of debt, its rate of change, and its acceleration – interact in complex ways that are best explained by an

analogy to driving in which the debt-to-GDP ratio is like the distance back to your starting point, its rate of change relative to GDP is like the speed of the car, and the Credit Impulse is like the car's acceleration or deceleration.

A low ratio of debt to GDP is like having taken a short drive – say, from Los Angeles to Phoenix (a distance of 370 miles). It's easy to get back to LA at any time, and the return journey is not something one has to plan all that much for. A high ratio is like a drive from LA to New York: it's a huge distance (2,800 miles), and the drive back – which corresponds to reducing the debt to GDP ratio – will take a long time.

The rate of change of debt (with respect to GDP) is like your speed of travel – the faster you drive, the sooner you'll get there – but there's a twist. On the way out, increasing debt makes the journey more pleasant – the additional spending increases aggregate demand – and this experience is what fooled neoclassical economists, who ignore the role of debt in macroeconomics, into believing that the economy was experiencing a 'Great Moderation.' But rising debt increases the distance you have to travel backwards when you want to reduce debt, which is what the USA is now doing. So rising debt feels great on the outward drive from LA east (increasing debt), but lousy when you want to head home again (and reduce debt).

The Credit Impulse is like acceleration – it's a measure of the g-forces, so to speak, generated by either rapid acceleration or rapid deceleration. Acceleration in the debt level felt great on the way up: it was the real source of the booms in the Ponzi economy that the USA has become. Equally, acceleration in the opposite direction – in effect going backwards at an accelerating speed – is terrifying: as the rate of decline in debt increases, the fall in aggregate demand increases and unemployment explodes.

The interactions of the level of debt, rate of growth of debt and the Credit Impulse are akin to those between distance, speed and acceleration as well – and here I'll limit my analogy to the last few years, when America went from increasing debt – the drive from LA to New York – and then abruptly changed direction into deleveraging.

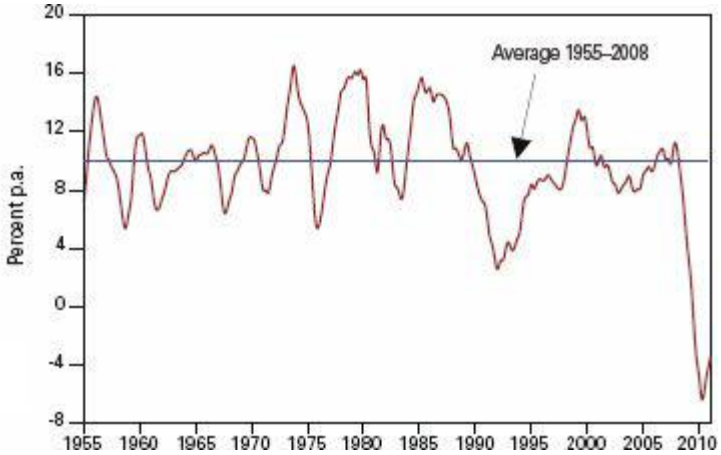
The reversal of direction necessarily involves your acceleration changing from zero or positive to negative, and it feels dreadful: imagine the feeling of slamming on the brakes, putting the car in reverse, and then driving backwards at an accelerating speed.

At some point, however, you will reach the maximum reverse speed of the car, and at that point the terrifying feeling of driving backwards more rapidly will give way to merely the unpleasant feeling of driving backwards at high speed. If you then start driving backwards less rapidly, you will actually feel a positive acceleration – *even though you are still driving backwards*. However, if you keep slowing down your reverse speed, then at some point you will reverse direction, and start heading back towards New York again. You can't maintain positive acceleration indefinitely without at some point changing from a negative to a positive velocity, and thus resuming your journey towards a place that you were initially trying to leave.

We can now get a handle on why this recession has been so extreme compared to its post-World War II predecessors, and why I believe that the crisis has many years to run.

First, all three debt indicators reached levels that are unprecedented in the post-World War II period. The debt-to-GDP ratio, which began the post-war period at barely 50 percent, increased by a factor of 6 in the subsequent five decades to reach a peak of 298 percent of GDP in early 2009.

Secondly, while private debt itself grew at a relatively constant if volatile 10 percent per annum between 1955 and 2008, the debt-financed proportion of aggregate demand rose from 5 percent in the 1950s to 28 percent in 2008.

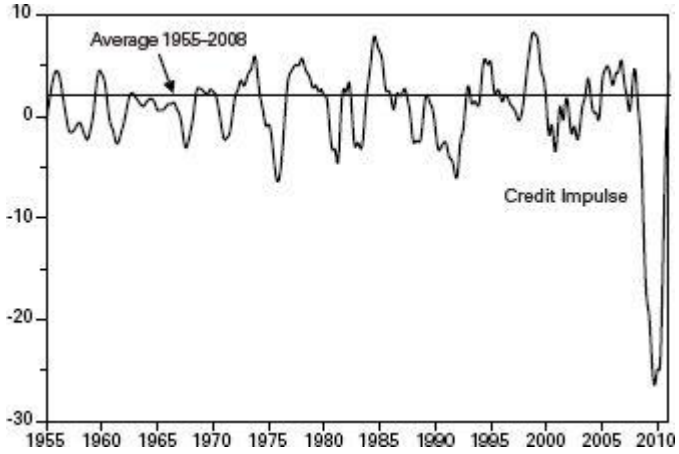


13.15 Relatively constant growth in debt

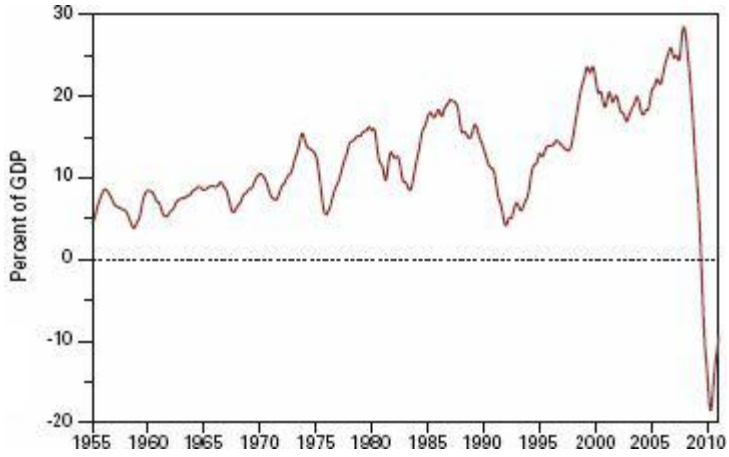
This occurred because the rate of growth of nominal debt was about 3 percent higher than that of nominal GDP from 1945 till 2008. The impact of

rising debt on aggregate demand therefore doubled every twenty-three years.[12]

It then plunged to minus 19 percent in early 2010 – an unprecedented event in post-World War II economic history. This debt level is still falling, though the rate of fall has slowed in recent times, from a peak rate of minus 19 percent of GDP in early 2010 to minus 12 percent in September 2010 (the last date at which debt data were available at the time of writing).



13.16 The biggest collapse in the Credit Impulse ever recorded



13.17 Growing level of debt-financed demand as debt grew faster than GDP

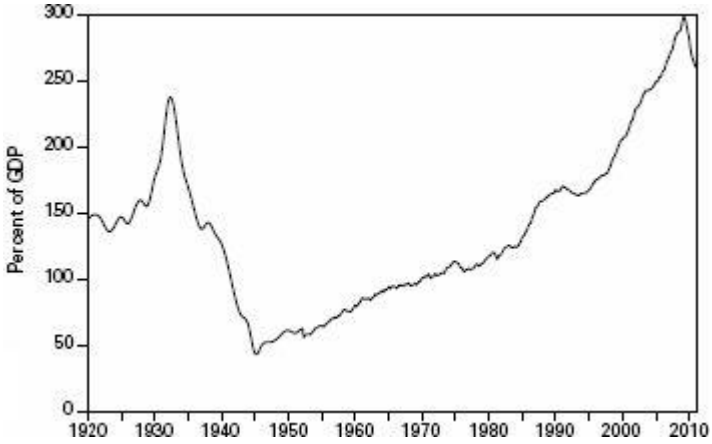
Thirdly, the Credit Impulse averaged plus 1.2 percent from 1955 till 2008, and then hit at an unprecedented minus 27 percent in 2009 at the depths of the downturn. It is now returning toward zero – which in part reflects its inevitable return toward zero as deleveraging becomes entrenched.[13] This puts far less drag on aggregate demand, but also

removes the 'turbo boost' that a positive Credit Impulse gave to growth in the previous half-century. The Credit Impulse will also tend to be negative while deleveraging continues, just as it tended to be positive when rising debt was boosting aggregate demand. This means the economy will have a tendency toward recessions rather than booms until the debt-to-GDP ratio stabilizes at some future date.

The interaction of these three factors will determine the economic future of the United States (and many other OECD nations, which are in a similar predicament).

The Credit Impulse, as the most volatile factor, will set the immediate economic environment. While it remains negative, the rate at which the USA is deleveraging accelerates, so it therefore had to rise again at some stage – as it has since mid-2009. This will accelerate aggregate demand, but it can't lead to a sustained rise in aggregate demand without causing the debt-to-GDP ratio to rise. That is extremely unlikely to happen, since even after the deleveraging of the last two years, the aggregate level of private debt is 100 percent of GDP higher than it was at the start of the Great Depression.

These dynamics of debt were the key cause of both the Great Moderation and the Great Recession, yet they were completely ignored by neoclassical economists because of their fallacious belief that changes in private debt have no macroeconomic effects (Bernanke 2000: 24). Therefore, far from making sure that 'It' won't happen again, as Bernanke asserted in 2002, by ignoring and in fact abetting the rise in private debt, neoclassical economists have allowed the conditions for another Great Depression to develop. Worse, a comparison of today's debt data to those from 1920–40 shows that the debt-deflationary forces that have been unleashed in the Great Recession are far larger than those that caused the Great Depression – see Figure 13.18.

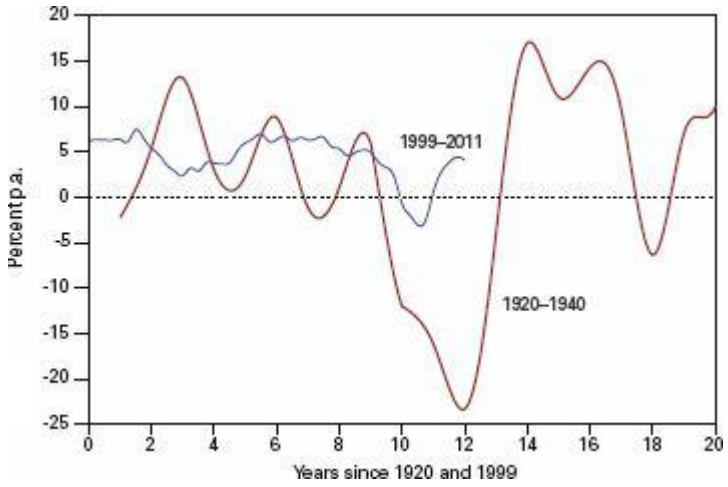


13.18 The two great debt bubbles

Debt deflation then and now

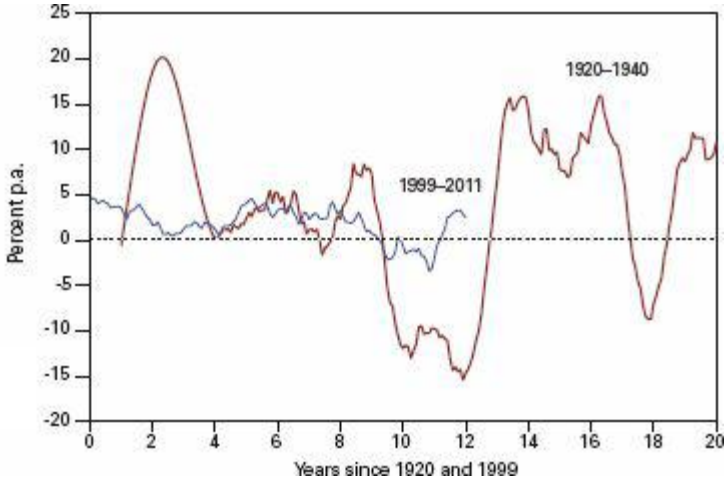
Comparing the 1920s–1940s to now – the Roaring Twenties and the Great Recession to the 'Noughty Nineties'[14] and the Great Recession – is feasible, but complicated both by differences in the economic circumstances at the time, and differences in the quality of the statistics.

A major complication is the extreme volatility in economic performance over the 1920s – no one was writing about 'the Great Moderation' back then. The decade began and ended with a depression, and recorded output fluctuated wildly. The average increase in nominal GDP over 1921–29 was 4.5 percent, but it fluctuated wildly from –2 to +13 percent, with a standard deviation of 4.4 percent. In contrast, the Noughty Nineties recorded a higher rate of nominal growth of 5.3 percent, and this was very stable, ranging between 2.6 and 6.6 percent with a standard deviation of only 1.4 percent.



13.19 Change in nominal GDP growth then and now

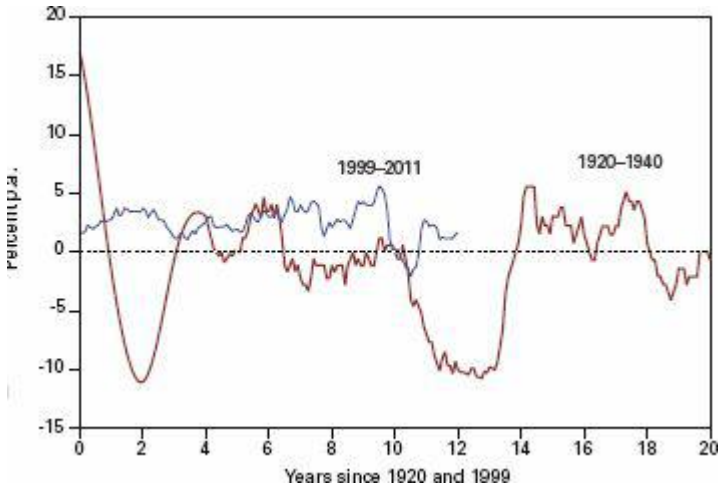
However, as well as being a decade of stock market speculation, the 1920s also saw serious Schumpeterian investment and 'creative destruction.' It was the decade of the Charleston and *The Great Gatsby*, but it was the decade of the production line, technological innovation in manufacturing and transportation, and the continuing transformation of American employment from agriculture to industry. The average rate of real economic growth was therefore higher in the 1920s than in the period from 1999 to 2009 – though disentangling this from the gyrations in the price level is extremely difficult.



13.20 Real GDP growth then and now

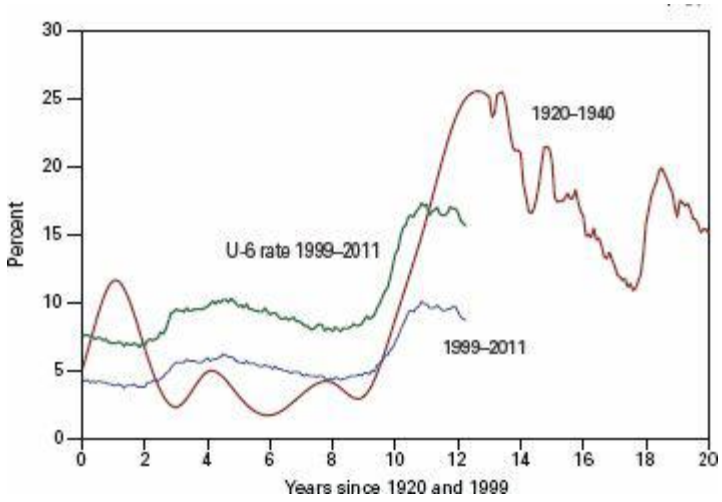
For example, the nominal rate of growth in 1922/23 was 13 percent, but the 'real' rate was an even higher 20 percent. This impossible level reflected

the simultaneous recovery from deflation of over 10 percent to inflation of 3 percent, and unemployment falling – and hence output rising – from 12 percent to 2.5 percent as the economy recovered from the depression of January 1920 to June 1921.



13.21 Inflation then and now

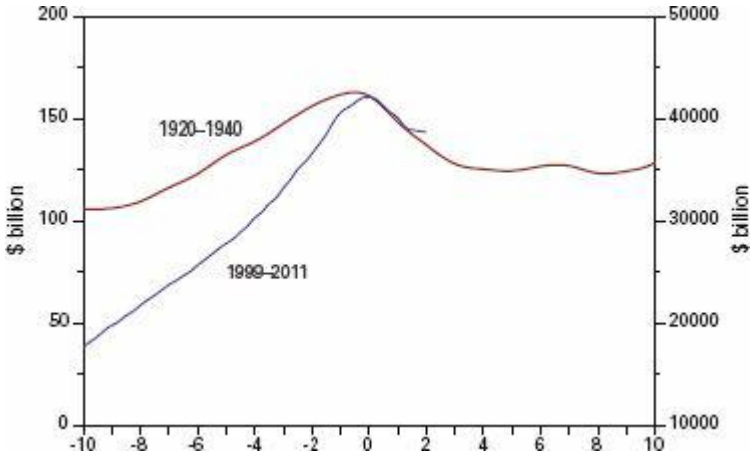
Overall, the rate of unemployment is the best means to compare the two periods, but here we run into the distortions caused by politically motivated redefinitions of the unemployment rate since the late 1970s (see Box 13.1). The U-3 measure for 1999–2009 averages 5 percent, only marginally higher than the average of 4.7 percent for 1920–29; but the U-6 measure for 1999–2009 averages 8.8 percent, and I regard this as a fairer comparison of the two periods.



13.22 Unemployment then and now

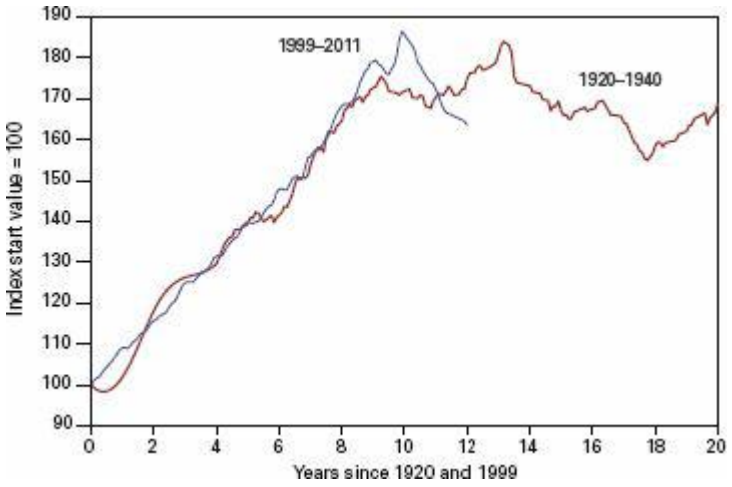
The upshot of all this is that the Roaring Twenties saw more real growth than the Noughty Nineties, and this masked the importance of debt at the time. But categorically, the fundamental cause of both the Great Depression and the Great Recession was the bursting of a debt-financed speculative bubble that had fueled the false but seductive prosperity of the previous decade.

The Great Depression remains the greatest economic crisis that capitalism has ever experienced, but on every debt metric, the forces that caused the Great Recession are bigger.[15] Private debt rose 50 percent over the 1920s, from $106 billion (yes, billion) in 1920 to $161 billion by 1930; it rose from $17 trillion to $42 trillion between 1999 and 2009 – a 140 percent increase.



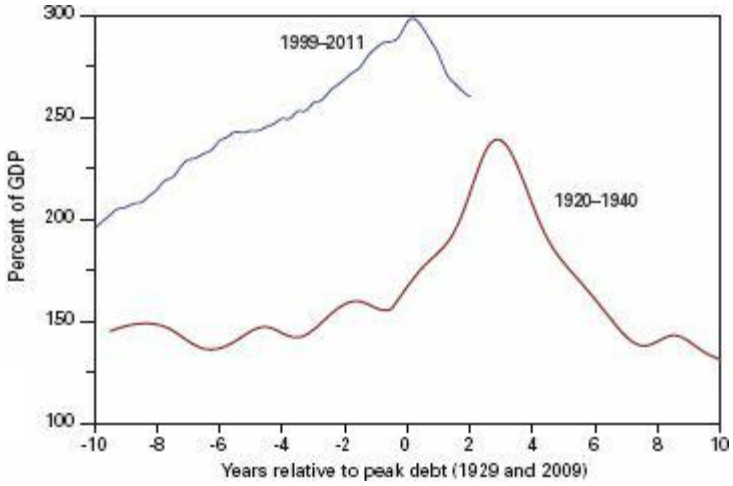
13.23 Nominal private debt then and now

In inflation-adjusted terms, the increase was very similar – a 72 percent increase over the Roaring Twenties versus an 85 percent increase from 1999 to 2009. Remarkably, the real level of debt grew at almost precisely the same rate for the first eight years in both periods – a rate of about 7 percent per year. This chimes with one implication of the monetary model of capitalism I outline in the next chapter: banks increase their profits by increasing debt, and they therefore have an incentive to increase debt as fast as is possible. The easiest way to do this is to fund Ponzi schemes, which were the hallmark of both the Roaring Twenties and the 'Noughty Nineties.'



13.24 Real debt then and now

Though the rate of growth of debt was similar, the level of debt compared to GDP is far higher now than in the 1930s. The debt-to-GDP ratio was 175 percent when the Great Depression began; it is over 100 percent higher today, and hit 298 percent before it began to reverse in 2009. The degree of deleveraging needed to eliminate the Ponzi overhang is therefore much higher today than it was in 1930.

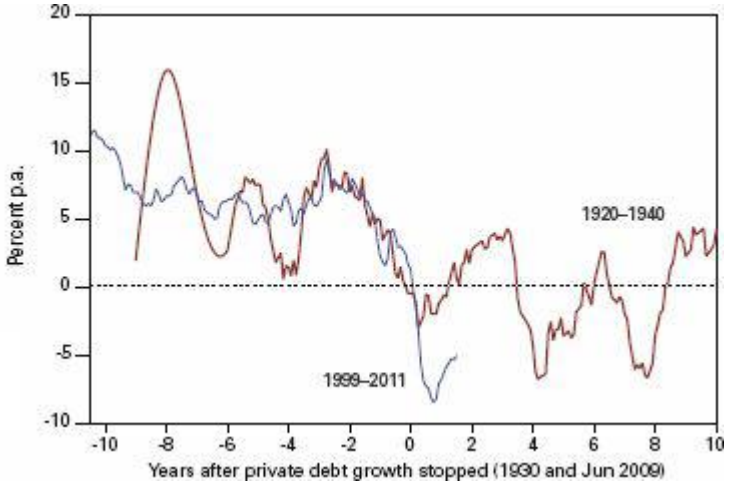


13.25 Debt to GDP then and now

Rising debt fueled the Roaring Twenties, just as rising debt fueled the false prosperity of the Internet and Subprime Bubbles in the 'Noughty Nineties.' Since the rate of real economic growth was higher back in the 1920s than today, the debt ratio itself remained roughly constant prior to the bursting of the Ponzi Scheme in the 1920s; however, debt grew as rapidly in

real terms in the 1920s as it did in the noughties, and the collapse of debt in real terms when the crisis hit was also remarkably similar.

But from there they diverge, because the second scourge of the 1930s – deflation – has yet to occur in a sustained manner during the Great Recession. Consequently, while the real burden of debt rose during the early 1930s even as the nominal level of debt was falling, so far the Great Recession has involved falling debt in both real and nominal terms.



13.26 Real debt growth then and now

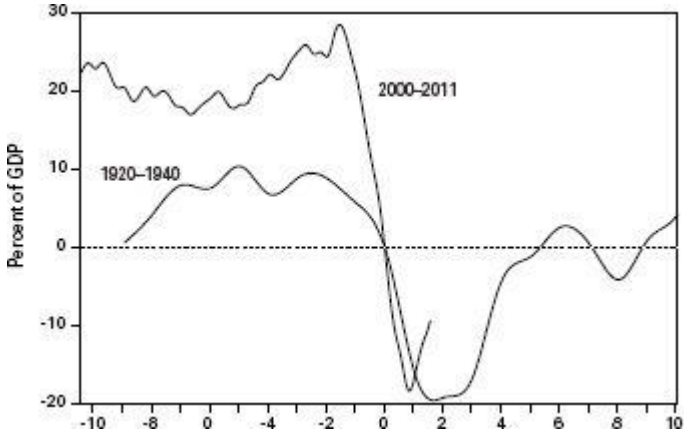
One possible reason for the marked difference in inflationary dynamics between the two periods is the composition of private debt. In the 1920s, the vast bulk of the debt was owed by business. Business debt was three times that of household debt, and four times that of the financial sector. Therefore, when the Roaring Twenties boom collapsed as debt-financing fell, businesses were the ones in serious financial difficulties. As Fisher surmised, individual businesses responded by cutting their markups to try to entice customers into their stores and not their competitors', leading to a general fall in the price level that actually increased the debt-to-GDP ratio, even as nominal debt levels fell.

Today the ranking is reversed in the insolvency stakes: the financial sector carried the highest level of debt leading into the Great Recession – virtually 125 percent of GDP, five times the level of debt it had in 1930. Households come second now, with a debt level of almost 100 percent of GDP, two and a half times the level they had in 1930. The business sector carried a modest debt level of 80 percent of GDP, when compared to its

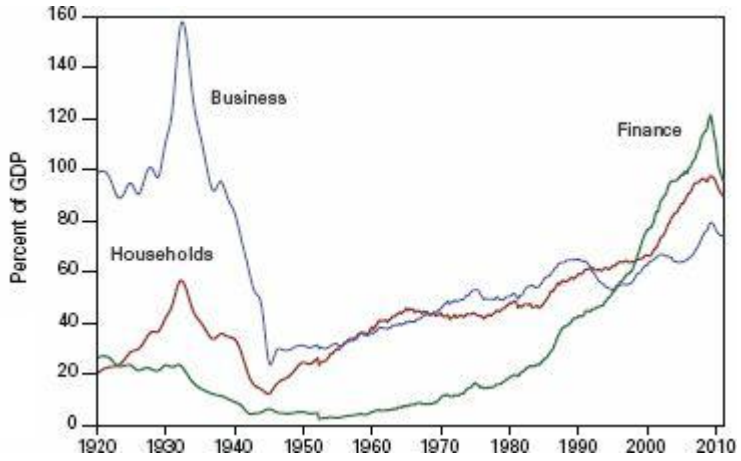
1930s level of 110 percent – though even this is more than twice its debt level during the 'Golden Age' of the 1950s and 1960s.

This composition difference may have implications for how the debt-deflationary dynamics of the Great Recession will play out. The prospects of a 1930s-style deflationary collapse are low, since businesses do not face the direct pressure of insolvency that they faced back then. However, their retail customers, the consumer sector, have never been this debt-encumbered, and it is far harder for households to reduce debt than it is for businesses: to put it colloquially, businesses can get out of debt by going bankrupt, ceasing investment, and sacking the workers. Bankruptcy is far more painful for individuals than companies; it is much harder to stop consuming than it is to stop investing, and households can't 'sack the kids.'

This implies a far less severe tendency to deflation, but a more intractable one at the same time since consumer demand will remain muted while debt levels remain high.



13.27 The collapse of debt-financed demand then and now

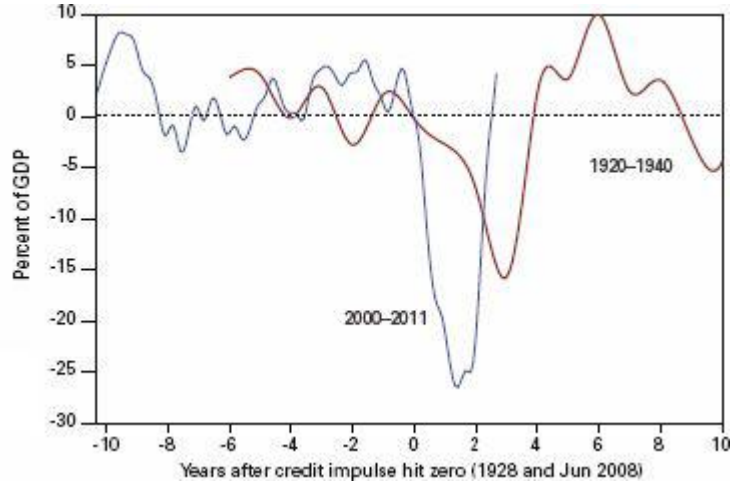


13.28 Debt by sector – business debt then, household debt now

Finally, though the Roaring Twenties became a reference period for frivolous speculation in popular culture, they have nothing on the Noughty Nineties. Debt-financed spending never exceeded 10 percent of GDP in the 1920s. In the noughties, it rarely fell below 20 percent of GDP. The popular culture of the twenty-first century may ignore the Roaring Twenties and see the Noughty Nineties as the hallmark of delusional economic behavior.

Given this much higher level of debt-financed speculation, the plunge into negative territory was far faster in 2008/09 than it was in 1929–31 – but the reversal of direction has also been far more sudden. The change in debt went from adding 28 percent of GDP to aggregate demand in 2008 to subtracting 19 percent from it in 2010, but the rate of decline turned around merely a year after the crisis began, compared to the three years that elapsed before the debt-financed contribution started to rise from the depths in the 1930s (see Figure 13.27) (a large part of this may be the product of the huge intervention by both the federal government and the Federal Reserve).

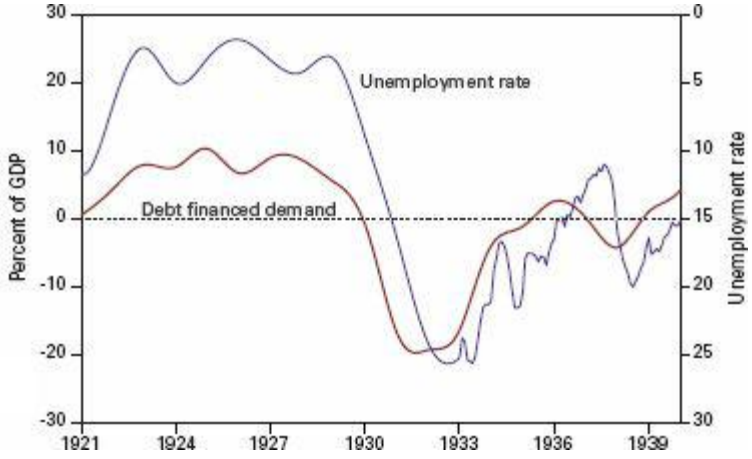
The Credit Impulse was also far more dramatic in the noughties than in the twenties: it was higher during the boom, and plunged far more rapidly and deeply during the slump. The Credit Impulse took four years to go from its positive peak of 2.5 percent before the Great Depression to –16 percent in 1931. It began from the much higher level of 5 percent in late 2007 and fell to a staggering –26 percent in late 2009 – a plunge of over 30 percent in just two years versus an 18 percent fall over four years in the Great Recession.



13.29 The Credit Impulse then and now

The collapse in debt-financed aggregate demand was the key factor behind both the Great Depression and the Great Recession. Though debt-financed demand played less of a role in the 1920s than it did in the noughties, the collapse in the Great Depression was as deep as today's, and far more prolonged, which caused unemployment to hit the unprecedented level of 25 percent in 1932. When the Credit Impulse finally rose again in 1933, so did employment, and unemployment fell to just over 11 percent in mid-1937 – leading to hopes that the depression was finally over.

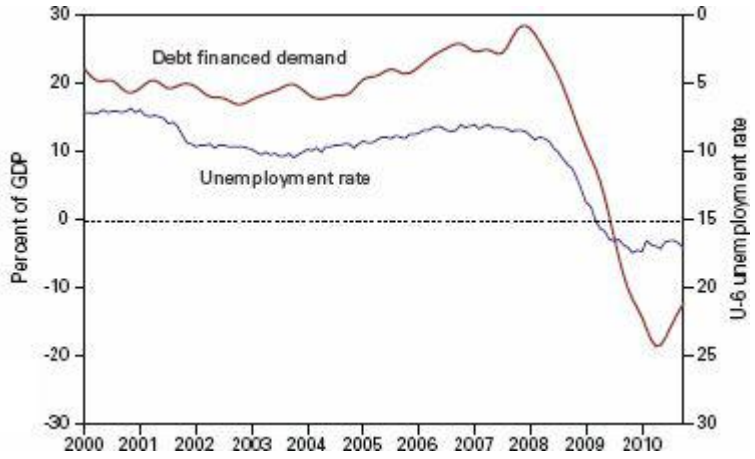
However, debt-financed demand turned negative once again in 1938, and unemployment rose with it to 20 percent. Only with the onset of the war with Japan did unemployment fall back to the average experienced during the 1920s.



13.30 Debt-financed demand and unemployment, 1920–40

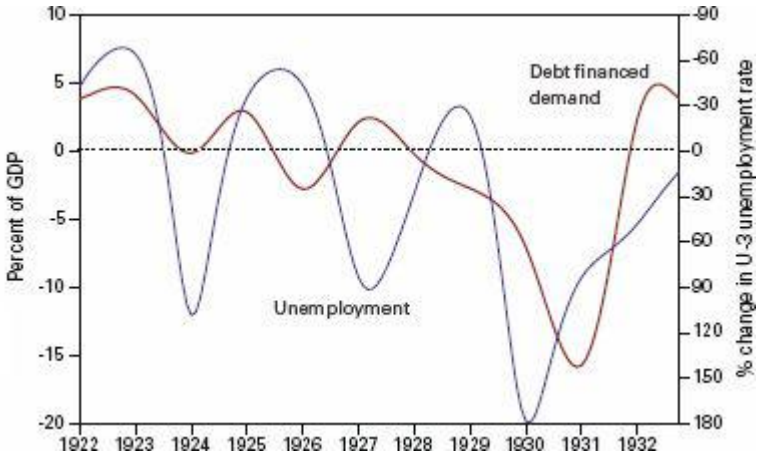
The same pattern has played out during the Great Moderation and Great Recession. When debt-financed demand collapsed, unemployment exploded to 10 percent on the U-3 measure, and 17 percent on the more comparable U-6 measure. Just as significantly, the unemployment rate stabilized when the decline in debt-financed demand turned around. Though the huge fiscal and monetary stimulus packages also played a role, changes in debt-financed demand dominate economic performance.

One statistical indicator of the importance of debt dynamics in causing both the Great Depression and the Great Recession and the booms that preceded them is the correlation coefficient between changes in debt and the level of unemployment. Over the whole period from 1921 till 1940, the correlation coefficient was minus 0.83, while over the period from 1990 till 2011, it was minus 0.91 (versus the maximum value it could have taken of minus one). A correlation of that scale, over time periods of that length, when economic circumstances varied from bust to boom and back again, is staggering.

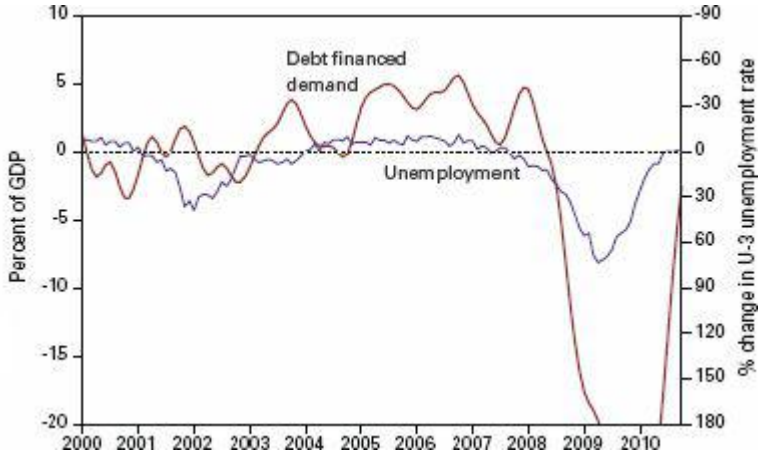


13.31 Debt-financed demand and unemployment, 1990–2011

The Credit Impulse confirms the dominant role of private debt. The correlation between the Credit Impulse and the rate of change of unemployment was minus 0.53 in 1922–40, and minus 0.75 between 1990 and 2011.

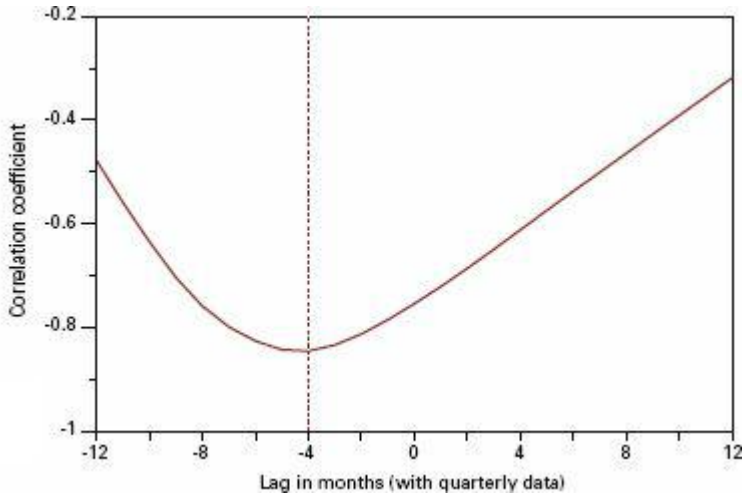


13.32 Credit Impulse and change in unemployment, 1920–40



13.33 Credit Impulse and change in unemployment, 1990–2010

Changes in the rate of change of credit also lead changes in unemployment. When the Credit Impulse is lagged by four months, the correlation rises to minus 0.85.



13.34 The Credit Impulse leads change in unemployment

This correlation is, if anything, even more staggering than that between debt-financed demand and the level of unemployment. The correlation between change in unemployment and the Credit Impulse is one between a rate of change and *the rate of change of a rate of change*. There are so many other factors buffeting the economy in addition to debt that finding any correlation between a first-order and second-order effect is remarkable, let alone one so large, and spanning such different economic circumstances – from the recession of the early 1990s, through the 'Great Moderation,' into the Great Recession and even the apparent beginnings of a recovery from it.

Fighting the Great Recession

The global economy won't return to sustained growth until debt levels are substantially reduced. With debt at its current level, the general tendency of the private sector will be to delever, so that the change in credit will deduct from economic growth rather than contributing to it. Any short-term boost to demand from the Credit Impulse – such as that occurring in early 2011 – will ultimately dissipate, since if it were sustained then ultimately debt levels would have to rise again. Since the household sector in particular is debt-saturated, credit growth will hit a debt ceiling and give way to deleveraging again. The US economy in particular is likely to be trapped in a never-ending sequence of 'double dips,' just as Japan has been for the last two decades.

There is a simple, but confrontational, way to stop this process: a unilateral write-off of debt.

This policy – which occurred regularly in ancient societies, where it was known as a Jubilee (Hudson 2000: 347) – goes strongly against the grain of a modern capitalist society, where paying your debts is seen as a social obligation. But the ancient and biblical practice addressed a weakness in those societies – the tendency for debtors to become hopelessly indebted given the enormous interest rates that were common then:

> Mesopotamian economic thought c. 2000 BC rested on a more realistic mathematical foundation than does today's orthodoxy. At least the Babylonians appear to have recognized that over time the debt overhead came to exceed the ability to pay, culminating in a concentration of property ownership in the hands of creditors. While debts grew exponentially, the economy grew less rapidly. The earning capacity of Babylonian rural producers hardly could be reconciled with creditor claims mounting up at the typical 33.333 percent rate of interest for agricultural loans, or even at the commercial 20 percent rate. Such charges were unsustainable for economies as a whole. (Ibid.: 348)

It would be foolish to deny that we have a similar weakness in modern capitalist society: our tendency to be sucked into Ponzi schemes by a banking sector that profits from rising debt.

As I explain in the next chapter, when lending is undertaken for investment or consumption, debt tends not to get out of hand. But when borrowing is undertaken to speculate on asset prices, debt tends to grow more rapidly than income. This growth causes a false boom while it is happening, but results in a collapse once debt growth terminates – as it has done now.

Though borrowers can be blamed for having euphoric expectations of unsustainable capital gains, in reality the real blame for Ponzi schemes lies with their financiers – the banks and the finance sector in general – rather than the borrowers. That was blindingly obvious during the Subprime Bubble in the USA, where many firms willfully wrote loans when they knew – or should have known – that borrowers could not repay them.

Such loans should not be honored. But that is what we are doing now, by maintaining the debt and expecting that debtors should repay debts that should never have been issued in the first place.

The consequences of our current behavior are twofold. First, the economy will be encumbered by a debt burden that should never have been generated, and will limp along for a decade or more, as has Japan. Secondly, the financial sector will continue to believe that 'the Greenspan Put' will absolve them from the consequences of irresponsible lending.

A debt jubilee would address both those consequences. First, debt repayments that are hobbling consumer spending and industrial investment would be abolished; secondly, this would impose the pain of bankruptcy and capital loss on the financial sector – a pain it has avoided in general thus far through all the rescues since Greenspan's first back in 1987.

Needless to say, this would not be an easy policy to implement.

Its biggest hurdle would be political: it is obvious that the major political force in the USA today – and much of the OECD – is the financial sector itself. Since widespread debt abolition would bankrupt much of this sector, and eliminate individual fortunes (those that have not already been salted away), it will be opposed ferociously by that sector.

The same was the case – though on a smaller scale than today – during the Great Depression. It took a Ferdinand Pecora (Perino 2010) to turn the tide against the bankers then, and a Franklin Roosevelt (Roosevelt 1933) to convert that tide into political power – and policies that included debt moratoria.

The recent Financial Crisis Inquiry Commission (Financial Crisis Inquiry Commission 2011) was a farce compared to Pecora's work, and Obama's administration to date has focused more on returning the financial sector to its old ways than on bringing it to account.

The policy would also need to re-establish the practice of banking providing working capital and investment funds for industrial capitalism. This should be the primary role of banking, but it virtually died out as the financial sector became more and more an engine for speculation, so that most companies today raise their funds on the commercial paper market.[16] A debt jubilee would bankrupt many banks, and put them into receivership;

though it would be painful, the receivers could also be required to re-establish this key but neglected banking practice.

It would also be necessary to compensate to some extent those not in debt as well – though they would also benefit from the sudden increase in spending power that such a policy would cause.

Such a policy would have to be accompanied by institutional reforms to finance that prevented a travesty like the Subprime Bubble from recurring; I discuss some possible reforms at the end of Chapter 14. It would also be far from a panacea for America's woes on its own, since it would also expose the extent to which the gutting of American industry in the last three decades has been disguised by the growth of the financial sector on the back of the Ponzi schemes of the stock and housing markets. The finance sector would shrink dramatically, and unlike in the 1930s, there would not be potential factory jobs awaiting unemployed financial advisors.

A debt jubilee, and the reforms I suggest in Chapter 14, is politically improbable now. But the alternative I believe is a decade or more of economic stagnation. At some stage we are going to have to accept the wisdom in Michael Hudson's simple phrase that 'Debts that can't be repaid, won't be repaid.'

Conclusion

The data on debt confirm the conclusions that can be reached from assessing the logical coherence – or lack of it – in neoclassical theory: every methodological choice neoclassical economics made was wrong. The belief that economics can be reduced to microeconomics is false; money and credit cannot be ignored, capitalism cannot be modeled as a single 'representative agent,' finance destabilizes the economy, and the economy is permanently in disequilibrium.

If we are to develop an economics that is relevant to capitalism, then it must be a strictly monetary, dynamic theory in which finance plays a fundamentally destabilizing role. In the next chapter, I show how such an economic theory can be developed, by building on the work of both the great non-neoclassical economists and recent empirical work by economists from the 'post-Keynesian' school of thought.

14 | A MONETARY MODEL OF CAPITALISM

Many of the foundations on which neoclassical macroeconomics is built arose from persevering with methodological choices that the nineteenth-century founding fathers of neoclassicism made out of expediency rather than preference. They assumed that all economic processes occurred in equilibrium, so that they could model the economy using comparative statics rather than using more difficult dynamic differential equations; they avoided thinking about money and modeled the simpler process of barter instead; they ignored uncertainty about the future and, as Keynes put it, tried to 'deal with the present by abstracting from the fact that we know very little about the future' (Keynes 1937: 215) and so on.

Though these choices made it easy to concoct simple parables about supply and demand, they actually made mathematical modeling of the economy harder, not easier. The absurdities that later neoclassicals added – from the fallacy of the horizontal demand curve to the intellectual travesty of the 'representative agent' – were products of clinging to these simple parables, despite the deep research that contradicted them.

Economists trained on these methods are now scrambling to make ad hoc modifications to the core neoclassical parable to produce hybrid models that mimic the real-world phenomenon of the Great Recession – which, according to the parables, cannot occur.[1] Though such models will superficially ape reality, they will do so for the reasons that Solow gave, that the addition of various 'imperfections' results in a model that 'sounds better and fits the data better' simply because 'these imperfections were chosen by intelligent economists to make the models work better […]' (Solow 2001: 26).

This is the difficult road to relevance – take a theoretical framework in which the real-world phenomenon you are trying to describe cannot happen, and tinker with it until something resembling reality emerges. It will not

last. Once the global economy emerges from this crisis, if this approach still dominates economics, then within decades these 'imperfections' will go the way of the dodo. Economists will return to the core parable, and the crisis we are now in will be seen as the result of bad Federal Reserve policy,[2] rather than a manifestation of capitalism's innate instability – amplified by a finance sector that is almost designed to generate Ponzi schemes.

We have to do better than that. We have to start with foundations from which the phenomena of reality emerge naturally by constructing monetary models of capitalism built on the melded visions of Marx, Schumpeter, Keynes and Minsky.

Methodological precepts

An essential first step towards a meaningful macroeconomics is to acknowledge the one profound lesson from the failure of the neoclassical experiment: that strong reductionism is a fallacy. Macroeconomic phenomena – and even phenomena within one market – are emergent properties of the dynamic, disequilibrium interactions of individuals and social groups in a rich institutional environment, constrained by the physical, temporal and environmental realities of production. These phenomena will not be predictable from the behavior of isolated individuals. Instead, macroeconomics is a self-contained field of analysis, and it must be reconstructed as such. The reductionist route must be abandoned.

There are basically two routes by which models of a new 'emergent phenomena' macroeconomics could be built: the 'bottoms-up' approach that has always dominated economics, but modified in the light of the modern knowledge of complex systems; or the 'tops-down' approach that typified the work of Marx, Schumpeter, Keynes and Minsky, in which the economy is described at the level of aggregates – evolutionary change, social classes, aggregate production, aggregate debt levels and so on.

The former approach takes the macroeconomic phenomena as given, and attempts to build computer-based multi-agent models in which those macroeconomic phenomena arise as emergent properties of the models. The latter works at the level of aggregates, and puts the verbal models of the great non-neoclassical thinkers into the form of dynamic equations.

Most economists who are trying to build macroeconomic models that transcend the neoclassical dead end are taking the former approach (Barr, Tassier et al. 2008; Seppecher 2010).[3] This approach is worthwhile, though there are inherent difficulties in it that I discuss briefly later. I have taken the latter approach of trying to put the Marx-Schumpeter-Keynes-Minsky vision directly into mathematical form.

Doing this turned out to be far easier than I expected, once I made money the starting point of my analysis of capitalism.

Endogenous money

One of the many issues on which Keynes failed to convince his fellow economists was the importance of money in modeling the economy. One reason for this was that money's explicit role in the *General Theory* itself was restricted largely to the impact of expectations about an uncertain future, and the difference between real and nominal wages. Keynes acknowledged that money did not feature heavily in his technical analysis, and that he saw a substantial continuity between monetary analysis and the Marshallian model of supply and demand:

> whilst it is found that money enters into the economic scheme in an essential and peculiar manner, *technical monetary detail falls into the background*. A monetary economy, we shall find, is essentially one in which changing views about the future are capable of influencing the quantity of employment and not merely its direction. But *our method* of analyzing the economic behavior of the present under the influence of changing ideas about the future is one which *depends on the interaction of supply and demand, and is in this way linked up with our fundamental theory of value*. We are thus led to a more general theory, which includes the classical theory with which we are familiar, as a special case. (Keynes 1936: xxii–xxiii; emphases added)

It is therefore difficult to attack neoclassical 'supply and demand'-oriented models of money as misrepresentations of Keynes. Nonetheless, the post-Keynesian school of thought has made the fundamental importance of money a byword of its analysis. An essential aspect of this has been the

empirically based analysis of how money is created (detailed in the previous chapter), which contradicts the conventional fractional reserve banking, 'Money Multiplier' model of money formation.

Having empirically eliminated one model of money creation, another was needed – but the initial attempts to create one were clumsy. Rather than the 'vertical money supply curve' of Hicks's IS-LM model, some post-Keynesian economists proposed a 'horizontal money supply curve' in which banks simply passively supplied whatever quantity of credit money firms wanted, at the prevailing interest rate. This model, known as 'Horizontalism' (Moore 1988b), led to a lengthy dispute within post-Keynesian economics over whether the money supply curve was horizontal, or sloped upwards (Dow 1997).

This dispute put the empirically accurate findings of post-Keynesian researchers into the same methodological straitjacket that neoclassical economics itself employed: the equilibrium analysis of intersecting supply and demand curves. Though this was hardly the intention of the originators of endogenous money analysis, it effectively made monetary analysis an extension of supply and demand analysis.

Participants in this debate were aware of the limitations of this approach – as Sheila Dow observed, '[T]he limitations of a diagrammatic representation of a non-deterministic organic process become very clear. This framework is being offered here as an aid to thought, but it can only cope with one phase of the process, not with the feedbacks' (ibid.: 74). But one of the great ironies of economics is that, because critics of neoclassical economics were themselves trained by neoclassical economists, most critics weren't trained in suitable alternative modeling methods, such as differential equations or multi-agent simulation.

For real analytic progress to be made, a watertight basis for Keynes's assertion that money 'enters into the economic scheme in an essential and peculiar manner' was required, as well as a methodological approach that captured the feedback effects that diagrams and equilibrium analysis could not.

The former was supplied by the 'Monetary Circuit' school in Europe, and specifically the Italian economist Augusto Graziani. Graziani argued that, if money is treated as just another commodity subject to the 'laws' of supply and demand, then the economy is effectively still a barter system: all

that has happened is that one more commodity has been added to the mix, or singled out as the commodity through which all barter must occur. This is quantitative change, not qualitative, and yet something qualitative must change if a monetary economy is to be distinguished from a barter system.

Graziani's brilliant insight was that, for a monetary economy to be clearly distinguished from a barter economy, *the monetary economy could not use a commodity as money*. Therefore money had to be a non-commodity – something that was intrinsically worthless, and which could not be simply produced as commodities themselves can: 'a commodity money is by definition a kind of money that any producer can produce for himself. But an economy using as money a commodity coming out of a regular process of production, cannot be distinguished from a barter economy' (Graziani 1989: 3). This then led to a simple but profound principle: 'A true monetary economy must therefore be using a token money, which is nowadays a paper currency' (ibid.: 3).

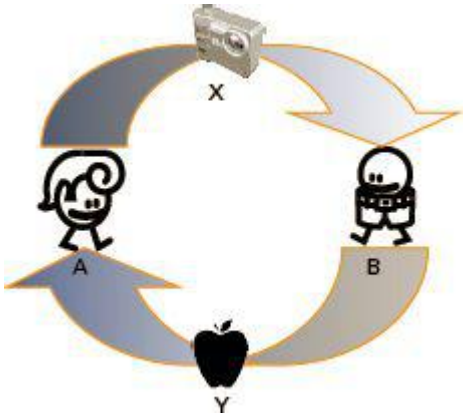
The fact that a monetary economy uses a token – something that is intrinsically worthless – as a means of exchange implies two further key conditions 'In order for money to exist':

b) money has to be accepted as a means of final settlement of the transaction (otherwise it would be credit and not money);

c) money must not grant privileges of seigniorage to any agent making a payment. (Ibid.: 3)

From this Graziani derived the insight that 'any monetary payment must therefore be a triangular transaction, involving at least three agents, the payer, the payee, and the bank':

> The only way to satisfy those three conditions is to have payments made by means of promises of a third agent, the typical third agent being nowadays a bank […] Once the payment is made, no debt and credit relationships are left between the two agents. But one of them is now a creditor of the bank, while the second is a debtor of the same bank. (Ibid.: 3; all emphases in original)

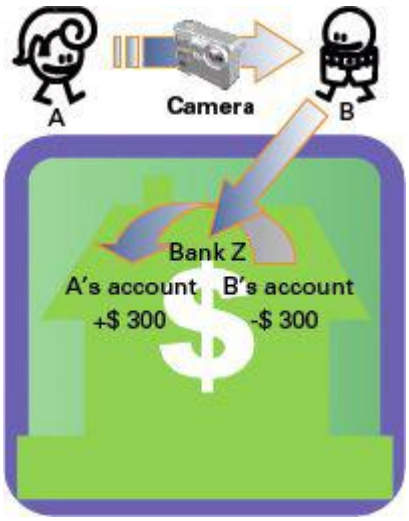


14.1 The neoclassical model of exchange as barter

This perspective clearly delineates a monetary vision of capitalism from the neoclassical barter paradigm. As shown in Figure 14.1, in the neoclassical world, transactions are two-sided, two-commodity, barter exchanges: person *A* gives person *B* one unit of commodity *X* in return for some number of units of commodity *Y*. Calling one of these 'the money commodity' does not alter the essentially barter personality of the transaction.

But in our monetary world, transactions are three-sided, single-commodity, financial exchanges, as portrayed in Figure 14.2: person *B* instructs bank *Z* to debit *Y* units of currency from *B*'s account, and credit *A*'s account with the same amount, in return for which person *A* gives person *B* one unit of commodity *X*.

Banks are thus an essential component of capitalism, and are inherently different to industrial firms. Firms produce goods (and services) for sale by combining labor and other commodities in a production process that takes both time and effort. Banks generate and honor promises to pay that are used by third parties to facilitate the sale of goods.[4] Therefore firms and banks must be clearly distinguished in any model of capitalism: 'Since in a monetary economy money payments necessarily go through a third agent, the third agent being one that specializes in the activity of producing means of payment (in modern times a bank), banks and firms must be considered as two distinct kinds of agents […] In any model of a monetary economy, banks and firms cannot be aggregated into one single sector' (ibid.: 4; emphasis in original).



14.2 The nature of exchange in the real world

This simple but profound perspective on what is the essence of a monetary capitalist economy yielded two essential requirements for a model of capitalism:

- all transactions involve transfer of funds between bank accounts;
- the minimum number of classes[5] in a model of capitalism is three: capitalists, workers and bankers.

It also implied that the best structure for modeling the financial side of capitalism is a double-entry system of bank accounts. This led me to develop a means to derive dynamic monetary models of capitalism from a system of double-entry bookkeeping accounts (Keen 2008, 2009b, 2010, 2011), and a remarkable amount of the Marx-Schumpeter-Keynes-Minsky perspective on capitalism arose naturally out of this approach.

I'll outline the simplest possible version of this model before expanding it to provide a monetary version of the Minsky model outlined in Chapter 13.

A 'pure credit' economy

Our modern monetary economy is a system of such complexity that it makes the outrageous contraptions of Rube Goldberg, Heath Robinson and Bruce Petty appear trite by comparison: the Bank of International Settlements, central banks, commercial banks; merchant banks, hedge

funds, superannuation funds, building societies; fiat money, credit money, multiple measures of money (base money, M_0, M_1, M_2, M_3, broad money); reserve ratios, Taylor Rules, Basel Rules …

Many of these components were instituted to try to control bank lending after the catastrophe of the Great Depression; many others were responses by the financial system to evade the intentions of these controls. To my cynical eye, the evasive maneuvers of the financial system have been far more effective than the regulatory structures themselves, and in essence our financial system approximates the behavior of the almost completely unregulated private banks of the 'free banking' period in the nineteenth century.

14.3 A nineteenth-century private banknote

For that reason, my base monetary model is a pure credit economy with no government or central bank, in which the private bank prints its own paper notes, and where transactions involve transferring paper notes from the accounts of the buyers to that of the sellers. There are three classes – workers, capitalists and bankers – and, in the simplest possible model with no Ponzi lending behavior, firms are the only borrowers, and they borrow in order to be able pay the wages needed to hire workers.

Five accounts are needed to describe the basic monetary flows in this system:

1 a vault, in which the bank stores its notes prior to lending;
2 a 'bank safe,' into and out of which interest payments are made;
3 deposit accounts for firms, into which money lent by the banks is put and through which all the firm sector's transactions occur;

4 deposit accounts for workers, into which their wages are paid; and
a loan register, which is *not* an account as such, but a ledger that records
5 the amounts that have been lent by the banks to firms, and on which loan interest is charged.

The basic monetary operations that occur in this simple model are:[6]

1 the banking sector makes loans to the firm sector;
2 the banks charge interest on outstanding loans;
3 firms pay the interest;
4 firms hire workers;
5 workers and bankers consume the output of the firms; and
6 firms repay their loans.

These operations are shown in Table 14.1, which (based on the standard accounting practice of showing 'assets minus liabilities equals equity') shows the economy from the point of view of the banks, with the banking sector's assets on the left-hand side of the ledger and its liabilities and residual equity on the right-hand side.[7]

Actual transfers of money are shown in normal text, while operations that are not money transfers but accounting operations – such as the bank recording that interest due on loans has been paid – are shown in italics.

TABLE 14.1 A pure credit economy with paper money

	Bank assets		Bank liabilities plus equity			
			Liabilities (deposits)			Equity
Operation	Vault	Loan ledger	Firms	Workers	Safe	
Lend money	– Lend money		+ Lend money			
Record loans		*– Lend money*				
Charge interest		*+ Charge interest*				
Pay interest			– Charge interest			+ Charge interest
Record payment		*– Charge interest*				
Deposit interest			+ Deposit interest			– Deposit interest
Hire workers			– Wages	+ Wages		
Bankers consume			+ Bankers' consumption			– Bankers' consumption
Workers consume			+ Workers' consumption	– Workers' consumption		
Loan repayment	+ Loan repayment		– Loan repayment			
Record repayment		*– Loan repayment*				

Since all the entries in this table indicate flows into and out of accounts (or additions and subtractions from the loan ledger), a remarkable thing is possible: a dynamic model of this monetary model can be derived just by 'adding up' the entries in the columns, as in Table 14.2.

TABLE 14.2 The dynamics of a pure credit economy with no growth

Rate of change of:	Equals:
Vault	– Lend money + Repay loans
Loan ledger	+ Loan money – Repay loans
Firm deposits	+ Loan money – Charge interest + Deposit interest – Wages + Bankers' consumption + Workers' consumption – Repay loans
Worker deposits	+ Wages – Workers' consumption
Safe	+ Charge interest – Deposit interest – Bankers' consumption

This model can be simulated if we put values on these flows. Some of these are obvious: the interest charged, for example, will equal the rate of interest on loans times the amount currently recorded on the loan ledger; interest paid is the rate of interest on deposits times the amount currently in the firms' deposit accounts.[8]

Others – lending from the vault, payment of wages, consumption by workers and bankers and loan repayment – will in the real world depend on a whole host of factors, but to model the simplest possible system, I relate them here to the balances in these other accounts, and use constants rather than variables simply to see whether the model is viable: obviously, if it's impossible to find a set of constants that makes this model viable, then no set of variables is likely to do it either.

Thus lending from the vault is modeled occurring at some constant rate times the amount of money in the vault; the flow of wages is some constant times the balance in firms' deposit accounts; workers' and bankers' consumption depend on the balances in the workers' deposit accounts and the safe respectively; while the flow of loan repayments is some constant times the amount of loans outstanding.

The constants (known as 'time constants' in dynamic modeling)[9] used tell us how many times in a year the given account will turn over – so a value of ½, for example, indicates that the balance in the relevant account will be turned over every two years. One obvious value here is that for workers' consumption: since workers' wages are paid on a weekly basis, and most of workers' incomes is expended on consumption, the constant for workers' consumption will be 26 – indicating that the balance in the workers' accounts turns over twenty-six times a year. For the sake of illustration, I use ½ for lending money (so that the vault turns over every two years), 3 for wages, 1 for bankers' consumption, 26 for workers'

consumption, 1/10 for loan repayment, and I set the rate of interest on loans to 5 percent and the rate of interest on deposits to 2 percent.

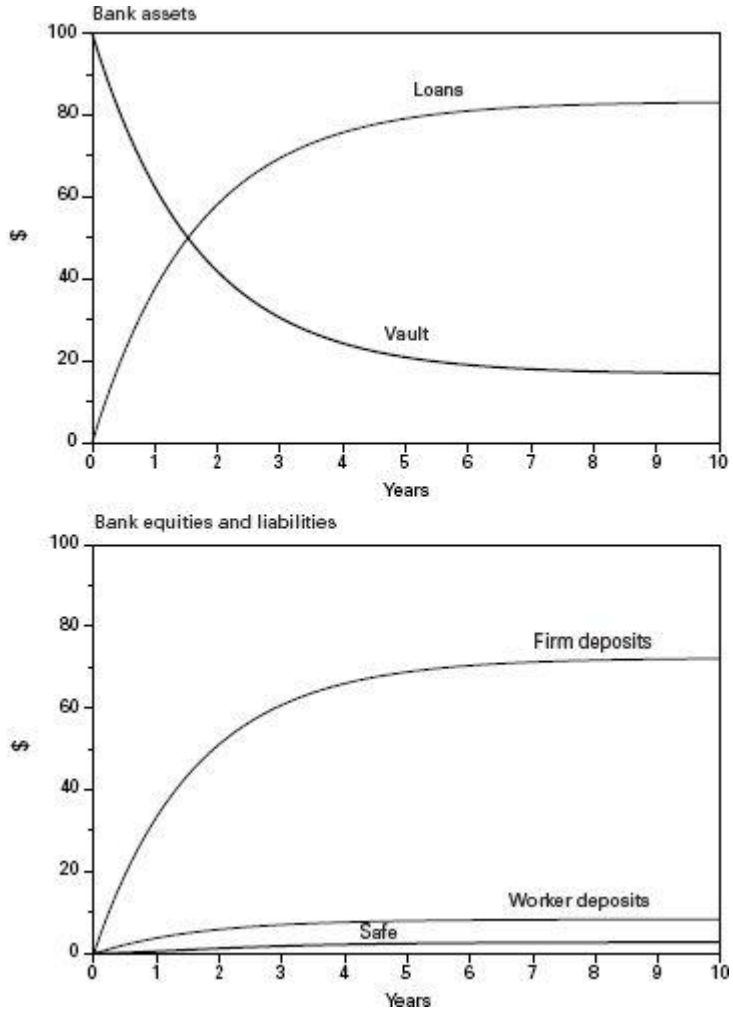
If the model starts with $100 million initially in the vault and no money in any other account, then after ten years, the amount in the vault falls to $16.9 million, with $83.1 million in outstanding loans, $2.7 million in the safe, $72.1 million in firm deposit accounts, and $8.3 million in the workers' deposit accounts – see Figure 14.4.[10] It is also possible to calculate the annual wages bill and bank earnings. The annual wages bill is the time constant for wage payments times the balance in the firms' deposit account, which is three times $72.1 million or $216.3 million, while bank gross earnings are the rate of interest on loans times the outstanding loan balance (5 percent times $83.1 million or $4.16 million) minus the rate of interest on deposits times the firms' deposit balance (2 percent times $72 million or $1.44 million), for a net bankers' income of $2.7 million per annum.

Capitalists' income isn't as obvious in this simple model, and to explain it properly will require incorporating production and pricing as well. But we can imply what profits are by realizing that net annual income in this simple model equals the sum of wages plus profits – the income of bankers cancels out and adds nothing to aggregate income (see Table 14.3).

TABLE 14.3 Net incomes

Class	Net Income components	Amounts
Workers	Wages	216.3
Capitalists	Profits minus Loan interest plus Deposit interest	72.1 – 4.16 + 1.44
Bankers	Loan interest minus Deposit interest	4.16 – 1.44
Total income	Wages plus Profits	288.4

Since wages represent part of the net surplus generated in production, profits must represent the remainder. If workers' wages represent, say, 75 percent of net income, then profits represent 25 percent – so in this numerical example they equal $72.1 million.[11]



14.4 Bank accounts

Annual income in this example is thus $288.4 million – almost three times the amount of money in the model, and precisely four times the amount of money in firms' deposit accounts. How can this be? Marx's insight into why Say's Law is invalid in a capitalist economy holds the key. Remember that Say's Law holds under simple commodity production (*Commodity→Money→Commodity*), but not in capitalism, because that also has the circuit *Money→Commodity→More Money*. Marx also pointed out that this 'Circuit of Capital' takes time: it involves getting money in the first place, using it to hire workers and buy inputs, combine them in a production process, ship the finished goods and finally sell them to customers. There is thus a time lag between outlaying *M* and earning *M+*, which Marx called the 'period of turnover.' This can be significantly shorter

than a year, though it's highly unlikely to be as short as the example Marx himself gave: 'Let the period of turnover be 5 weeks, the working period 4 weeks […] In a year of 50 weeks […] Capital I of £2,000, constantly employed in the working period, is therefore turned over 12½ times. 12½ times 2,000 makes £25,000' (Marx 1885: ch. 16).

Expressed as a fraction of a year, Marx's example gives a value of 1/12.5 for the period of turnover – and in general, the smaller the number, the faster a given amount of money turns over, and the more profit (and wages) that can be generated. Marx's numerical example was extreme, but the basic insight is correct, that a given sum of money can finance several times as much turnover in a given year.

The period of turnover can also be derived for our example, using the facts that the value of the time constant for wages is 3, and 75 percent of national income goes to workers as wages. Total income – wages plus profits – is thus four times the amount of money in the firms' deposit accounts. The turnover period is therefore one year divided by 4: it takes three months, in this toy economy, to go from *M* to *M+*.

Though the turnover period is an unfamiliar concept, it's related to the well-known if less well-defined concept of the velocity of money. The turnover period tells us how often the money in firms' deposit accounts turns over; the velocity of money in this model is the value of wages plus profits (GDP, which is $288.4 million in this example) divided by either the total money supply ($100 million) or the money in active circulation, which is the sum of the amounts in the deposit accounts plus the safe ($83.1 million). Measured the former way, the velocity of money is 2.88; measured the latter way, it's 3.47.

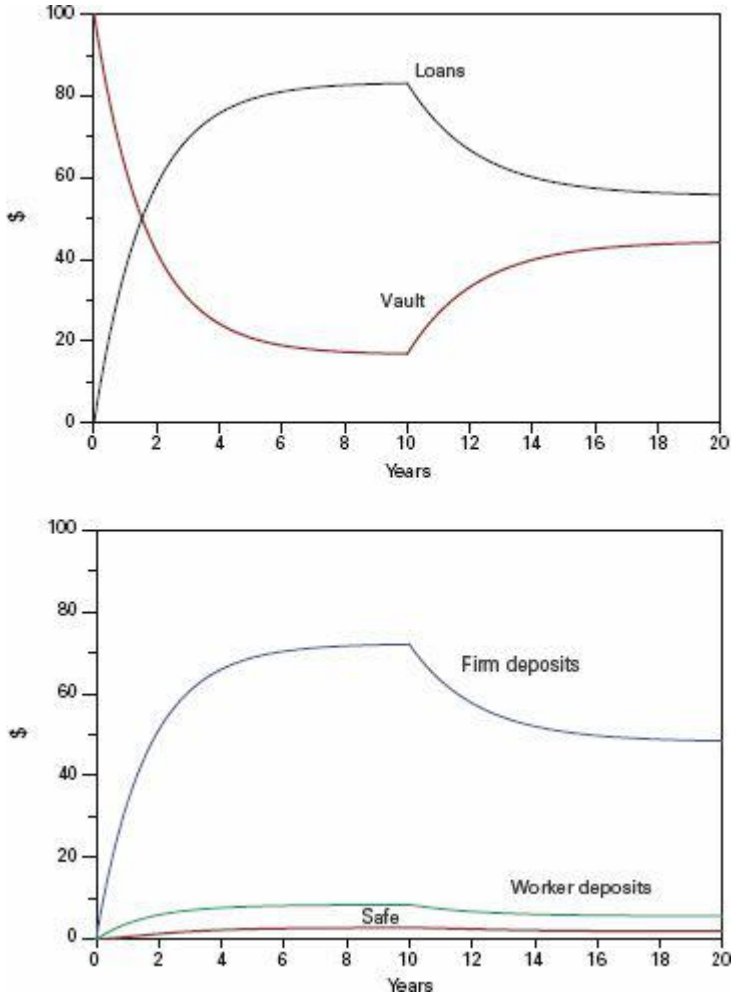
This is an incredibly simple system, but even at this point it can give us some insights into why Bernanke's QE1 was far less effective than he had hoped – and why it would have been far more effective if the money had been given to the debtors rather than to the banks.

A credit crunch

The crisis of 2007 was not merely a credit crunch (where the problem is liquidity) but the end point in the process of Ponzi lending that made much of the US economy insolvent. However, the credit-crunch aspect of this

crisis can be simulated in this model by halving the rate at which the bank lends from the vault, and doubling the speed at which firms try to repay their debts. The time constant for bank lending therefore drops from ½ to ¼ – so that the amount in the vault turns over every four years rather than every two – while that for repaying debts goes from 1/10 to 1/5 – so that loans are repaid every five years rather than every ten.

The credit crunch has a drastic impact upon both bank account balances and incomes. The level of loans drops from over $83 million to under $56 million, while the amount in the vault – and therefore inactive – rises from $16.9 million to $44.1 million.



14.5 A credit crunch causes a fall in deposits and a rise in reserves in the bank's vault

All incomes drop substantially as well: wages drop from $216 million to $145 million per year, profits drop from $72 million to $48.5 million, and

bank income drops from $2.7 million to $1.8 million – a 32.8 percent drop.

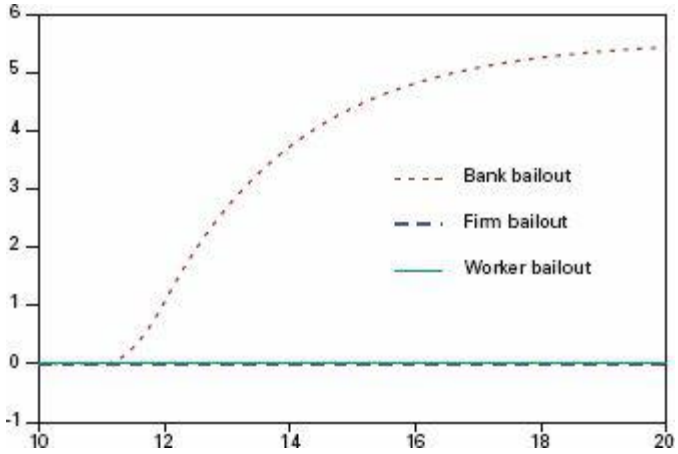
Now let's consider what would happen if an injection of $10 million was made one year after the crunch began, into either the vault, or into the deposit accounts of the firms. The former approximates what Bernanke did in his attempt to exploit the mythical 'Money Multiplier,' the latter approximates what might have happened if the bailout had gone to debtors rather than to the banks – and this is also very similar to what was in fact done in Australia, where the Rudd government effectively gave every Australian with a pulse $1,000 to spend.[12]

The results are intriguing, complex even though the model itself is simple, and the reverse of what Obama was told would happen by his neoclassical advisors.

Whose bailout works best?

The bank bailout injects $10 million into the vault over a one-year period; the firm and worker bailouts inject the same amount of money over the same period of time into the deposit accounts of the firms or workers.

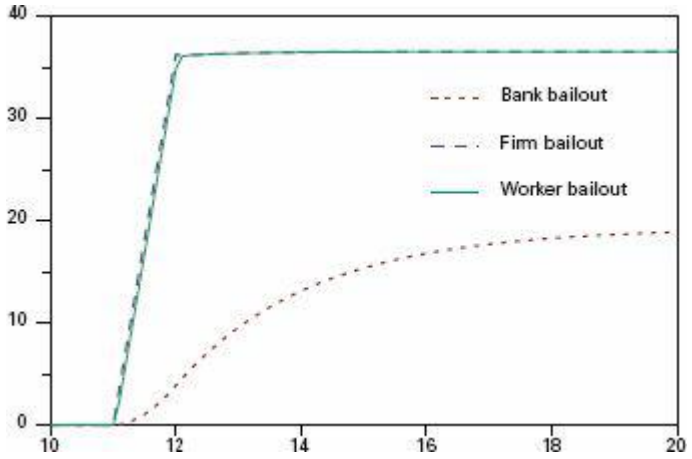
If you believed that the most important thing was to get lending going again after a credit crunch, then the bank bailout wins hands down: neither the firm nor the worker bailouts affect the level of loans at all, which remain on the depressed credit-crunch trajectory, while the bank bailout leads to loans falling less steeply, so that ten years after the crunch, they are $5.5 million higher than they would have been without the bailout.



14.6 A bank bailout's impact on loans

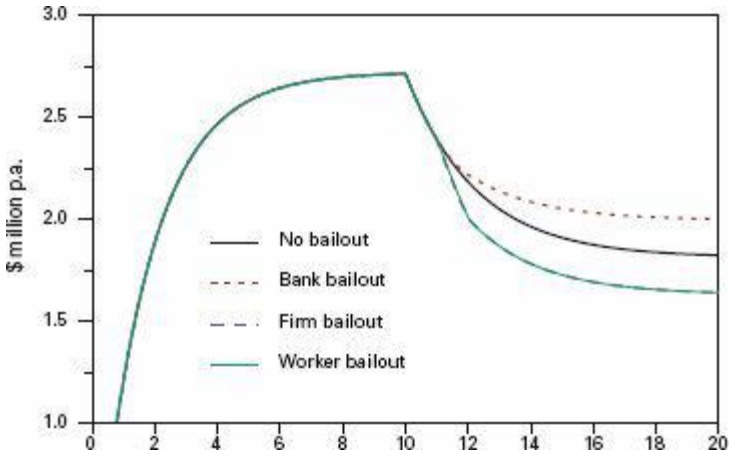
However, if you believed that the most important thing was to restore economic activity, then the bank bailout is the least effective way to do this!

Profits and wages do rise because of the bank bailout, but the rise in income is far greater when the firms or workers receive the bailout than when the banks do.[13] The increase in incomes is immediate and large in the case of the firms' bailout, versus gradual and modest for the bank bailout.



14.7 A bank bailout's impact on incomes

The only people that do better if the bailout goes to the bankers … are the bankers. Not only do they do better under their bailout than if nothing is done, they do worse if the bailout goes to firms or workers than if there is no bailout at all! The reason is that the firm (or worker) bailout increases the deposit accounts of the banks while leaving their loans unaffected. Their payment of interest to the rest of the economy therefore increases, while their receipts of interest payments remain the same.



This is a very basic and incomplete model, and much more needs to be added to it before any definitive implications could be drawn about the impact of a government bailout during a credit crunch.[14] But the differences between this simple dynamic model, and the even simpler but false Money Multiplier model that lay behind Obama's decision to bail out the banks rather than the public, tempt me to write what Obama *could* have said, if his advisers were not neoclassical economists:

> And although the banks have argued that government money would be more effective if it were given to them to lend, rather than going directly to families and businesses– 'where's our bailout?' they ask – the truth is that an additional dollar of capital in a bank will dribble out slowly through the choked arteries of our sclerotic financial system, while that same dollar, if given to families and businesses, will enter circulation rapidly, a process that will cause a faster pace of economic growth.

But that's enough of fantasy. Let's bring this model up to date in terms of how money is created endogenously today, and extend it to include production, prices and growth.

A modern credit crunch

The model we've just considered has a fixed amount of money in it, and since it's a paper-money system, the banks would need to print more notes if they wanted to expand the money supply. However, the majority of banking transactions have always involved the buyer writing a check drawn on an account in a bank, rather than handing over paper notes in return for goods – and today's innovation of electronic transfer banking has taken this one step farther. The fact that these promises by banks to pay are accepted as money in their own right is what makes it possible for banks to expand the money supply simply by creating a new loan. The new loan creates a debt between the borrower and the bank, and it also creates additional spending power.

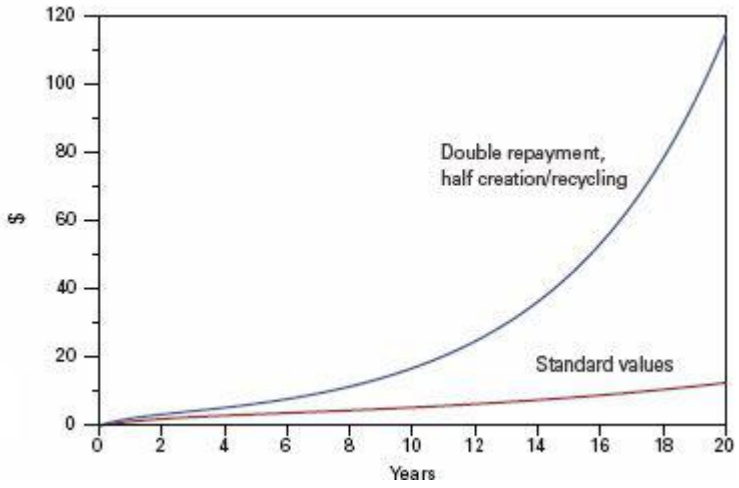
It's this capacity to create money 'out of nothing' which state policies like Reserve Requirements and Basel Rules attempted to control, but the empirical evidence shown in the last chapter shows that these control mechanisms have failed: the banks create as much new money as they can get away with, because, fundamentally, banks profit by creating debt.

TABLE 14.4 A growing pure credit economy with electronic money

	Bank assets		Bank liabilities plus equity			
			Liabilities (deposits)		Equity	
Operation	Vault	Loan ledger	Firms	Workers	Safe	
Lend money	– Lend money		+ Lend money			
Record loans		– Lend money				
Charge interest		+ Charge interest				
Pay interest			– Charge interest		+ Charge interest	
Record payment		– Charge interest				
Deposit interest			+ Deposit interest		– Deposit interest	
Hire workers			– Wages	+ Wages		
Bankers consume			+ Bankers' consumption		– Bankers' consumption	
Workers consume			+ Workers' consumption	– Workers' consumption		
Loan repayment	+ Loan repayment		– Loan repayment			
Record repayment		– Loan repayment				
Lend new money			+ New loan			
Record new loan		+ New loan				

We can model this endogenous creation of both debt and new money (in a check-account or electronic-money banking system) by adding two new rows to the table – one in which the firms' deposit accounts are credited with new money, the second in which the new debt the firms have to the banks is recorded on the loan ledger (see Table 14.4).

This extension helps explain why banks are so willing to create debt, and discourage its repayment: the source of bank profits is interest on outstanding debt, and the more debt that is out there, the more they make. The amount of outstanding debt will rise if existing money is turned over more rapidly, if new money is created more rapidly, and if debts are repaid *more slowly*. Banks therefore have an innate desire to create as much debt as possible – which is why it is unwise to leave the level of debt creation up to the financial sector. As the Great Recession shows, they will be willing to create as much debt as they can, and if they can persuade borrowers to take it on – which is easy to do when banks finance a Ponzi scheme – then the economy will ultimately face a debt crisis where the banks' willingness to lend suddenly evaporates.



14.9 Bank income grows if debt grows more rapidly

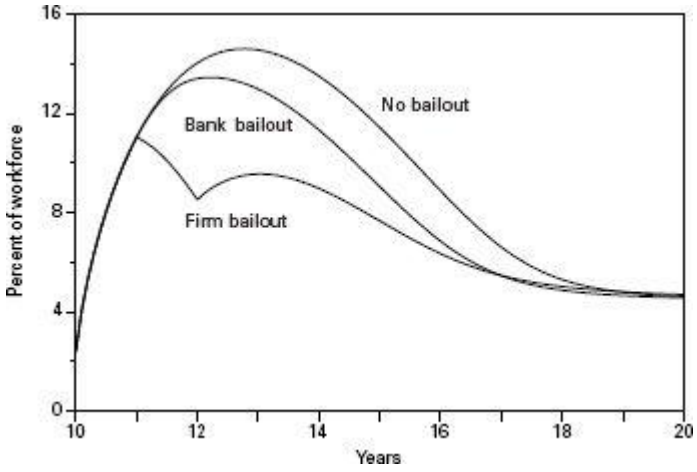
The extension also provides the means to link this purely monetary model to the cyclical Minsky model I outlined in the previous chapter, in a manner that is consistent with the argument that aggregate demand is the sum of income *plus the change in debt*.

In the model above, we were in a 'Say's Law' world in which aggregate demand equaled aggregate supply, and there was no change in debt. However, we now consider firms that wish to invest, and which are willing to take on new debt to finance it – which also causes new money to be created. Aggregate demand is now income plus the change in debt, where incomes finance consumption, and the change in debt finances investment.

The new loans thus provide the money needed to finance the investment that was an integral part of the Minsky model.

For simplicity, I assume that new money is created at a constant rate relative to the current level of debt (which halves when the credit crunch strikes); in the full Minsky model, this is a function of the rate of profit.

To link the two models, one more component is needed: a formula that describes how prices are set. For obvious reasons, this doesn't involve working out where 'marginal cost equals marginal revenue.' However, the equation I use is based on the proposition that prices will tend to converge to a level that equates the monetary value of demand and the monetary value of supply. At the same time, the equation conforms to the empirical research into how firms set prices (see Chapter 4) – that they involve a markup on the wage cost per unit of output – which is the theory of price-setting used by post-Keynesian economists (Lee 1998; Downward 1999).[15]

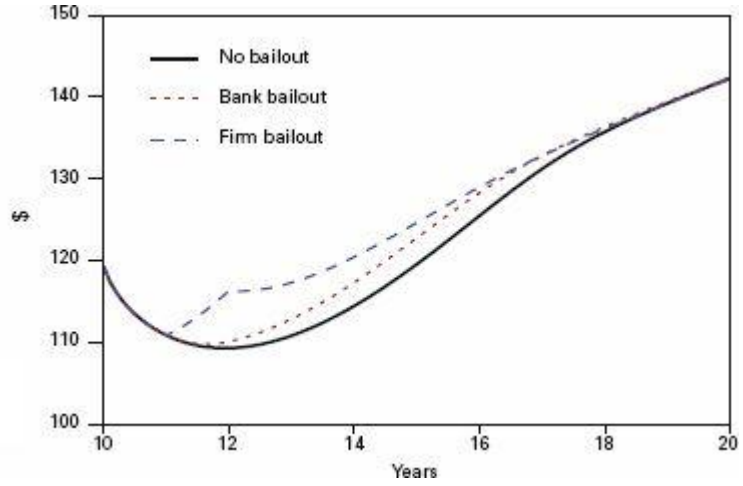


14.10 Unemployment is better with a debtor bailout

We also need an explanation of how wages are set, and this raises the vexed issue of 'the Phillips Curve.' As explained earlier, a properly specified Phillips Curve should have three factors in determining money wages – the employment rate, its rate of change, and a feedback from inflation – but for simplicity here I'll just use the first factor (all three are used later in my monetary Minsky model).

The results of this model amplify the case made in the money-only, no-growth model. The firms' bailout works better on every front, on every metric – except one (any guesses which one?).

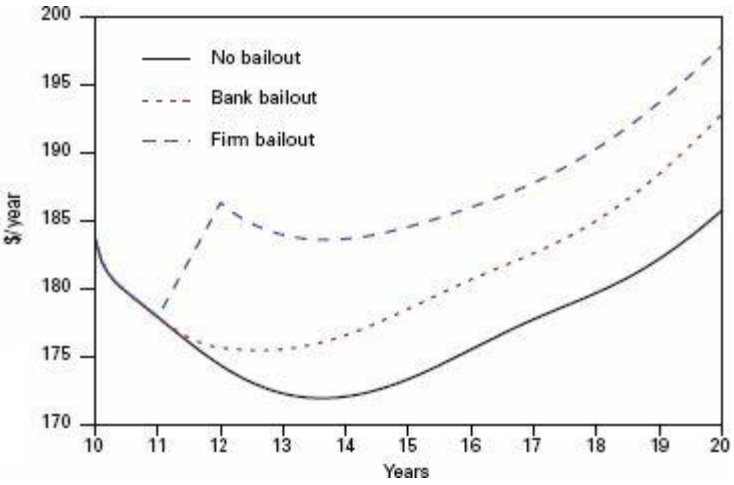
Loans recover more rapidly when the firms are bailed out rather than the banks.



14.11 Loans grow more with a debtor bailout

The rate of unemployment is turned around almost instantly with the firm bailout, and never reaches the extreme levels that apply with the bailout going to the banks (see Figure 14.10).

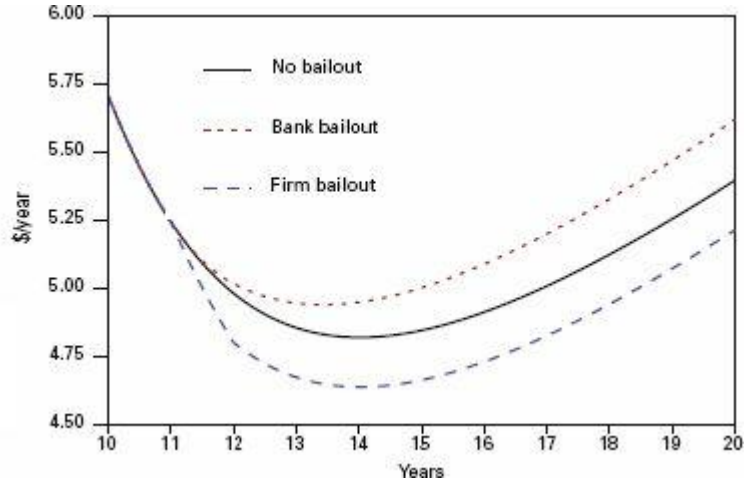
Both profits and wages are higher if the firms get the bailout money rather than the banks.



14.12 Profits do better with a debtor bailout

The only losers from the bailout going to the firms rather than to the banks are … the banks (did you guess right?). Once again, not only do they

do worse if the firms get the bailout rather than them, they do worse under the firms' bailout than they do from no policy intervention at all.



14.13 Bank income does better with a bank bailout

This is still a very simple model, and much more needs to be done to complete it – from replacing time constants with variables (which I do in the Minsky model to come), through to properly modeling government finances as well as those of private banks (which I haven't yet done). But again it reaches results that are the opposite of the neoclassical 'Money Multiplier' model that Obama, acting on the advice of his neoclassical advisors, actually followed. Given the poor response of the economy to the stimulus and QE1, I think it's reasonable to argue that it's time Obama – and politicians in general – looked elsewhere for their economic advice.

From tranquility to breakdown

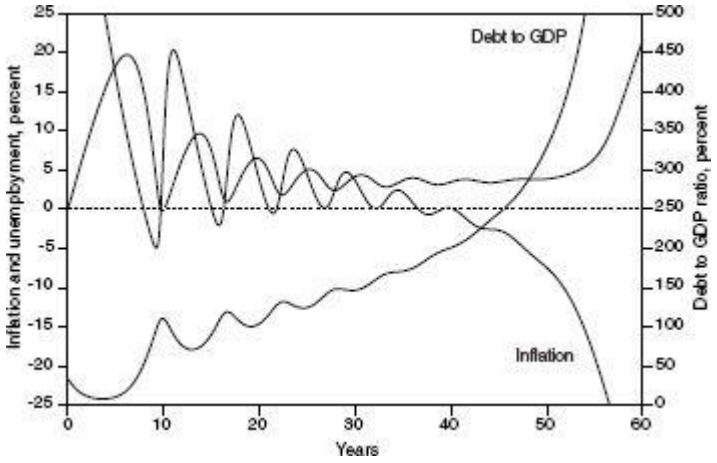
To a neoclassical economist, the most striking aspect of the Great Recession was the speed with which apparent tranquility gave way to sudden breakdown. With notable, noble exceptions like Nouriel Roubini, Robert Shiller, Joe Stiglitz and Paul Krugman, economists paid little attention to the obvious Bonfire of the Vanities taking place in asset markets, so in a sense they didn't see the warning signs, which were obvious to many others, that this would all end in tears.

My model, in contrast, is one in which the Great Moderation and the Great Recession are merely different phases in the same process of debt-financed speculation, which causes a period of initial volatility to give way

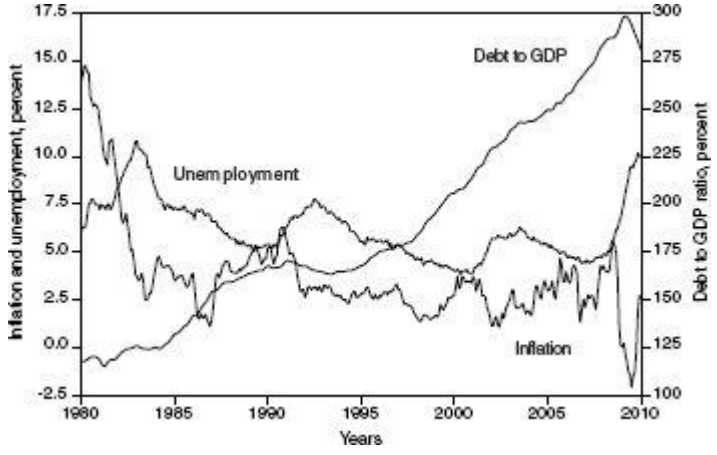
to damped oscillations as rising debt transfers income from workers to bankers, and then total breakdown occurs when debt reaches a level at which capitalists become insolvent.

The fixed parameters used in the previous models are replaced by functions where the rates of money creation and relending and debt repayment depend on the rate of profit, and where the rate of change of wages depends on the level of employment, its rate of change, and the rate of inflation. The link between the monetary and physical models is the creation of new money, which finances investment.

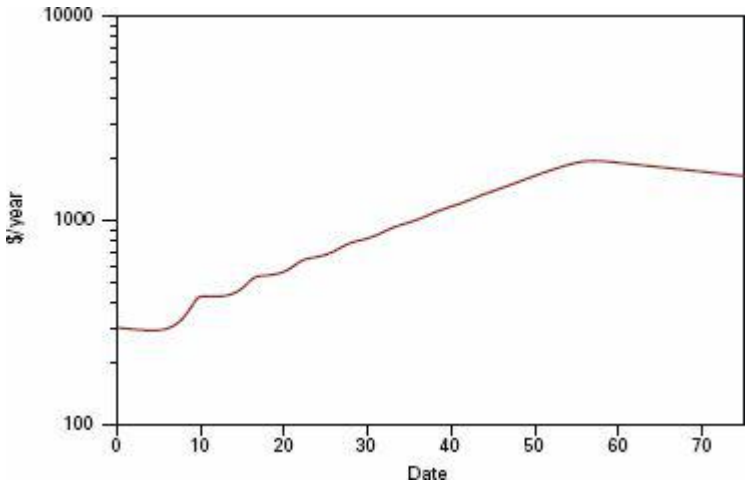
The model generates as sudden a turnaround in output as any neoclassical model hit by 'exogenous shocks,' but unlike in those models there is continuity between the Great Moderation and the Great Recession.



14.14 Modeling the Great Moderation and the Great Recession – inflation, unemployment and debt



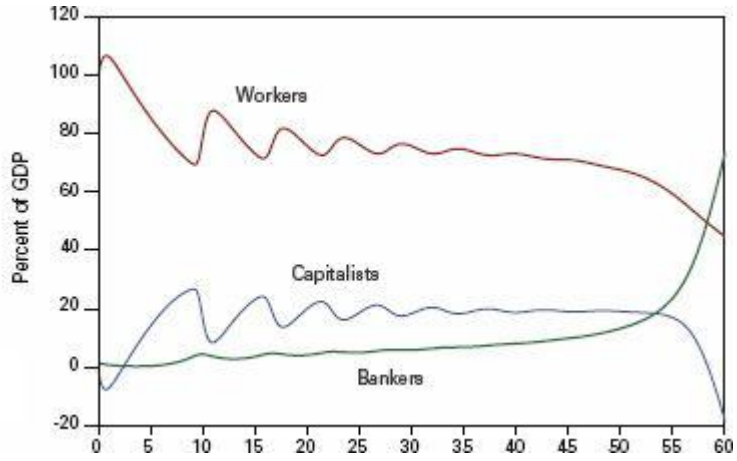
14.15 The Great Moderation and the Great Recession – actual inflation, unemployment and debt



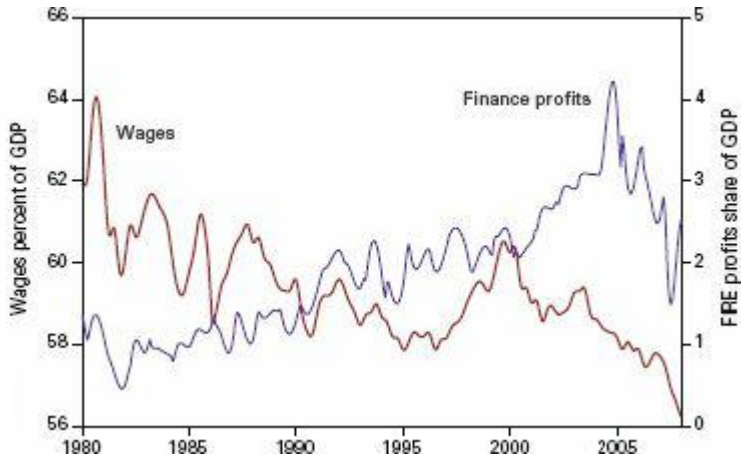
14.16 Modeling the Great Moderation and the Great Recession – output

The model's numbers and the magnitude of its crash are hypothetical,[16] and the main question is whether its qualitative behavior matches that of the US economy – which it clearly does. A period of extreme cycles in unemployment and inflation is followed by diminishing cycles which, if they were the only economic indicators one focused upon, would imply that a 'Great Moderation' was occurring. But the third factor ignored by neoclassical economics – the ratio of debt to GDP – rises in a series of cycles until it takes off exponentially (see Figure 14.14).

The qualitative similarity of this pattern to the actual US data (prior to the massive intervention by both the government and the Federal Reserve) is striking – see Figure 14.15. As in my 1995 model, though capitalists are the ones who actually take on debt, in practice the workers pay for it via a fall in their share of national income.



14.17 Income distribution – workers pay for the debt

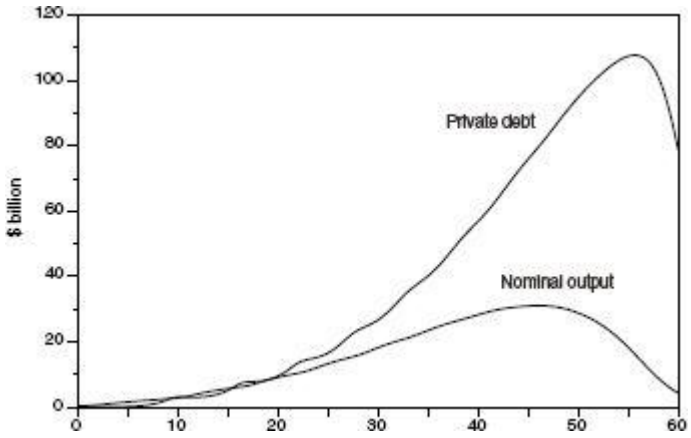


14.18 Actual income distribution matches the model

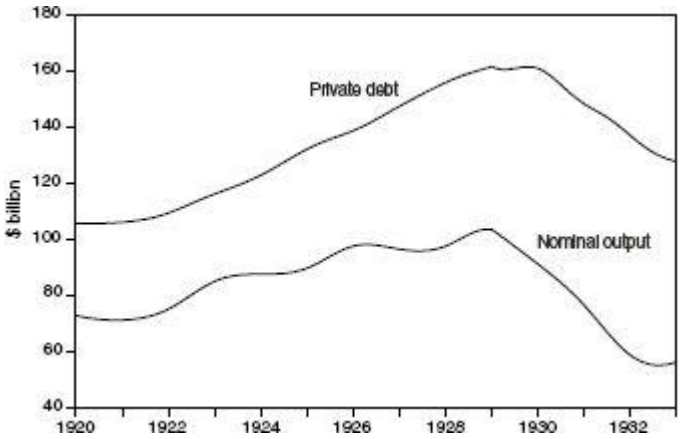
This strictly monetary model generates one aspect of Minsky's hypothesis that my 1995 model could not: the 'deflation' part of the process of debt deflation. Debt rises in a series of booms and busts as in my 1995 paper, but as well the rate of inflation falls in a cyclical manner until it becomes accelerating deflation.

This generates the phenomenon observed in the early years of the Great Depression: the debt-to-GDP ratio continues to rise, even though nominal debt is falling (see Figure 14.19).

The model dynamic is more extreme than the data because the model doesn't yet include the impact of bankruptcy – which reduces debt during a depression. But again, the qualitative similarity between the model and the empirical data is striking – see Figure 14.20.



14.19 Debt and GDP in the model



14.20 Debt and GDP during the Great Depression

Making monetary modeling accessible: QED

I originally developed the models in this chapter using differential equations, and I found it very difficult to extend them, or explain them to other economists who weren't familiar with this approach to mathematics. Then a chance challenge to the accuracy of my models – Scott Fullwiler asserted that there must be errors in my models from the point of view of double-entry bookkeeping – inspired me to see whether I could in fact explain my models using double-entry bookkeeping.

Not only did that prove possible, it also transpired that a double-entry bookkeeping layout of financial flows could be used to generate the models in the first place.

This overcame a major problem that I had with using system dynamics programs like Vissim (www.vissim.com) and Simulink (www.mathworks.com/products/simulink/) to build models of the financial sector. While these technologies were brilliant for designing engineering products like cars, computers and airplanes, they were poorly suited to modeling financial flows.

These programs use 'wires' to link one variable to another, and this is fine for physical processes where, for example, a wire from the fuel injector module to the cylinder module indicates a flow of gas from one point to another, and only one such link exists per cylinder. However, in a model of financial flows, the same term could turn up as often as three times in one diagram: once for the source account for some monetary transfer, once for its destination, and once to record it on a ledger. This resulted in almost incomprehensible models, and made 'wiring up' such a model extremely tedious.

I now use my double-entry bookkeeping methodology to develop models like the one in this chapter, and a simulation tool has also been developed for me to showcase this method. It's free, fairly easy to use, and you can both simulate the models I've shown in this chapter and build your own using it.

It's called QED – which stands for Quesnay Economic Dynamics – and can be downloaded from my blog at www.debtdeflation.com/blogs/qed/.

Conclusion

There are many aspects of this model of which I am critical. For example, it doesn't distinguish borrowing for investment from borrowing for speculation, the government sector isn't incorporated, and many factors that are variable in reality (such as interest rates and the markup that sets prices) are constants in the model. But these missing aspects can be easily introduced into later extensions of the model – a topic that I will take up in my next book, *Finance and Economic Breakdown* – without needing to make the absurd assumptions that neoclassical economics makes when it tries to combine more realism with the fantasy that everything happens in equilibrium.

It is also possible – indeed it is essential – to make this theory one not merely of macroeconomics, but of finance as well. In counterpoint to the false neoclassical dichotomy between macroeconomics and finance on the basis of the counterfactual proposition that debt has no macroeconomic effects, a valid economic theory has to explain the behavior of both the macroeconomy and the financial markets. Such a coherent theory has not yet been developed. However, there are several realistic models of the behavior of financial markets themselves, which we'll now consider.

15 | WHY STOCK MARKETS CRASH

The Efficient Markets Hypothesis says that the stock market's volatility is due to the random arrival of new information that affects the equilibrium value of shares. Allegedly, if it were not for the arrival of new information from outside the market, the market itself would be quiescent.

However, there are alternative explanations that attribute most (though not all) of the market's volatility to its own internal dynamics. Remarkably, these two explanations can predict statistical outcomes for share market prices that are almost indistinguishable from each other.

The kernel

If financial markets aren't efficient, then what are they? According to Behavioral Finance, they are markets where agents make systematically irrational choices, thus resulting in both inefficiency and trading opportunities for the more rational. According to the Fractal Markets Hypothesis, they are highly unstable dynamic systems that generate stock prices which appear random, but behind which lie deterministic patterns. According to the Inefficient Markets Hypothesis, they are systems which overreact to good news and bad, leading to excessive asset price volatility which inhibits the performance of the real economy. According to the burgeoning field of Econophysics, they are akin to nuclear reactors or tectonic plates, where interdependent interactions between speculators can occasionally give rise to runaway processes like nuclear reactions or earthquakes.

All these non-neoclassical theories support the argument that unless finance markets are institutionally tamed, capitalism will remain subject to potentially catastrophic breakdown caused by the finance sector.

The roadmap

In this chapter I outline four different but consistent non-equilibrium theories of finance – 'Behavioral Finance,' the 'Fractal Markets Hypothesis,' the 'Inefficient Markets Hypothesis,' and 'Econophysics.' The chapter concludes with two proposals to institutionally limit the capacity of the finance sector to entice us into debt.

Behavioral finance

Given the failure of the Efficient Markets Hypothesis (EMH), which is predicated on the belief that investors are 'rational' as neoclassical economists define the word, it is little wonder that the most popular response to the failure of the EMH has been to argue instead that investors are in fact irrational – or rather that their behavior deviates from pure rationality in systematic ways. This is then used as part of the explanation as to why the stock market is not efficient – as the Efficient Markets Hypothesis defined the word – so that asset prices deviate from their fundamental values in systematic ways.

As you can imagine, I have rather more sympathy for this approach – which is known as Behavioral Finance – than I do for the EMH. But there are several aspects of this approach that make me rather less enthusiastic than you might expect. I'll detail these before I move on to the legitimate contributions that Behavioral Finance has made to understanding the behavior of finance markets.

What is rational? The development of Behavioral Finance was motivated by the results of experiments in which people were presented with gambles where their decisions consistently violated the accepted definition of rational behavior under conditions of risk, which is known as 'expected utility theory.' Under this theory, a rational person is expected to choose an option that maximizes their expected return – and expected return is simply the sum of the returns for each outcome, multiplied by the odds of that outcome actually happening.

For example, say you were asked whether you'd be willing to take the following 'heads or tails' bet:

Heads: You win $150

Tails: You lose $100

Most people say 'no thanks!' to that gamble – and according to expected utility theory, they're being irrational. Why? Because the 'expected value' of that gamble is greater than zero: a 50 percent chance of $150 is worth $75, while a 50 percent chance of minus $100 is worth minus $50. The sum is plus $25, so that a person who turns the gamble down is walking away from a positive expected value.

Do you think it's irrational to turn that gamble down? I hope not! There's at least one good reason to quite sensibly decline it.[1]

This is that, if you take it, you don't get the 'expected value': you get *either* $150 *or* minus $100. Though you can know the odds of a particular random event like a coin toss, those odds are *almost* irrelevant to any given outcome.[2] Whether the coin will come down heads or tails in any given throw is an *uncertain* event, not a risky one. The measurement of risk is meaningful only when the gamble is repeated multiple times.

This is easily illustrated by modifying the bet above so that if you chose it, you have to play it 100 times. Think carefully now: would you still turn it down?

I hope not, because the odds are extremely good that out of 100 coin tosses, you'll get more than 40 heads, and 40 is the breakeven point. There is only a 1 percent chance that you'd get fewer than 40 heads and therefore lose money. If you get the most common outcome of 50 heads (which occurs 8 percent of the time), you'll make $2,500, while your odds of making between zero (from 40 heads) and $5,000 (from 60 heads) are better than 19 out of 20.

In other words, you get the expected value if, and only if, you repeat the gamble numerous times. But the expected value is irrelevant to the outcome of any individual coin toss.

The concept of expected value is thus not a good arbiter for rational behavior in the way it is normally presented in Behavioral Economics and Finance experiments – why, then, is it used?

If you've read this far into this book, you won't be surprised to learn that it's because economists have misread the foundation research on this topic by the mathematician John von Neumann, and his economist

collaborator Oskar Morgenstern, *The Theory of Games and Economic Behavior* (Von Neumann and Morgenstern 1953).

Misunderstanding von Neumann John von Neumann was one of the greatest intellects of all time, a child prodigy who went on to make numerous pivotal contributions to a vast range of fields in mathematics, physics, and computer science. He was a polymath at a time when it was far more difficult to make contributions across a range of fields than it had been in earlier centuries. One of the fields he dabbled in was economics.

His collaboration with Oskar Morgenstern resulted in whole fields of economic theory being developed by later researchers – including Game Theory, much of neoclassical finance theory, and ultimately Behavioral Economics – but one key thing he actually wanted to achieve never happened: *he wanted to eliminate indifference curves and immeasurable utility from economics*. He regarded these concepts as a sign of the immaturity of economic theory – primarily because it was so lacking in sound empirical data. His observations on this front are sadly even more relevant today:

> In some branches of economics the most fruitful work may be that of careful, patient description; indeed, this may be by far the largest domain for the present and for some time to come […] the empirical background of economic science is definitely inadequate. Our knowledge of the relevant facts of economics is incomparably smaller than that commanded in physics at the time when the mathematization of that subject was achieved. Indeed, the decisive break which came in physics in the seventeenth century, specifically in the field of mechanics, was only possible because of previous developments in astronomy. It was backed by several millennia of systematic, scientific, astronomical observation, culminating in an observer of unparalleled caliber, Tycho de Brahe. Nothing of this sort has occurred in economic science. It would have been absurd in physics to expect Kepler and Newton without Tycho – and there is no reason to hope for an easier development in economics. (Ibid.: 2, 4)

Von Neumann was particularly disparaging about the role that the concept of immeasurable utility took in economic theory. You'll remember from Chapter 1 that early economists imagined that there was a measurable unit of utility they called the 'util,' but that this idea of measurable or 'cardinal' utility gave way to the concept of 'ordinal' utility – in which the satisfaction gained from different bundles of commodities could be ranked, but not measured – because measurement of individual subjective utility was deemed impossible.

Von Neumann disagreed, and proved that in situations in which it was possible to define indifference curves, it was also possible to calculate numerical values for utility by using gambles.

His idea was to set an arbitrary starting point for utility – for example, to define that, for a given individual, one banana was worth one 'util' – and then present that individual with a gamble where the options were either one banana, or a gamble between zero bananas and two bananas with a variable probability. The probability at which the consumer is willing to accept the gamble then lets you derive a numerical estimate of the utility of two bananas. As von Neumann and Morgenstern put it:

> The above technique permits a direct determination of the ratio q of the utility of possessing 1 unit of a certain good to the utility of possessing 2 units of the same good. The individual must be given the choice of obtaining 1 unit with certainty or of playing the chance to get two units with the probability a or nothing with the probability 1–a …; if he cannot state a preference then a=q. (Ibid.: 18–19, n. 3)

For example, if you were willing to accept a gamble that gave you either 2 bananas or zero when the odds of getting 2 bananas was 6 out of 10, then the ratio of the utility of 1 banana to the utility of 2 bananas for this consumer was 0.6. A bit of algebraic manipulation shows that this consumer gets 1.67 utils of utility from consuming two bananas, compared to 1 util from one banana. A hypothetical example of using this procedure to provide a numerical measure of utility is shown in Table 15.1.

TABLE 15.1 Von Neumann's procedure for working out a numerical value for utility Consumer: Joan Cheng

Number of bananas				
Certain	Gamble	Accepted odds (%)	Utility	Marginal utility
0			0.00	
1	0 or 2	60	1.00	1.00
2	0 or 3	78	1.67	0.67
3	0 or 4	92	2.14	0.47
4	0 or 5	97	2.32	0.19
5	0 or 6	99	2.39	0.07

An essential element of this procedure was that it had to be repeatable, and for obvious reasons. If it were done just once, and the experimental subject was hungry, then he might be unwilling to take the risk of starving that the gamble implied, if the outcome were that he had to forgo the banana he already had.

Von Neumann was emphatic about this: to make sense, his procedure had to be applied to *repeatable experiments only*:

> Probability has often been visualized as a subjective concept more or less in the nature of an estimation. Since we propose to use it in constructing an individual, numerical estimation of utility, the above view of probability would not serve our purpose. The simplest procedure is, therefore, to *insist upon the alternative, perfectly well founded interpretation of probability as frequency in long runs*. (Ibid: 19; emphasis added)

Unfortunately, both neoclassical and behavioral economists ignored this caveat, and applied the axioms that von Neumann and Morgenstern developed to situations of one-off gambles, in which the *objective risk* that would apply in a repeated experiment was replaced by the *subjective uncertainty* of a single outcome. Neoclassical economists combined the concept of expected utility with their ordinal, 'indifference curve' theory of consumer choice to develop the Capital Assets Pricing Model, despite the fact that von Neumann was adamant that he wanted to replace the concept of indifference curves with his concept of cardinal utility:

> we hope we have shown that the treatment by indifference curves implies either too much or too little: if the preferences of the individual are not at all comparable, then the indifference curves do not exist. If the individual's preferences are all comparable, then we can even obtain a (uniquely defined) numerical utility which renders the indifference curves superfluous. (Ibid.: 19–20)

Behavioral economists, on the other hand, developed all sorts of 'paradoxes of irrational behavior' from how people's behavior in experiments violated von Neumann's 'Axioms of Expected Utility' – but all of these paradoxes evaporate when the correct, objective, 'frequency in long runs' version of probability is used.

The four axioms were Completeness, Transitivity, Independence and Continuity:[3]

Completeness: A subject can always decide whether he prefers one combination to another, or is indifferent between them.

Transitivity: Choices are consistent so that if shopping trolley A is preferred to trolley B, and B to C, then A is preferred to C.

Independence: Adding two gambles together doesn't change the rankings that apply when the gambles are undertaken separately. And

Continuity: If A is preferred to B and B to C, then there must be some combination of the best (A) and worst (C) option that is as desirable as the middle option (B).

One alleged instance of a violation of these axioms is the famous 'Allais Paradox,' named after the French economist Maurice Allais. The violations definitely occur when a single experiment is all that is conducted, but would disappear if the experiment were repeated multiple times, as von Neumann intended.

Allais compared two experiments, the first of which is shown in Table 15.2:

TABLE 15.2 The Allais 'Paradox': Experiment 1

Option 1A		Option 1B	
Winnings	Odds	Winnings	Odds
$1 million	100%	$1 million	89%
		Nothing	1%
		$5 million	10%

The expected value of Option 1B is higher than that of 1A: 1B is worth $1.39 million (0.89 times $1 million plus 0.1 times $5 million, or $890,000 plus $500,000), so according to expected utility theory, a rational person should choose option A over option B. But in practice, most people choose A – presumably because people prefer a sure thing of a million dollars against even the slightest chance of walking away with nothing.

Rather than calling this behavior irrational, behavioral economists say that this shows 'risk-averse' behavior.

The second experiment is shown in Table 15.3:

TABLE 15.3 The Allais 'Paradox' Part 2: Experiment 2

Option 2A		Option 2B	
Winnings	Odds	Winnings	Odds
Nothing	89%	Nothing	90%
$1 million	11%	$5 million	10%

Here the expected value of Option B is higher than that of A: B is worth $500,000 whereas A is worth $110,000. And here, most people in fact choose option B rather than option A. So in this experiment, most people are consistent with expected utility theory, whereas in the first experiment, most people are inconsistent.

Much was then made of this alleged inconsistency. It was said that it displayed people switching from risk-averse to risk-seeking behavior, that it was provably inconsistent with the Independence Axiom, and so on – the Wikipedia entry on the Allais Paradox gives quite a reasonable summary.

However, these 'inconsistencies' disappear when one uses the 'frequency in long runs' approach that von Neumann *insisted* upon – see his words above. Imagine now that you are offered the chance of repeating Experiment 1 a thousand times. The person who picked option A would certainly walk away a billionaire, but anyone who chooses B will probably

walk away about $400 million richer. Ditto with Experiment 2: Option A would see you probably end up with $100 million, while your wealth via option B would be of the order of half a billion. Only Option B makes any sense in both experiments now – it would clearly be a sign of poor reasoning to choose A instead.

The 'Allais Paradox' is thus not a paradox at all, but a typical case of economists misreading their own literature. I have a similar attitude to all other 'paradoxes' in the behavioral economics literature.

However, this doesn't mean that this entire literature is a waste of time, because the exercises do point out the difference between an uncertain outcome and a risky one – and it is clearly the uncertain outcome which is relevant to people's behavior in stock markets. Uncertainty introduces an asymmetry into people's reactions to losses and gains, and this results in a multitude of ways in which people's behavior deviates from the predictions of the Efficient Markets Hypothesis – which, in their own peculiar way, are similar to the predictions of this misreading of von Neumann.

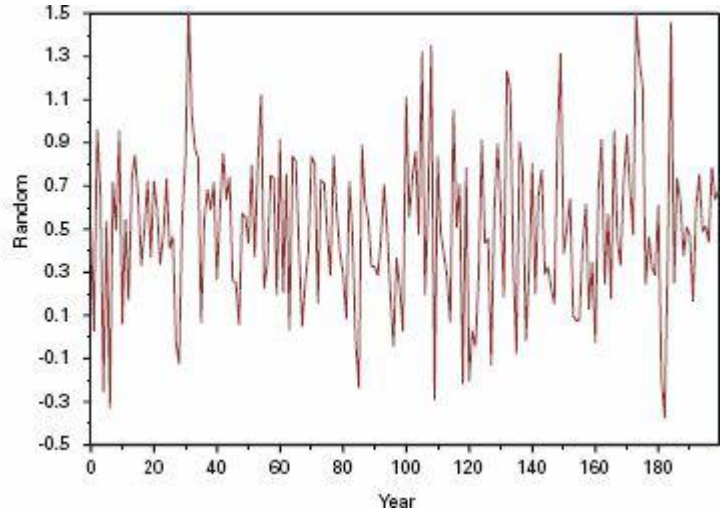
Many of these behaviors are also clearly counterproductive in the context of stock market gambling, and in turn they make it highly likely that market prices will deviate substantially from 'innate value.' These effects also form part of the Inefficient Markets Hypothesis, so I'll delay discussion of them until then.

The inherent instability of stock markets

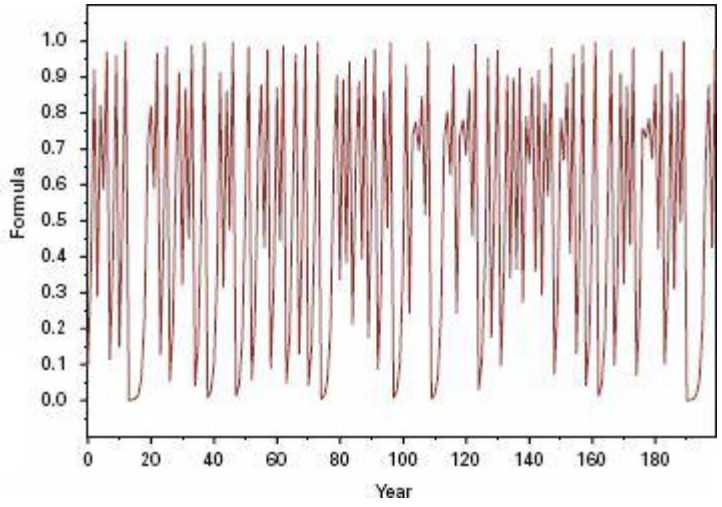
The Efficient Markets Hypothesis explains the price fluctuations that characterize financial markets as rational reactions by the markets to the random arrival of new information affecting the future prospects of companies. The three different approaches to finance outlined in this chapter all argue that these price fluctuations are due to the markets' own internal dynamics. These are two fundamentally different explanations for the same phenomenon: one based on exogenous shocks – the random arrival of external economic news – the other on internal dynamics – today's market prices being a reaction to yesterday's. How can two such different explanations account for the same data?

An analogy might help here. Some animal populations – for example, lemmings – are known to fluctuate wildly from year to year. There could be

two explanations: the environment in which lemmings live could be so volatile that it causes extreme variations in population from one year to the next, and without this environmental volatility, lemming numbers could be constant. Or, the environment could be relatively stable, but the population dynamics of lemmings could be so volatile that they cause huge fluctuations in numbers from year to year.



15.1 Lemming population as a constant subject to exogenous shocks



15.2 Lemming population as a variable with unstable dynamics

It turns out that it's very difficult to know which process is generating a given set, just from the numbers themselves: an unstable dynamic process can generate numbers which are very difficult to distinguish from a set of

random numbers – unless you have a very large data set. The Efficient Markets Hypothesis claimed that the movements in stock prices would be random, and at least initially this contention did seem to be supported by the data from a small sample (between 1950 and 1966). But stock market data actually support a far different contention: that the stock market is inherently unstable.

The Efficient Markets Hypothesis was also developed before the scientific world became reacquainted with the concept of chaos,[4] and it fitted neatly with the economic predilection to see everything in terms of equilibrium. It also meant that economists working in finance theory could avail themselves of all the mathematical and statistical tools devised by mathematicians and scientists to study random processes.

This was an intellectual bonanza – though it simultaneously meant that stock market speculators had to be told that, sadly, there was no bonanza for them hidden in the daily data of the stock exchange. Technical analysts, those looking for trends and waves within waves, were wasting their time.

However, as time went on, more and more data turned up which were not consistent with the EMH. As I detail in the next section, this led to something of a 'siege mentality' by supporters of the EMH, as they fought to defend their theory from attack. But it also inspired other researchers to develop alternative theories of stock market movements.

The Fractal Markets Hypothesis

The Fractal Markets Hypothesis is primarily a statistical interpretation of stock market prices, rather than a model of how the stock market, or investors in it, actually behave. Its main point is that stock market prices do not follow the random walk predicted by the EMH,[5] but conform to a much more complex pattern called a fractal. As a result, the statistical tools used by the EMH, which were designed to model random processes, will give systematically misleading predictions about stock market prices.

The archetypal set of random numbers is known as the 'normal' distribution, and its mathematical properties are very well known. A normal distribution with an average value of zero and a standard deviation of 1 will throw up a number greater than 1 15 percent of the time, a number greater than 2 just over 2 percent of the time, and a number greater than 3 only

once every 750 times, and so on. The chance of a 'far from average' event occurring diminishes rapidly and smoothly the farther the event is from the average.

The standard deviation of daily movements on the Dow Jones Industrial Average is roughly 1 percent. If stock market prices were generated by a normal process, then extreme movements – say a fall of more than 5 percent in just one day – would be vanishingly rare. The odds of any such event having occurred even once during the twentieth century would be just over 1 in a 100.

In fact, there were over sixty such daily downward movements (and over fifty daily upward movements of 5 percent or more) during the twentieth century.

The fact that extreme movements occurred roughly 10,000 times more often than for a random process is fairly strong evidence that the process is not random at all (and there's lots more evidence besides this morsel).

A fractal set of numbers, on the other hand, is a far more pernicious beast. Specifically, it is much more likely to generate extreme events than a normal distribution, and one large movement is likely to be followed by another large movement – another feature of stock markets which the EMH finds very difficult to explain.[6] A fractal pattern also displays 'self-similarity': the data pattern looks the same regardless of whether you are looking at a short data period – such as one day, or a week – or longer periods, such as a year or even a century.

The basic idea behind a fractal is that each number in the series is a simple but nonlinear function of previous numbers in the series. This differs from a true 'random number generator' such as dice, where the next number is independent of all previous numbers – rolling a 6 now doesn't change the odds of rolling a 6 on your next throw, they will still be 1 in 6.

Applying this to the stock market, it is quite possible that each price movement is a complex function of previous price movements.

This might seem to imply that, if the fractal markets hypothesis is correct, it should be easy to make money out of the stock market – in which case the hypothesis would be invalid, since it isn't easy to profit as a trader. However, there is another key aspect of fractal systems which comes into play here, which is known as 'sensitive dependence on initial conditions.'

Even if you knew precisely the 'system' which generated the Dow Jones Industrial Average, you could never know the precise value of the index because of rounding error. Let's say your initial measure of its value was out by 1/10th of a percent – rather than being, say, 10396.5, it was actually 10396.6.

One day (or iteration) later, your model would be wrong by (say) 1 percent; one day later by 10 percent; and a day after that, it would be completely useless as a means of predicting the following day's value. This is because any measurement errors you make in specifying the initial conditions of a fractal model grow exponentially with time, whereas for a random model the errors normally grow linearly (and can even fall with time for a stable system). As Ott puts this dilemma: 'The exponential sensitivity of chaotic solutions means that, as time goes on, small errors in the solution can grow very rapidly (i.e., exponentially). Hence, after some time, effects such as noise and computer roundoff can totally change the solution from what it would be in the absence of these effects' (Ott 1993).

Ott gives the example of a chaotic function called the Henon Map being simulated on a computer which is accurate to fifteen decimal places: the smallest difference it can record between two numbers is 0.00000000000001. He shows that if your initial measurement of the system was out by precisely this much, then after forty-five iterations of the model, your estimate of where the system is would be completely wrong. Attempting to overcome this problem by more computing power is futile: 'Suppose that we wish to predict to a longer time, say, twice as long. Then we must improve our accuracy by a tremendous amount, namely 14 orders of magnitude! In any practical situation, this is likely to be impossible. Thus, the relatively modest goal of an improvement of prediction time by a factor of two is not feasible' (ibid.).

Applying this to the stock market, it is possible to hold two apparently contradictory attitudes simultaneously: the market is driven largely by endogenous processes in which previous price movements determine future price movements; and it is impossible – or very difficult – to predict which way the market will move, and by how much.

Much of the Fractal Markets Hypothesis is directed at critiquing the notion that price movements in the stock market are random – as I noted earlier, it is primarily a way to characterize the properties of the statistics

the market throws up, rather than a theory of how the market actually behaves. However, it makes one important behavioral observation that runs directly counter to the EMH's assumptions about investors.

This is that the market will be stable when it allows investors with different time horizons to trade smoothly. As a result, heterogeneity – the fact that all investors are not the same – is a vital part of this theory. As Peters puts it:

> Take a typical day trader who has an investment horizon of five minutes and is currently long in the market. The average five-minute price change in 1992 was –0.000284 per cent [it was a 'bear' market], with a standard deviation of 0.05976 per cent. If, for technical reasons, a six standard deviation drop occurred for a five minute horizon, or 0.359 per cent, our day trader could be wiped out if the fall continued. However, an institutional investor – a pension fund, for example – with a weekly trading horizon, would probably consider that drop a buying opportunity because weekly returns over the past ten years have averaged 0.22 per cent with a standard deviation of 2.37 per cent. In addition, the technical drop has not changed the outlook of the weekly trader, who looks at either longer technical or fundamental information. Thus the day trader's six-sigma [standard deviation] event is a 0.15-sigma event to the weekly trader, or no big deal. The weekly trader steps in, buys, and creates liquidity. This liquidity in turn stabilizes the market. (Peters 1994)

The Fractal Markets Hypothesis thus explains the stability of the market by the realistic assumption that traders differ in their time horizons. It also alleges that instability is likely to occur if all investors suddenly switch to the same time horizon.

The Fractal Markets Hypothesis is thus more consistent with stock market data, more robust, and completely untainted by any assumption that the market is in, or tends toward, equilibrium. But it still doesn't provide an answer to what is actually generating the data: what is the system behind the fractal? To answer that question, we have to return to the kind of institutional analysis that Keynes provided in 1936. Two such analyses have

been provided: by Robert Haugen in the 'Inefficient Markets Hypothesis,' and Hyman Minsky in the 'Financial Instability Hypothesis,' as discussed in Chapter 13.

The Inefficient Markets Hypothesis

After a long career as an academic finance economist, Bob Haugen presents the diametrically opposite case from the Efficient Markets Hypothesis with gusto in three short books: *The Beast on Wall Street*, *The New Finance*, and *The Inefficient Stock Market*. Anyone who is or is thinking of speculating in the market – or is suffering from having done so – should read all three. Amid an extensive catalogue of data that contradicts the Efficient Markets Hypothesis, Haugen presents the alternative case for 'a noisy stock market that overreacts to past records of success and failure on the part of business firms, and prices with great imprecision' (Haugen 1999b).[7]

Though Haugen makes no reference to Keynes, the reasons he gives for the market behaving in this way echo the arguments Keynes made back in 1936 – that in the real world of uncertainty, few if any stock market speculators trade on the basis of new information. Instead, they trade on the basis of how they think other market participants will, on average, expect the market to react to news. Unlike the efficient market hypothesis, this 'news' can include the most recent movements of stock prices themselves.

In fact, in today's stock market, the major news will always be the most recent movements in stock prices, rather than 'real' news from the economy.

Haugen argues that there are three sources of volatility: event-driven, error-driven, and price-driven (the Efficient Markets Hypothesis models only the first, on the belief that the other two can't exist in the equilibrium of an efficient market). The second results from the market overreacting to news, then over-adjusting itself once the initial mistake has become obvious.

The third is the phenomenon of the market reacting to its own volatility, building price movements upon price movements, in the same way that neighborhood dogs can sometimes keep yelping almost indefinitely after

one of them has started. Haugen argues that this endogenous instability accounts for over three-quarters of all volatility.

He also argues that the market's endogenous instability has a severe and deleterious impact on the functioning of a modern capitalist economy.

First, if the stock market has any role at all in directing investment funds, then its valuations will direct them very badly. Price-driven volatility will lead to some companies which will in the long term turn out to be worthless being given massive funding – which will then be wasted – while potentially worthy ventures will be starved of funds. According to Haugen, the managers of a firm that has been seriously overvalued by the market over-invest: 'Consumers get what they don't want.' On the other hand, an undervalued firm 'would invest to produce a product that consumers really want, if it could raise capital at a fair price, but in this market, it can't' (Haugen 1999a). Overall, by providing too much money to ventures which, in the long run, are going to turn out not to be all that profitable, while providing too little money to those which, in the long run, will be worthwhile, the market causes the economy to grow less rapidly and less smoothly than it could.

Secondly, Haugen argues that, as well as causing investment to be badly apportioned, the stock market's endogenous volatility reduces the overall level of investment. Over the long term, the risk-free real rate of return has averaged about 1 percent, whereas the risk premium for investing in stocks has averaged about 6 percent. This means that investors have required a return of about 7 percent on their investments – with the result that investments which predict a lower expected rate of return don't get funded.

Haugen argues that investors require this higher return to compensate them for the risk of investing on the market, yet most of this risk results from the endogenous instability of the market itself. Since his statistical research indicates that price-driven volatility accounts for almost 95 percent of all volatility, he argues that this risk premium would be substantially lower – perhaps as low as just 0.4 percent, versus the 6 percent it has been historically (ibid.). If the risk premium could be reduced to this level, then the rate of investment would be substantially higher: 'Price-driven volatility has greatly inhibited investment spending over the years. Ultimately, it has acted, and acts, as a serious drag on economic growth' (ibid.).

At the individual level, Haugen argues that the market's tendencies to overreact to news, and to be consumed with endogenous instability in prices, provide opportunities for non-institutional investors to profit from the market. However, at the macroeconomic level, Haugen believes, as did Keynes, that the economy would benefit if the market were restrained. His recommendation, again very similar to Keynes's, is to reduce the length of the trading day, or to limit trading to just one computer-assisted auction per day. He hopes that this would eliminate the phenomenon of price volatility driven by the market reacting to its own every move.

Econophysics

Broadly speaking, Econophysics is the application of the analytic techniques of modern physics to the social sciences. This is rather ironic, since the founders of neoclassical economics themselves aped what they thought were the methods of physicists in the nineteenth century.

What Walras and others attempted to mimic then was physics before it had developed a number of key innovations, including not merely quantum mechanics but the concept of entropy, which introduced the notion of irreversible change into physics. Mirowski coined the term 'proto-energetics' to describe the type of physics on which neoclassical economic theory modeled itself:

> From now on I shall need a term that will serve to identify a type of physical theory that includes the law of conservation of energy and the bulk of rational mechanics, but excludes the entropy concept and most post-1860 developments in physics. This collection of analytical artifices is more an historical than a systematic subset of physics: it includes the formalisms of vector fields, but excludes Maxwell's equations, or even Kelvin's mechanical models of light.
>
> Since this resembles the content of the energetics movement, I trust it will not do the phenomena too much violence to call it 'proto-energetics.' Classical thermodynamics diverges from proto-energetics in one very critical aspect: Thermodynamic processes only change in one direction. In proto-energetics, time is isotropic, which means that no physical laws would be violated

> if the system ran backward or forward in time. (Mirowski 1989: 63)

Since then, physics has evolved rapidly, while economics has developed rather as can the language of a group of migrants who, separated from their home country, hang on to terms that have become obsolete in the original language.

The new incursion of physics into economics is being led by physicists themselves, and motivated partly by the innate curiosity that physicists like Cheng Zhang had about economic issues, and partly by the fact that 'we'd run out of things to do in physics.'[8] Though called Econophysics, a more accurate term for this school of thought at present would be 'Finaphysics' – since the vast bulk of its research has concerned the behavior of financial markets, rather than the broader economy.

This orientation reflects the inherently empirical nature of physics, and the fact that its analytic techniques have been developed to process enormous amounts of data generated by non-equilibrium experiments in physics. Economics does not generate a sufficient volume of data, but financial markets do in abundance, with the price and volume data of financial transactions; as Joe McCauley put it, 'the concentration is on financial markets because that is where one finds the very best data for a careful empirical analysis' (McCauley 2004: xi).

Given that it is a relatively new field, there are numerous explanations of the volatility of financial markets within Econophysics – including Power Law models of stock market movements (Gabaix, Gopikrishnan et al. 2006), Didier Sornette's earthquake-based analysis (Sornette 2003), Joe McCauley's empirically derived Fokker-Planck model (McCauley 2004), and Mandelbrot's fractal geometry (Mandelbrot and Hudson 2004) – and it would require another book to detail them all.

A unifying theme is that the behavior of financial markets is driven by the interactions of numerous market participants with each other, and these generate a highly unstable and therefore relatively unpredictable time series in financial data themselves. These characteristics resemble the behavior of fissile materials in a nuclear reactor, or tectonic plates in an earthquake zone, physical processes for which physicists have developed an enormous

arsenal of mathematical analytic techniques in the last century. Econophysics is essentially the application of these techniques to financial data.

This Econophysics explanation of the unpredictability of finance markets is thus diametrically opposed to the explanation that neoclassical economics has given of precisely the same phenomenon – the difficulty of predicting the market – and Econophysicists react with incredulity to the simplistic 'random disturbances to an equilibrium process' explanation that neoclassical economists provide:

> Three states of matter – solid, liquid, and gas – have long been known. An analogous distinction between three states of randomness – mild, slow and wild – arises from the mathematics of fractal geometry. Conventional finance theory assumes that variation of prices can be modeled by random processes that, in effect, follow the simplest 'mild' pattern, as if each uptick or downtick were determined by the toss of a coin. What fractals show […] is that by that standard, real prices 'misbehave' very badly. A more accurate, multifractal model of wild price variation paves the way for a new, more reliable type of financial theory. (Mandelbrot and Hudson 2004: v)

> Economists teach that markets can be described by equilibrium. Econophysicists teach that markets are very far from equilibrium and are dynamically complex […] equilibrium is never a good approximation […] market equilibrium does not and cannot occur […] (McCauley 2004: 185)

> Uncertainties and variabilities are the key words to describe the ever-changing environments around us. Stasis and equilibrium are illusions, whereas dynamics and out-of-equilibrium are the rule. The quest for balance and constancy will always be unsuccessful. (Sornette 2003: xv)

I'll single out Didier Sornette's work here, not because it will necessarily be 'the' approach of Econophysics, but because he is making a

direct challenge to one tenet of conventional finance: that the market cannot be predicted. Using his model that the behavior of stock markets follows the 'log-periodic' pattern of earthquakes, he has made predictions about future stock market crashes that can be verified after the predicted crashes have (or have not) occurred: 'The Financial Crisis Observatory (FCO) is a scientific platform aimed at testing and quantifying rigorously, in a systematic way and on a large scale, the hypothesis that financial markets exhibit a degree of inefficiency and a potential for predictability, especially during regimes when bubbles develop' (Sornette 2011).

The result of the FCO can be tracked at the website www.er.ethz.ch/fco. The voluminous literature of the Econophysics movement can be tracked from its website unifr.ch/econophysics.

Conclusion: progress versus ossification

There are thus numerous vigorous alternatives to the failed paradigm of neoclassical finance – but students of economics are unlikely ever to learn of them, if all they do is study the textbooks of neoclassical finance courses. Despite the manifest failures of the Efficient Markets Hypothesis, and the recanting of it by the very same economists who developed it in the first place (Fama and French 2004), and the numerous stock market booms and crashes of the past quarter-century that could not have happened if the EMH were correct, textbooks continue to teach that finance markets are 'efficient,' in the bastardized way that economists use the term. This extract from a brand-new 2011 text – published seven years after the developers of the EMH concluded that 'the failure of the CAPM in empirical tests implies that most applications of the model are invalid' (ibid.: 25) – is typical:

> A financial market is informationally efficient if prices reflect all available information […] there are likely to be noise traders […] who trade on information unrelated to the true value of shares. If the information they trade on is random, they will tend to cancel each other out, leading to efficient market prices. However, it is likely that they trade on similar information, so that noise trading will lead to either an undervaluation or an overvaluation […] it would pay arbitragers to take an offsetting position […] This

> process will cause share prices to stay close to their true values […]
>
> Academic studies usually conclude that the share market is efficient. (Valentine, Ford et al. 2011: 245–7)

The unwillingness – and possibly even the inability – of neoclassical economists to admit that their paradigm has failed means that, if change is left to them alone, it will not occur.

Reforming finance?

The results of the non-neoclassical theories of stock market behavior surveyed in this chapter emphasize one point: asset markets perform their alleged role of the allocation of investment capital very poorly.[9] In this they echo Keynes's dictum during capitalism's last major crisis, that speculation should not be allowed to dominate capital formation and allocation:

> Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation. *When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done*. The measure of success attained by Wall Street, regarded as an institution of which the proper social purpose is to direct new investment into the most profitable channels in terms of future yield, cannot be claimed as one of the outstanding triumphs of laissez-faire capitalism – which is not surprising, if I am right in thinking that the best brains of Wall Street have been in fact directed towards a different object. (Keynes 1936: 159; emphasis added)

Though deregulation of the financial sector was far from the sole cause of the financial crisis that began in 2007, removing the fetters from the financial sector resulted in a crisis that was more extreme than it would have been had the previous regulations been kept in place. The USA's 'shadow banking' sector could not have invented and sold nearly so many 'weapons of financial mass destruction' as it did – to use Warren Buffett's evocative phrase – had Glass-Steagall not been abolished during Bill Clinton's term, for example.

I expect that history will judge that signing that bill into law was a far more reckless act than anything Clinton did with Monica Lewinsky. The comments of the handful of senators who opposed its repeal back in 1999 make interesting reading today:

> 'I think we will look back in 10 years' time and say we should not have done this but we did because we forgot the lessons of the past, and that that which is true in the 1930's is true in 2010,' said Senator Byron L. Dorgan, Democrat of North Dakota. 'I wasn't around during the 1930's or the debate over Glass-Steagall. But I was here in the early 1980's when it was decided to allow the expansion of savings and loans. We have now decided in the name of modernization to forget the lessons of the past, of safety and of soundness.'
>
> Senator Paul Wellstone, Democrat of Minnesota, said that Congress had 'seemed determined to unlearn the lessons from our past mistakes.'
>
> 'Scores of banks failed in the Great Depression as a result of unsound banking practices, and their failure only deepened the crisis,' Mr. Wellstone said. 'Glass-Steagall was intended to protect our financial system by insulating commercial banking from other forms of risk. It was one of several stabilizers designed to keep a similar tragedy from recurring. Now Congress is about to repeal that economic stabilizer without putting any comparable safeguard in its place.' (Labaton 1999)

In contrast, the beliefs of those who campaigned to end the Act have the ring of delusion:

> 'The world changes, and we have to change with it,' said Senator Phil Gramm of Texas, who wrote the law that will bear his name along with the two other main Republican sponsors, Representative Jim Leach of Iowa and Representative Thomas J. Bliley Jr. of Virginia. 'We have a new century coming, and we have an opportunity to dominate that century the same way we dominated this century. Glass-Steagall, in the midst of the Great Depression, came at a time when the

> thinking was that the government was the answer. In this era of economic prosperity, we have decided that freedom is the answer.' (Ibid.: 2)

Far from strengthening America, the financial follies that followed the repeal of Glass-Steagall have left it crippled at the start of the twenty-first century, and facing an economic eclipse by China. Far from reducing the role of the government in the US economy, the collapse of the Subprime Bubble has resulted in the government taking a larger role in the economy than it did even during the Great Depression.

Back in 2000, in the first edition of this book, I sided with the opponents of deregulation, noting that though they were 'mooted as "reforms" by their proponents, […] they were in reality retrograde steps, which have set our financial system up for a real crisis' (Keen 2001a: 255). That real crisis duly arrived eight years after the repeal of Glass-Steagall.

However, blocking the abolition of Glass-Steagall wouldn't have prevented the crisis, since its underlying cause was a debt bubble that had already driven the USA to the brink of Great Depression debt levels by 1989. Deregulation simply allowed the debt bubble to continue growing for another two decades, from the 170 percent of GDP level it reached as the 1990s recession began, and the 200 percent level it was at when Glass-Steagall was abolished, to the 300 percent of GDP peak hit ten years later in 2009.

I also wrote in 2000 that 'I can only hope that, if the crisis is serious enough, then genuine reform to the finance sector will be contemplated' (ibid.: 256), but the first and second response of government to this crisis has been to try to restore the 'business as usual' that applied prior to the crisis.

This is to be expected. Politicians, as Keynes observed long ago, are just as beholden to the ideas of neoclassical economics as are professional economists: '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' (Keynes 1936: 383).

It takes time before a real reformer comes along and challenges, not merely the belief systems that gave rise to mistakes like the abolition of Glass-Steagall, but the beneficiaries of those belief systems as well. We await a politician who is willing to not merely try to resuscitate the financial sector but to challenge it, as Roosevelt was during the Great Depression.

> [A] host of unemployed citizens face the grim problem of existence, and an equally great number toil with little return. Only a foolish optimist can deny the dark realities of the moment.
>
> Yet our distress comes from no failure of substance […] Plenty is at our doorstep, but a generous use of it languishes in the very sight of the supply. Primarily this is because the rulers of the exchange of mankind's goods have failed, through their own stubbornness and their own incompetence, have admitted their failure, and abdicated. Practices of the unscrupulous money changers stand indicted in the court of public opinion, rejected by the hearts and minds of men.
>
> True they have tried, but their efforts have been cast in the pattern of an outworn tradition. *Faced by failure of credit they have proposed only the lending of more money. Stripped of the lure of profit by which to induce our people to follow their false leadership, they have resorted to exhortations, pleading tearfully for restored confidence*. They know only the rules of a generation of self-seekers. They have no vision, and when there is no vision the people perish.
>
> The money changers have fled from their high seats in the temple of our civilization. We may now restore that temple to the ancient truths. The measure of the restoration lies in the extent to which we apply social values more noble than mere monetary profit. (Roosevelt 1933; emphasis added)

The reforms enacted in Roosevelt's era clearly worked, but as subsequent history has indicated, the problem with real reforms of our financial system is that, if successful, they will be abolished. The era of financial tranquility they usher in will be misinterpreted – particularly if

economists continue to believe in the fantasy world of neoclassical economics – as inherent to capitalism, and not merely the product of regulations that inhibit the financial system's innate tendency to create too much debt.

Politicians who did not live through the crisis that caused these regulations to be enacted will then weaken these regulations over time, and we will be back in a crisis again.

Fundamentally, reforms of the financial sector fail because they try to constrain the sector's innate desire to create debt. They will work for a while in the aftermath to a crisis like the Great Depression or the Great Recession, where the carnage wreaked by a financial crisis is so great that the sector behaves prudently for a while. However, the incentives to create debt are so great for this sector that, over time, a debt-driven culture will replace prudence.

Institutional control of finance is also flawed, for reasons that should be obvious from our current crisis: 'regulatory capture.' Not only are regulators slower to move than the organizations they are intended to control, they often become advocates rather than monitors of those organizations. There is little doubt that Greenspan's actions in rescuing the financial sector from itself after numerous crises, in championing the development of financial assets now universally regarded as toxic, and in restricting the development of new regulations to control new financial instruments, turned a potentially garden-variety would-be depression in 1987 into the near-death experience of the Great Recession. The regulators, by delaying the inevitable for two decades, have made this crisis more intractable than it would have been without them.[10]

Reforms also fail because they do not recognize that the financial system has what Kornai called a 'soft budget constraint' (Kornai, Maskin et al. 2003).[11] A bank is not constrained in its lending by its reserves, but by the willingness of borrowers to take on additional debt (see Holmes 1969; Moore 1979). It therefore faces a 'soft budget constraint': to expand its operations, all it has to do is to persuade borrowers (firms and households) to borrow more money, and its income will grow – as will the level of debt.

This growth in bank income and debt is in turn dependent on the willingness of borrowers to incur debt. If this is based solely on their

income, then the 'hard budget constraint' that households and firms face will put a limit on the amount of debt they will take on.

If, however, a Ponzi scheme develops in some asset class – so that people are willing to borrow money in the expectation of future capital gain – then the amount of borrowing will no longer be constrained by incomes. While capital gains are made, the borrowers also operate with a soft budget constraint: any deficiency of revenue over costs can be covered by selling an asset whose price has been inflated by the increase in leverage.

Initially banks – after they have forgotten the previous crisis – will be willing to fund this process, since it increases their incomes. But inevitably a crisis will result because the borrowing is adding to debt levels without increasing the capacity of the economy to service those debts. Though individuals can operate with a soft budget constraint while the price bubble lasts, the entire economy is stuck with the hard budget constraint that, in the long run, the debt must be serviced from income.[12]

If we are to prevent this process playing out yet again in the future, then we need to prevent the formation of Ponzi schemes in the first place. Unfortunately, the way that financial assets are currently defined contains the seeds of not one Ponzi scheme but two.

Because shares currently have an indefinite lifespan, it is quite possible for someone to assert, as Henry Blodget did about Amazon in 1998, that a given company's shares will go from $1 to $400 in a matter of a year (Blodget 2010). Faced with those hypothetical gains, ordinarily sane people are liable to succumb to the euphoria that produces them and be willing to borrow to speculate.

Because there is no effective limit to the debt that can be secured against a property, property prices reflect the leverage that people are willing to incur to buy them. When houses are bought as residences, that isn't a problem. But when they are bought as speculative assets, then again people's willingness to borrow can become unhinged from their incomes.

We therefore need not merely reforms, but changes to the incentives that encourage people into debt – because so long as those incentives exist, we can be sure that at some point the financial sector will find a way to entice the public into debt, leading to yet another financial crisis.

I have two simple proposals to achieve this objective. Neither of them has any chance of being implemented immediately, but there is some prospect that they might be considered more seriously if, as I expect, this crisis causes a prolonged slump for America that resembles Japan's two 'Lost Decades' since its bubble economy collapsed in the early 1990s. They are:

1. Jubilee shares: To redefine shares so that, if purchased from a company directly, they last for ever (as all shares do now), but once these shares are sold by the original owner, they last another fifty years before they expire; and
2. Property Income Limited Leverage: To limit the debt that can be secured against a property to ten times the annual rental of that property.

Jubilee shares Ninety-nine percent of all trading on the stock market involves speculators selling pre-existing shares to other speculators. Valuations are ostensibly based on the net present value of expected future dividend flows, but in reality based on the 'Greater Fool' principle, where rising debt funds the Greater Fool. Anticipated capital gain is the real basis of valuation, and the overwhelming source of that capital gain is not increased productivity, but increased leverage. This trading adds zip to the productive capacity of society, while promoting bubbles in stock prices as leverage drives up prices, encouraging more leverage, leading to a crash after price-to-earnings ratios reach levels even the Greater Fool regards as ridiculous. When the share market crashes, prices fall but the debt that drove prices up remains.

If instead shares on the secondary market lasted only fifty years, then even the Greater Fool couldn't be enticed to buy them with borrowed money – since their terminal value would be zero. Instead a buyer would purchase a share on the secondary market only in order to secure a flow of dividends for fifty years (or less). One of the two great sources of rising unproductive debt would be eliminated.

This reform would dramatically tilt the balance in favor of the raising of capital via primary share issues, force valuations to be based on prospective earnings rather than capital gain, and make leveraged speculation on the value of shares on the secondary market much less attractive.

Jubilee shares could be introduced very easily, if the political will existed – something that is still years away in practice. All existing shares could be grandfathered on one date, so that they were all ordinary shares; but as soon as they were sold, they'd become Jubilee shares with an expiry date of fifty years from the date of first sale.

Property Income Limited Leverage Obviously some debt is needed to purchase a house, since the cost of building a new house far exceeds the average wage. But debt past a certain level drives not house construction, but house price bubbles: as soon as house prices start to rise because banks offer more leverage to home buyers, a positive feedback loop develops between house prices and leverage, and we end up where Australia and Canada are now, and where America was before the Subprime Bubble burst: with house prices out of reach of ordinary wage earners, and leverage at ridiculous levels so that 95 percent or more of the purchase price represents debt rather than owner equity.

Property Income Limited Leverage ('the *PILL*') would break the positive feedback loop that currently exists between leverage and property prices. With this reform, all would-be purchasers would be on equal footing with respect to their level of debt-financed spending, and the only way to trump another buyer would be to put more non-debt-financed money into purchasing a property.

This doesn't happen under our current system because the amount extended to a borrower is allegedly based on his/her income. During a period of economic tranquility that is initiated after a serious economic crisis has occurred and is finally over – like the 1950s after the Great Depression and World War II – banks set a responsible level for leverage, such as the requirement that borrowers provide 30 percent of the purchase price, so that the loan-to-valuation ratio was limited to 70 percent. But as economic tranquility continues, banks, which make money by extending debt, find that an easy way to extend more debt is to relax their lending standards, and push the loan-to-valuation ratio (LVR) to, say, 75 percent.

Borrowers are happy to let this happen, for two reasons: borrowers with lower income who take on higher debt can trump other buyers with higher incomes but lower debt in bidding on a house they desire; and the increase

in debt drives up the price of houses on sale, making the sellers richer and leading all current buyers to believe that their notional wealth has also risen.

Ultimately, you get the runaway process that we saw in the USA, where leverage rises to 95 percent, 99 percent, and even beyond – to the ridiculous level of 120 percent, as it did with Liar Loans at the peak of the subprime frenzy. Then it all ends in tears when prices have been driven so high that new borrowers can no longer be enticed into the market – since the cost of servicing that debt can't be met out of their incomes – and as existing borrowers are forced into bankruptcy by impossible repayment schedules. The housing market is then flooded by distressed sales, and the bubble bursts. The high house prices collapse, but as with shares, the debt used to purchase them remains.

If we instead based the level of debt on the income-generating capacity of the property being purchased, rather than on the income of the buyer, then we would forge a link between asset prices and incomes that is currently easily punctured by rising debt. It would still be possible – indeed necessary – to buy a property for more than ten times its annual rental. But then the excess of the price over the loan would be genuinely the savings of the buyer, and an increase in the price of a house would mean a fall in leverage, rather than an increase in leverage as now. There would be a negative feedback loop between house prices and leverage. That hopefully would stop house price bubbles developing in the first place, and take dwellings out of the realm of speculation back into the realm of housing, where they belong.

Conclusion

As the above 'bubble on a whirlpool of speculation' quote from Keynes indicates, this is not the first time that the conventional theory of finance has been attacked. What is unique about these most recent critiques is that the contribution from physicists in particular turns one of the alleged strengths of neoclassical economics against it: mathematics.

In the past, neoclassical economists have disparaged their critics with the assertion that they object to neoclassical theory because they don't understand mathematics. This time, however, they are under attack, not

merely from critics who eschew the use of mathematics, but from those to whom mathematical thinking is second nature.

The impact of this power inversion can be seen in the physicist Joe McCauley's observations about the need to reform economics education:

> The real problem with my proposal for the future of economics departments is that current economics and finance students typically do not know enough mathematics to understand (a) what econophysicists are doing, or (b) to evaluate the neo-classical model (known in the trade as 'The Citadel') critically enough to see, as Alan Kirman put it, that 'No amount of attention to the walls will prevent The Citadel from being empty.'
>
> I therefore suggest that the economists revise their curriculum and require that the following topics be taught: calculus through the advanced level, ordinary differential equations (including advanced), partial differential equations (including Green functions), classical mechanics through modern nonlinear dynamics, statistical physics, stochastic processes (including solving Smoluchowski–Fokker–Planck equations), computer programming (C, Pascal, etc.) and, for complexity, cell biology.
>
> Time for such classes can be obtained in part by eliminating micro- and macro-economics classes from the curriculum. The students will then face a much harder curriculum, and those who survive will come out ahead. So might society as a whole. (McCauley 2006: 607–8)

This amplifies a point that, as a critic of economics with a reasonable grounding in mathematics myself, has long set me apart from most other critics: neoclassical economics is not bad because it is mathematical per se, but because it is bad mathematics.

16 | DON'T SHOOT ME, I'M ONLY THE PIANO

Why mathematics is not the problem

Many critics of economics have laid the blame for its manifest failures at the feet of mathematics. Mathematics, they claim, has led to an excessive formalism in economics, which has obscured the inherently social nature of the subject.

While it is undeniable that an inordinate love of mathematical formalism has contributed to some of the intellectual excesses in economics, generally this reaction is as erroneous as blaming the piano for the discordant notes of a bad piano player. If anything should be shot, it is the pianist, not the piano.

Though mathematics has definite limitations, properly used it is a logical tool that should illuminate, rather than obscure. Economists have obscured reality using mathematics because they have practiced mathematics badly, and because they have not realized the limits of mathematics.

The kernel

If you divided the world's population into those who dislike mathematics, those who like it, and those who were indifferent, I suspect that 95 percent would be in the first camp, 5 percent in the second, and 0 percent in the third. Neoclassical economists come almost exclusively from the 'like it' camp, and therefore their arguments are almost always expressed in mathematical form. Most critics of economics come from the 'dislike it' camp, and frequently criticism of mathematics per se forms part of their criticism of economics.

Call me weird: I'm a critic of neoclassical economics who likes mathematics. But I am not alone. There are numerous mathematically

inclined critics of neoclassical economics, and in many ways this book was written to convey their critiques to a non-mathematical audience. Not only is it possible to simultaneously like mathematics and dislike mainstream economics, but a sound knowledge of mathematics makes you an even more confirmed opponent – because much of the mathematics in conventional economic theory is unsound.

At one level, it is unsound because conditions that economists assume contradict other conditions needed for their models, so that the theory is built on a mathematical error. For example, as shown in Chapter 4, one crucial assumption in the neoclassical argument in favor of small competitive firms over monopolies violates one of the most basic rules of calculus, the Chain Rule.

At a second level, it employs the wrong mathematical tools to analyze the dynamic processes that characterize a market economy – employing complicated comparative static equilibrium analysis when dynamic systems analysis is not only more appropriate but frankly easier.

At a third and more profound level, conventional economics is mathematically unsound because it has not learnt the lesson which true mathematicians have learnt in the last century: that there are limits to mathematical logic.

The roadmap

In this chapter I argue that conventional economics has abused mathematics in two main ways: by practicing bad mathematics, and by not acknowledging the limitations of mathematics. Many economic theorems result in logical contradictions which economists fail to recognize as such, and many other theorems are derived by falsely assuming that different quantities are in fact equal. Modern mathematics has also realized that there are limits to mathematical logic, but economists have evaded this realization by effectively but unintentionally isolating themselves from mainstream mathematical science.

Bad mathematics

In a classic instance of 'those who live by the sword die by the sword,' the school of economics that prides itself on being mathematical is subject

to the indictment that its mathematics is erroneous. There are numerous theorems in economics that rely upon mathematically fallacious propositions. There are basically four ways in which this manifests itself in economic theory:

- logical contradiction, where the theory is allegedly 'saved' by an assumption which in fact contradicts what the theory purports to show;
- omitted variables, where an essential aspect of reality must be ignored to derive the mathematical results that economists wish to prove;
- false equalities, where two things that are not quite equal are treated as if they are equal; and
- unexplored conditions, where some relation is presumed without exploring what conditions are needed to make this relation feasible.

Logical contradiction The case outlined in Chapter 2 – the failed attempt to aggregate individual preferences to form community preferences with the same properties – is an excellent example of logical contradiction.

The economic theory of consumer behavior begins with the proposition that it is possible to aggregate individual demand curves to derive a market demand curve that has the same characteristics as an individual's demand curve. Economists have proved that this is possible only when the Sonnenschein-Mantel-Debreu conditions apply: (a) that all consumers have the same preference map; and (b) that preferences do not change with income.

Condition (a) effectively means that there is only one consumer. Condition (b) effectively means that there is only one commodity. Aggregation is therefore strictly possible if there is only one consumer and only one commodity.

Clearly this is not aggregation at all.

A good mathematician would recognize this as proof by contradiction (Franklin and Daoud 1988). This is a clever technique whereby, to prove a statement, you assume its opposite and then show a contradiction. Therefore the statement is true.

For example, consider the problem that confronted Pythagoras and friends when they tried to work out the length of the hypotenuse of a right-angled triangle whose other sides were both one unit long. According to Pythagoras's theorem that 'the square of the length of the hypotenuse is equal to the sum of the squares of the other sides,' this meant that the length of the hypotenuse was the square root of 2.

These Ancient Greeks initially believed that all numbers were 'rational,' which meant that any number could be expressed as the fraction of two integers: thus 1.5, for example, is the integer 3 divided by the integer 2. But they could never accurately measure the value of the square root of 2 in terms of the ratio of two integers: every more accurate measurement led to a different fraction.

The reason that they couldn't find the 'right' two integers is that the square root of 2 is an irrational number: it can't be defined as the ratio of two integers.

This can be proved quite easily using proof by contradiction. You start with the opposite assumption – that it is possible to express the square root of 2 as a ratio of two integers. You then work on from this point, to show this leads to a contradiction. Therefore you show that the square root of 2 is irrational.[1]

The proofs which led to the Sonnenschein-Mantel-Debreu conditions can easily be described in the same fashion. You wish to prove that it is *not* possible to aggregate individual preferences to derive community preferences which display the same characteristics as individual preferences.

You start with the opposite assumption – that it is possible to aggregate the preferences of two or more different consumers over two or more different commodities to market demand curves that have the same characteristics as individual demand curves. You then show that this is possible only if there is only one individual and only one commodity. This contradicts your starting assumption that there are multiple different consumers and different commodities. Therefore you have proved, by contradiction, that it is not possible to aggregate individual preferences to derive market demand curves that obey the 'Law' of Demand.

The trouble is that economists were hoping that they could prove that it was possible to aggregate. In this sense, they were in the same situation as the ancient Pythagoreans, who were trying to prove that all numbers were rational – they were not at all pleased to find that they were wrong.

At least the Pythagoreans relented: they abandoned the belief that all numbers were rational, and accepted that there were numbers which could not be described as the ratio of two integers. Mathematics thus absorbed the existence of irrational numbers, and went on from there to many other discoveries.

Economists, on the other hand, have been unwilling to abandon their concept of rationality. Faced with an equivalent discovery – that society cannot be understood as the sum of the rational individuals within it – economics has instead enshrined these and similar logical contradictions as 'intuitively reasonable' (Gorman 1953) abstractions that are needed to forge a link between individual and collective utility.

This is bad mathematics. It has led to bad economics, which has avoided the more complex but richer visions of the economy that flow from coming to terms with the myriad contradictions of the simplistic notions underlying neoclassical economics.

Omitted variables

Bad mathematics also shows up in such hallowed economic concepts as maximizing profit by equating marginal cost and marginal revenue. As is shown in Chapter 4, this mantra of the everyday economist is false even on its own terms, but it is doubly so if we ignore time. Once time is rightly included in the analysis, then it is mathematically evident that, to maximize profits over time, a firm should ensure that its marginal revenue exceeds its marginal cost.

Many critics of conventional economics have previously argued that time is the most crucial variable left out of economic analysis, but most of these critics have then eschewed mathematics itself as a result. However, good mathematical economics incorporates time as an essential aspect of reality, and results in a type of economic analysis that is profoundly different from conventional neoclassical economics.

Time is not the only vital factor omitted by neoclassical economists, of course. Other notable examples include uncertainty, and the formation of expectations under uncertainty, and, most importantly of all, given the debt-induced crisis we are now in, money and debt. The basis of this is the so-called 'money illusion,' which is drummed into new economics students in their first year – ordinarily when most are too naive about the world to see the fallacy behind it[2] – resulting in macroeconomic models that ignore the role of money and debt in our fundamentally credit-driven market economies.

False equalities

One popular but erroneous step in conventional economic argument is to assert that something that is extremely small can be treated as zero. This is especially so when economists then pretend to 'aggregate up' from the individual firm to derive a result that applies at the aggregate level. What results is not mathematical analysis, but a mathematical sleight of hand – the intellectual equivalent of a magician's trick.

The model of perfect competition illustrates this nicely. The argument starts with the market having a downward-sloping demand curve and an upward-sloping supply curve. Step one of the trick is to omit showing the downward-sloping marginal revenue curve, which must be there if the market demand curve slopes downward, and which is distinctly different from, and steeper than, the demand curve. Step two is to break the market demand curve into lots of tiny bits, each of which must also slope downwards if the whole curve slopes downwards, but to persuade the audience that the slope of each of these little lines is so flat that they can be treated as horizontal. Hence for the individual firm, the demand curve and the marginal revenue curve are identical. The final stage of the trick is to return to the market level by adding up all the individual firm's marginal cost curves, and to show that price is set by the intersection of the demand curve and the supply curve. The troublesome market marginal revenue curve has been made to disappear, and the trick is complete.

The special irony of this piece of magic is that the magician doesn't realize that a trick is being pulled. Economists are so used to presuming that

an infinitesimal amount is equivalent to zero that they don't even realize they are breaching one of the fundamental rules of mathematics.

Unexplored conditions

There are numerous examples of this phenomenon. The comparison of monopoly to perfect competition presumes that the supply curve for the competitive industry is equivalent to the marginal cost curve for the monopoly. However, this is possible only if the two 'curves' are the same horizontal straight line (Keen and Standish 2010: 89–91). The theory of the labor market presumes that the supply curve of labor is upward sloping; Chapter 5 showed that this is not a necessary outcome of the neoclassical theory of labor supply. The analysis of production requires that the money value of capital is an adequate measure of the amount of capital; Chapter 6 showed that it is not, once we acknowledge that machines are produced by other machines and labor.

That these (and doubtless many other) logical conundrums exist indicates that economists do not explore the logical foundations of their beliefs. This in itself is not necessarily unscientific; as discussed in Chapter 7, it is a standard practice that scientists in a given school within a science do not challenge what Lakatos describes as the 'hard core' of their beliefs. But it is a sign of how fragile the neoclassical hard core is that so many elements of it can be shown so easily to be internally inconsistent.

It is also unscientific that, when such logical flaws are either pointed out by critics (as with Sraffa's critique in Chapter 6) or discovered by believers (as with the Sonnenschein-Mantel-Debreu conditions in Chapter 2), neoclassical theory adopts 'ancillary assumptions' which are clearly absurd (such as the 'machines as putty' notions which were put forward during the debate with Sraffa and his supporters, or the SMD conditions used to save the theory of consumption, which amount to assuming that all consumers have identical, income-independent tastes). This, to Lakatos, is the sign of a degenerative scientific research program which is preoccupied with adjusting its ancillary beliefs to defend its hard core, whereas a truly progressive program would be expanding the range of real-world phenomena which its theory explains. The school of economics which gives pride of place to the word 'rational' hardly displays rational behavior when

its core beliefs are challenged. I expect that the new logical conundrums pointed out in this book will generate further displays of irrational behavior by conventional economists.[3]

Mathematics is therefore not the reason why conventional economics is so bad. Instead, bad economics is supported by bad mathematical practice. But this is only half of the story about how economics has abused mathematics. Economics has also accidentally inoculated itself against many of the advances of modern mathematics. One essential aspect of modern mathematics that economics has not realized is that mathematics today has a humility that economics lacks, because mathematicians have proved that mathematics has limits.

The limits to mathematics

Economics remains perhaps the only area of applied mathematics that still believes in Laplace's dictum that, with an accurate enough model of the universe and accurate enough measurement today, the future course of the universe could be predicted.

For mathematicians, that dictum was dashed in 1899 by Poincaré's proof of the existence of chaos. Poincaré showed that not only was it impossible to derive a formula which could predict the future course of a dynamic model with three or more elements to it, but even any numerical approximation to this system would rapidly lose accuracy. The future could be predicted only if the present was known to infinite accuracy, and this was clearly impossible.

Today, mathematicians are quite comfortable with the proposition that most mathematical problems cannot be explicitly solved in a manner which yields the kind of didactic statements which economics makes as a matter of course – such as 'perfect competition gives superior welfare outcomes to monopoly,' 'free trade is superior to protection,' and so on. Such definitive pronouncements are generally only possible when the problem is essentially the same as working out where two straight lines intersect. This class of models is known as algebraic.

Some algebraic equations are rather difficult to solve because there is no standard formula to apply. But there are standard formulas available to solve systems of algebraic equations where all the equations are 'straight

lines.' This is the type of mathematics which economic theory generally tries to apply to economic problems.

However, this body of mathematics is an appropriate model of only a tiny subset of real-world systems – and that subset does not include true economic analysis.[4]

The more appropriate starting point for mathematical models of the economy is dynamic equations, in which the relationships between variables cannot be reduced to straight lines. These are known in mathematics as nonlinear differential equations. The vast majority of these cannot be solved, and once three or more such equations interact, they are impossible to solve.

Table 16.1 summarizes the situation. Economic theory attempts to analyze the economy using techniques appropriate to the upper left-hand part of Table 16.1 (with italic text), when in fact the appropriate methods are those in the lower right-hand part (with cells shaded gray).

Other developments, such as Gödel's proof that a mathematical system cannot be self-contained – so that it must take some axioms on faith – and the proof that there were some mathematical problems which could not be solved, added to this realization by mathematicians and physicists that mathematics and science had innate limits. As a result, in place of Laplace's grand conceit, there is a humility to modern mathematics. The future cannot be known, mathematics cannot solve every problem, some things may not be knowable.

TABLE 16.1 The solvability of mathematical models (adapted from Costanza 1993)

	Linear			Non-linear		
Equations	One	Several	Many	One	Several	Many
Algebraic	*Trivial*	*Easy*	*Possible*	Very difficult	Very difficult	Impossible
Ordinary differential	Easy	Difficult	Essentially impossible	Very difficult	Impossible	
Partial differential	Difficult	Essentially impossible	Impossible			

But these epiphanies passed economists by: they continue to believe in a clockwork universe, in which a proper specification of the conditions of today could enable you to predict the future for all time. Nowhere is this vanity more obvious than in the school's defining works, Walras's *Elements of Pure Economics* and Debreu's *Theory of Value*.

Walras's arrogance towards those economists who would not practice mathematics is still the typical attitude today held by economists towards those who criticize their use of mathematics:

As for those economists who do not know any mathematics, who do not even know what is meant by mathematics and yet have taken the stand that mathematics cannot possibly serve to elucidate economic principles, let them go their way repeating that 'human liberty will never allow itself to be cast into equations' or that 'mathematics ignores frictions which are everything in social science' and other equally forceful and flowery phrases. They can never prevent the theory of the determination of prices under free competition from becoming a mathematical theory. Hence, they will always have to face the alternative either of steering clear of this discipline and consequently elaborating a theory of applied economics without recourse to a theory of pure economics or of tackling the problems of pure economics without the necessary equipment, thus producing not only very bad pure economics but also very bad mathematics. (Walras 1954 [1874])

As this book has shown, it is neoclassical economists who have been guilty of very bad mathematics. But just as important is the fact that they do not appreciate the limits to mathematics.

At least Walras could be forgiven for not being aware of Poincaré's theorem of 1899 – though he had sought out Poincaré in a forlorn attempt to garner support for his mathematization of economics. Debreu's opus pre-dated the rediscovery of chaos by Lorenz, but that he could even conceive of modeling the economy as a system in which all production and exchange decisions were 'made now for the whole future,' and in which the theory of uncertainty was 'free from any probability concept and formally identical with the theory of certainty,' betrayed a profound lack of appreciation of the mathematics of his day (not to mention the real world).

The modern manifestation of this ignorance of the limits of mathematics is a widespread – though not universal – failure by economists to appreciate the importance of nonlinear analysis and chaos theory. If I had a cent for

every time I heard an economist comment that 'chaos theory hasn't amounted to much' – well, I wouldn't be wealthy, but I could afford an expensive meal or two.

Chaos theory has 'not amounted to much' in economics because its central tenets are antithetical to the economic obsession with equilibrium. In other sciences, chaos theory, complexity analysis and their close cousin evolutionary theory have had profound impacts. It shows how isolated economics has become from the scientific mainstream of the late twentieth and early twenty-first century that such ignorant views could be commonplace.

The recurring nightmare of straight lines

Virtually every critique detailed in this book has led to the result that some relationship between phenomena that economics argued was curved had to instead be a straight line.

The economic theory of consumer behavior argued that a person's consumption of a commodity could change in any direction as his income rose: if it was a luxury, consumption would rise relative to other commodities; if a necessity, consumption could fall. Instead, the Sonnenschein-Mantel-Debreu conditions show that if the theory is to aggregate from the individual to the market demand curve, Engel curves must be straight lines.

The economic theory of production argues that output is subject to diminishing marginal returns, so that as the variable input rises, output rises less than proportionately – the relationship is curved. Sraffa's critique shows that, in general, output should rise proportionately: the relationship should be a straight line.

The economic theory of competition argues that perfect competition is superior to monopoly. But the only conditions under which the comparison is watertight involve a straight-line relationship between inputs and outputs – not the curved relationship asserted by the concept of diminishing marginal productivity.

Why do straight lines haunt economic theory, when it is forced to be logical?

The answer to this dilemma has a lot to do with one of the basic notions of economics, the belief that society is no more than the sum of its parts. This asserts that to work out the whole, all you have to do is add the parts up. This requires that the interactions between the parts are either zero or negligible. The only interaction that one variable can have with another is the one neoclassical economists want to use, simple addition: your utility plus my utility equals social utility; your output plus my output equals industry output; and so on.

This categorically rules out interactions where one variable is multiplied by another (where both are likely to be large numbers). One obvious such interaction occurs in working out a firm's revenue. A firm's revenue equals the number of units it sells, times the sale price. Economics argues that the quantity a firm will sell is a function of price – to invoke a higher supply from the firm, the price has to rise to offset the effect of diminishing marginal productivity.

If both the price and the quantity are treated as variables, then the firm's revenue is equal to one variable (price) times another (quantity). This can't be allowed if economists are to treat society as no more than the sum of its parts, so economists assume that the price a competitive firm faces is a constant. Then the firm's revenue equals a constant (price) times a variable (quantity).

However, neoclassicals then want it both ways: they want price for the entire industry to be a variable, but price for the individual firm to be a constant, without the firms interacting in any way. This just can't happen mathematically – as discussed in Chapter 4. So if they force this situation by making an invalid assumption, it inevitably means that something else that they want to treat as a variable has to instead be treated as a constant. Hence the recurrent nightmare of a straight line.

The future of mathematics in economics

There is little doubt that the close identification of neoclassical economics with mathematical analysis has given mathematics a bad name among critical economists, and members of the ordinary public who are critical of economics.

This is likely to lead to a backlash against mathematics in economics, if the discipline ever rids itself of its dominance by neoclassical economics. This would be a great pity, since, as I hope this book has shown, properly used, mathematical reasoning debunks unsound economics. Furthermore, with its limitations fully appreciated, it and computer simulations can assist in the construction of sound alternative analyses. But if mathematics is avoided for its own sake, in reaction to how economics embraced it for its own sake, then the development of a meaningful economics will be stymied.

I now turn to some of the alternatives to conventional economics that do exist – warts and all. We begin with the most radical alternative – Marxian economics. You may, if you have typecast me as 'left-wing,' expect me to praise Marxian analysis. If so, you are in for a surprise.

17 | NOTHING TO LOSE BUT THEIR MINDS

Why most Marxists are irrelevant, but most of Marx is not

Marxian economics is clearly one of the alternatives to the neoclassical way of 'thinking economically,' and by rights I should be discussing it in the next chapter, which looks at alternatives to conventional economics. However, in an illustration of the fact that conservative economists do not have a monopoly on unsound analysis, Marxian economics, as conventionally understood, is hobbled by a logical conundrum as significant as any of those afflicting neoclassical economics.

This conundrum has split non-orthodox economists into two broad camps. One tiny group continues to work within what they see as the Marxian tradition, and spends most of its time trying to solve this conundrum. The vast majority largely ignore Marx and Marxian economics, and instead develop the schools of thought discussed in the next chapter.

I find this ironic, since if Marx's philosophy is properly understood, the conundrum disappears, and Marx provides an excellent basis from which to analyze capitalism – though bereft of the revolutionary message that makes Marx both so appealing to his current followers, and anathema to so many others.

The kernel

One defining belief in conventional Marxian economics is that labor is the only source of profit: while machines are necessary for production, labor alone generates profit for the capitalist. This proposition is a key part of the radical appeal of Marxism, since it argues that capitalist profit is based upon exploitation of the worker.

Marxists argue that labor is the only source of profit because it is the only commodity where one can distinguish between 'commodity' and

‘commodity-power.’ When any other commodity is sold, the purchaser takes it lock, stock and barrel. But with labor, the capitalist ‘purchaser’ does not own the worker. Instead, he pays a subsistence wage, which can be represented by a bundle of commodities; this is the cost of production of the ability to work, which Marxists describe as the commodity ‘labor-power.’ The capitalist then puts the laborer to work for the length of the working day, during which time the worker produces a different bundle of commodities that is worth more than his subsistence wage. The difference between the output of labor and the cost of maintaining labor-power is the source of profit.

Since no such distinction can be made for machinery, the capitalist ‘gets what he paid for’ and no more when he buys a machine, whereas with labor, he gets more than he paid for. Therefore machines transfer their value only to the product.

This proposition has been shown to lead to severe logical problems, so the vast majority of critical economists have in practice abandoned Marx’s logic. However, a minority of economists continue to swear allegiance to what they perceive as Marx’s method, and continue to strive to invent ways in which the proposition that labor is the only source of profit can be maintained.

The critiques which have been made of this notion on mathematical grounds are cogent, but have been challenged by Marxian economists on philosophical or methodological grounds.

However, there are philosophical reasons why the proposition that labor is the only source of profit are invalid, and these reasons were first discovered by Marx himself. Unfortunately, Marx failed to properly understand his own logic, and instead preserved a theory that he had in fact shown to be erroneous.

Once Marx’s logic is properly applied, his economics becomes a powerful means of analyzing a market economy – though not one which argues that capitalism must necessarily give way to socialism. Unfortunately, given the ideological role of Marxism today, I expect that Marxian economists will continue to cling to an interpretation of Marx that argues for capitalism’s ultimate demise.

The roadmap

In this chapter I explain the classical economics concept of 'value,' and the manner in which Marx developed this into the labor theory of value. I illustrate the logical problems with the proposition that labor is the only source of value. I then outline Marx's brilliant philosophical analysis of the commodity, and show that this analysis contradicts the labor theory of value by arguing that all inputs to production are potential sources of value.

Marxian economics and the economics of Marx

If a nineteenth-century capitalist Machiavelli had wished to cripple the socialist intelligentsia of the twentieth century, he could have invented no more cogent weapon than the labor theory of value. Yet this theory was the invention, not of a defender of capitalism, but of its greatest critic: Karl Marx.

Marx used the labor theory of value to argue that capitalism harbored an internal contradiction, which would eventually lead to its downfall and replacement by socialism. However, Marx's logic in support of the labor theory of value had an internal contradiction that would invalidate Marx's critique of capitalism if it could not be resolved. Consequently, solving this enigma became the 'Holy Grail' for Marxist economists. Whereas nineteenth-century revolutionaries spent their time attempting to overthrow capitalism, twentieth-century revolutionaries spent theirs attempting to save the labor theory of value. Capitalism itself had no reason to fear them.

Despite valiant efforts, Marxist economists failed in their quest – and they achieved little else. As a result, while Marx's thought still has considerable influence upon philosophers, historians, sociologists and left-wing political activists, at the beginning of the twenty-first century, Marx and Marxists are largely ignored by other economists.[1] Most non-orthodox economists would acknowledge that Marx made major contributions to economic thought, but it seems that overall Samuelson was right: Marx was a 'minor Post-Ricardian' – someone who took classical economics slightly farther than had David Ricardo, but who ultimately led it into a dead end.

This conclusion is false. Properly understood, Marx's theory of value liberates classical economics from its dependence on the labor theory of value, and makes it the basis for a deep and critical understanding of capitalism. But in a truly Machiavellian irony, the main factor obscuring

this richer appreciation of Marx is the slavish devotion of Marxist economists to the labor theory of value.

To see why Marx's theory of value is not the labor theory of value, we have to first delve into the minds of the great classical economists Adam Smith and David Ricardo.

Value – a prelude

The proposition that something is the source of value raises two questions: what is 'value' anyway, and why should any one thing be the source of it?

A generic definition of value – one which encompasses the several schools of thought in economics which have used the term – is that value is the innate worth of a commodity, which determines the normal ('equilibrium') ratio at which two commodities exchange. One essential corollary of this concept is that value is unrelated to the subjective valuation which purchasers put upon a product. In what follows, I'll use 'value' in this specific sense, not in any of its more colloquial senses.

The classical economists also used the terms 'value in use' (or 'use-value') and 'value in exchange' (or 'exchange-value') to distinguish between two fundamental aspects of a commodity: its usefulness, and the effort involved in producing it. Value in use was an essential aspect of a commodity – why buy something which is useless? – but to the classical economists, it played no role in determining price.

Their concept of usefulness was also objective, focusing upon the commodity's actual function rather than how it affected the user's feelings of well-being. The use-value of a chair was not how comfortable it made you feel, but that you could sit in it.

In contrast, the neoclassical school argues that value, like beauty, is 'in the eye of the beholder' – that utility is subjective, and that the price, even in equilibrium, has to reflect the subjective value put upon the product by both the buyer and the seller. Neoclassical economics argues that the equilibrium ratio at which two products exchange is determined by the ratio of their marginal utilities to their marginal costs.

As we have seen in Chapters 3 and 4, there are serious problems with the economic theory of pricing. But it has some appeal in comparison to the

classical approach, since it seems reasonable to say that price should be determined both by the innate worth of a product, however that is defined, and by the buyer's subjective valuation of it.

The general classical reply to this concept was that, sure, in the short run and out of equilibrium, that would be true. But the classical school was more interested in 'long run' prices, and in the prices of things which could easily be reproduced.

In the long run, price would be determined by the value of the product, and not by the subjective valuations of the buyer or seller. For this reason, the classical school tended to distinguish between price and value, and to use the former when they were talking about day-to-day sales, which could be at prices which were above or below long-run values.

As well as having some influence out of equilibrium, subjective utility was the only factor that could determine the value of rare objects. As Ricardo put it:

> There are some commodities, the value of which is determined by their scarcity alone. No labor can increase the quantity of such goods, and therefore their value cannot be lowered by an increased supply. Some rare statues and pictures, scarce books and coins, wines of a peculiar quality, which can be made only from grapes grown on a particular soil, of which there is a very limited quantity, are all of this description. Their value is wholly independent of the quantity of labor originally necessary to produce them, and varies with the varying wealth and inclinations of those who are desirous to possess them. (Ricardo 1817)

Thus where scarcity was the rule, and the objects sold could not easily be reproduced, price was determined by the seller's and buyer's subjective utilities. But this minority of products was ignored by the classical economists.

Marx gave an additional explanation of why, in a developed capitalist economy, the subjective valuations of both buyer and seller would be irrelevant to the price at which commodities exchanged.

This was the historical argument that, way back in time, humans lived in small and relatively isolated communities, and exchange between them

was initially a rare and isolated event. At this stage, the objects being exchanged would be items that one community could produce but the other could not. As a result, one community would have no idea how much effort had gone into making the product, and the only basis for deciding how to exchange one product for another was the subjective valuation that each party put upon the products. As Marx put it:

> The exchange of commodities, therefore, first begins on the boundaries of such communities, at their points of contact with other similar communities, or with members of the latter. So soon, however, as products once become commodities in the external relations of a community, they also, by reaction, become so in its internal intercourse. The proportions in which they are exchangeable are at first quite a matter of chance. What makes them exchangeable is the mutual desire of their owners to alienate them. Meantime the need for foreign objects of utility gradually establishes itself. The constant repetition of exchange makes it a normal social act. In the course of time, therefore, some portion at least of the products of labor must be produced with a special view to exchange. From that moment the distinction becomes firmly established between the utility of an object for the purposes of consumption, and its utility for the purposes of exchange. Its use-value becomes distinguished from its exchange-value. On the other hand, the quantitative proportion in which the articles are exchangeable, becomes dependent on their production itself. (Marx 1867)

The most famous example of two products being exchanged on the basis of the perceived utility rather than their underlying value is the alleged exchange of the island of Manhattan for a bunch of beads.[2] This price would never have been set if trade between the Dutch and the Indians had been a long-established practice, or if the Indians knew how little work it took to produce the beads.

In an advanced capitalist nation, factories churn out mass quantities of products specifically for exchange – the seller has no interest in the products his factory produces. The sale price reflects the cost of production, and the subjective utility of the buyer and seller are irrelevant to the price.[3]

There is thus at least a prima facie plausibility to the argument that value alone determines the equilibrium ratio at which commodities are exchanged. The problem comes with the second question: what is the source of value?

Physiocrats

The first economists to systematically consider this question[4] answered that the source of all value was land.

The argument, in a nutshell, was that land existed before man did. Therefore man – or more specifically, man's labor – could not be the source of value. Instead, value came from the land as it absorbed the energy falling on it from the sun. Man's labor simply took the naturally generated wealth of the land and changed it into a different form. Land generated a surplus, or net product, and this enabled both growth and discretionary spending to occur.

Manufacturing, on the other hand, was 'sterile': it simply took whatever value the land had given, and transformed it into different commodities of an equivalent value. No formal proof was given of this latter proposition, beyond an appeal to observation:

> Maxims of Economic Government. I: Industrial work does not increase wealth. Agricultural work compensates for the costs involved, pays for the manual labor employed in cultivation, provides gains for the husbandmen, and, in addition, produces the revenue of landed property. Those who buy industrial goods pay the costs, the manual labor, and the gain accruing to the merchants; but these goods do not produce any revenue over and above this. Thus all the expenses involved in making industrial goods are simply drawn from the revenue of landed property – no increase of wealth occurs in the production of industrial goods, since the value of these goods increases only by the cost of the subsistence which the workers consume. (Quesnay, cited in Meek 1972)

Since land determined the value of commodities, and the price paid for something was normally equivalent to its value, the ratio between the prices

of two commodities should be equivalent to the ratios of the land needed to produce them.

Smith (and Ricardo)

The physiocratic answer to the source of value reflected the school's origins in overwhelmingly rural France. Adam Smith, a son of Scotland and neighbor to the 'nation of shopkeepers,' was strongly influenced by the physiocrats. But in *The Wealth of Nations* (which was published in the year in which the first steam engine was installed) Smith argued that labor was the source of value. In Smith's words: 'The annual labor of every nation is the fund which originally supplies it with all the necessaries and conveniences of life which it annually consumes, and which consist always either in the immediate produce of that labor, or in what is purchased with that produce from other nations' (Smith 1838 [1776]).

The growth of wealth was due to the division of labor, which increased because the expansion of industry allowed each job to be divided into ever smaller specialized sub-tasks. This allowed what we would today call economies of scale: an increase in the size of the market allowed each firm to make work more specialized, thus lowering production costs (his most famous example of this was of a pin factory; this passage, which is better known than it is read, is reproduced on the web at Marx/More).[5]

Smith therefore had an explanation for the enormous growth in output which occurred during the Industrial Revolution. However, he had a dilemma: for reasons discussed below, Smith knew that, though labor was the source of value, it could not possibly determine price. Yet value was supposed to determine the ratio at which two commodities exchanged.

The dilemma arose because two commodities could exchange only on the basis of the amount of direct labor involved in their manufacture if only labor was required for their production. Smith gave the example of exchange in a primitive hunting society:

> In that early and rude state of society which precedes both the accumulation of stock and the appropriation of land, the proportion between the quantities of labor necessary for acquiring different objects seems to be the only circumstance which can afford any rule for exchanging them for one another. If among a nation of hunters, for

> example, it usually costs twice the labor to kill a beaver which it does to kill a deer, one beaver should naturally exchange for or be worth two deer. (Ibid.)

However, once there had been an 'accumulation of stock' – once a market economy had evolved – then paying for the labor alone was not sufficient; the price had also to cover profit:

> As soon as stock has accumulated in the hands of particular persons, some of them will naturally employ it in setting to work industrious people, whom they will supply with materials and subsistence, in order to make a profit by the sale of their work, or by what their labor adds to the value of the materials. In exchanging the complete manufacture either for money, for labor, or for other goods, over and above what may be sufficient to pay the price of the materials, and the wages of the workmen, something must be given for the profits of the undertaker of the work who hazards his stock in this adventure. (Ibid.)

So Smith was forced to concede that the price had to be high enough to pay for not just the hours of labor involved in making something, but also a profit. For example, if the deer hunter was an employee of a deer-hunting firm, then the price of the deer had to cover the hunter's labor, and also a profit margin for the firm.

The problem became more complicated still when land was involved. Now the price had to cover labor, profit, and rent. Smith's statement of this reveals that this 'father of economics' was rather more cynical and critical of market relations than some of his descendants: 'As soon as the land of any country has all become private property, the landlords, like all other men, love to reap where they never sowed, and demand a rent even for its natural produce' (ibid.).

In the end, Smith was reduced to an 'adding up' theory of prices: the price of a commodity represented in part payment for labor, in part payment for profit, and in part payment for rent. There was therefore no strict relationship between value and price.

Ricardo Though he paid homage to his predecessor, Ricardo was, to say the least, critical of Smith's treatment of the relationship between value and price. He began his *Principles of Political Economy and Taxation* (Ricardo 1817) with an emphatic statement of the belief he shared with Smith, that labor was the determinant of the value of a commodity: 'The value of a commodity, or the quantity of any other commodity for which it will exchange, depends on the relative quantity of labor which is necessary for its production' (ibid.). However, he was much more aware than Smith of the need for precise definitions, and of the difficulties in going from value to price.

Smith had used two measures of the amount of labor contained in a product: 'labor embodied' and 'labor commanded.' Labor embodied was the amount of direct labor time it actually took to make a commodity. Labor commanded, on the other hand, was the amount of labor-time you could buy using that commodity.

If, for example, it took one day for a laborer to make a chair, then the chair embodied one day's labor. However, that chair could well sell for an amount equivalent to two days' wages – with the difference accounted for by profit and rent. The chair would therefore command two days' labor.

Ricardo argued that the former measure was far less volatile than the latter. He believed, in common with most classical economists, that workers received a subsistence wage. Since this would always be equivalent to a fairly basic set of commodities – so much food, clothing, and rental accommodation – it would not change much from one year to the next. The latter measure, however, reflected the profit earned by selling the worker's output, and this would vary enormously over the trade cycle.

His solution for the value/price dilemma was that the price of a commodity included not just direct labor, but also the labor involved in producing any tools. Ricardo took up Smith's deer and beaver example and elaborated upon it. Even in Smith's example, some equipment had to be used to kill the game, and variations in the amount of time it took to make the equipment would affect the ratio in which deer and beavers were exchanged:

> Even in that early state to which Adam Smith refers, some capital, though possibly made and accumulated by the hunter himself,

> would be necessary to enable him to kill his game. Without some weapon, neither the beaver nor the deer could be destroyed, and therefore the value of these animals would be regulated, not solely by the time and labor necessary to their destruction, but also by the time and labor necessary for providing the hunter's capital, the weapon, by the aid of which their destruction was effected.
>
> Suppose the weapon necessary to kill the beaver was constructed with much more labor than that necessary to kill the deer, on account of the greater difficulty of approaching near to the former animal, and the consequent necessity of its being more true to its mark; one beaver would naturally be of more value than two deer, and precisely for this reason, that more labor would, on the whole, be necessary to its destruction. (Ibid.)

Thus the price of any commodity reflected the labor which had been involved in creating it, and the labor involved in creating any means of production used up in its manufacture. Ricardo gave many numerical examples in which the labor involved in producing the means of production simply reappeared in the product, whereas direct labor added additional value over and above its means of subsistence – because of the difference between labor embodied (which equaled a subsistence wage) and labor commanded (which included a profit for the capitalist).[6]

However, Smith and Ricardo were both vague and inconsistent on key aspects of the theory of value.

Though he generally argued that labor was the source of value, on several occasions Smith counted the work of farm animals as labor.[7] Though he failed to account for the role of machinery in the creation of value, he also argued that machines could produce more value than it took to produce them – which would mean that machinery (and animals) would be a source of value, in addition to labor: 'The expense which is properly laid out upon a fixed capital of any kind, is always repaid with great profit, and increases the annual produce by a much greater value than that of the support which such improvements require' (Smith 1838 [1776]).

Ricardo more consistently implied that a machine added no more value to output than it lost in depreciation, but he also occasionally lapsed into completely ignoring the contribution of machinery to value.[8]

Marx's labor theory of value

Where his forebears implied and were vague, Marx stated and was emphatic: labor was the only source of value, in the sense that it could add 'more value than it has itself' (Marx 1867). Marx called this difference between the value embodied in a worker and the value the worker added to production 'surplus value,' and saw it as the sole source of profit.

He was critical of Ricardo for not providing an explanation of why this difference existed – in Ricardo's terms, for not having a systematic explanation of why labor embodied differed from labor commanded. As Marx put it:

Ricardo starts out from the actual fact of capitalist production. The value of labor is smaller than the value of the product which it creates – The excess of the value of the product over the value of the wages is the surplus-value – For him, it is a fact, that the value of the product is greater than the value of the wages. How this fact arises, remains unclear. The total working-day is greater than that part of the working-day which is required for the production of wages. Why? That does not emerge. (Marx 1968 [1861]: Part II)

The best that Ricardo could offer, Marx claimed, was that:

> [t]he value of labor is therefore determined by the means of subsistence which, in a given society, are traditionally necessary for the maintenance and reproduction of the laborers.
>
> But why? By what law is the value of labor determined in this way?
>
> Ricardo has in fact no answer, other than – the law of supply and demand – He determines value here, in one of the basic propositions of the whole system, by demand and supply – as Say notes with malicious pleasure. (Ibid.)

Similarly, Marx rejected Smith's musings on the productivity of machinery, and concurred with Ricardo that a machine only added as much value to output as it lost through depreciation:

The maximum loss of value that they can suffer in the process, is plainly limited by the amount of the original value with which they came into the process, or in other words, by the labor-time necessary for their production. Therefore, the means of production can never add more value to the product than they themselves possess independently of the process in which they assist. However useful a given kind of raw material, or a machine, or other means of production may be, though it may cost £150, or, say 500 days' labor, yet it cannot, under any circumstances, add to the value of the product more than £150. (Marx 1867)

Marx likewise concurred with Ricardo's definition of value, cited above, that it 'depends on the relative quantity of labor which is necessary for its production.' Value in turn determined the price at which commodities exchanged, with commodities of an equivalent value – commodities containing an equivalent quantity of labor[9] – exchanging for the same price (in equilibrium).

This exchange of equivalents nonetheless still had to enable capitalists to make a profit, and Marx was disparaging of any explanation of profits which was based on 'buying cheap and selling dear':

> To explain, therefore, the general nature of profits, you must start from the theorem that, on the average, commodities are sold at their real values, and that profits are derived by selling them at their values, that is, in proportion to the quantity of labor realized in them. If you cannot explain profit upon this supposition, you cannot explain it at all. (Marx 1847)

Marx gave two explanations for the origin of surplus value. One was a 'negative' proof, by a process of elimination based on the unique characteristics of labor. The other was a 'positive' proof, based on a general theory of commodities. Most Marxist economists are aware of only the negative proof.

The origin of surplus value (I)

This was that labor was a unique commodity, in that what was sold was not actually the worker himself (which would of course be slavery), but his capacity to work, which Marx called labor-power. The value (or cost of production) of labor-power was the means of subsistence, since that is what it took to reproduce labor-power. It might take, say, six hours of labor to produce the goods which are needed to keep a worker alive for one day.

However, what the capitalist actually received from the worker, in return for paying for his labor-power, was not the worker's capacity to work (labor-power), but actual work itself. If the working day was twelve hours long (as it was in Marx's day), then the worker worked for twelve hours – twice as long as it actually took to produce his value. The additional six hours of work was surplus labor, which accrued to the capitalist and was the basis of profit. As Marx put it:

> The laborer receives means of subsistence in exchange for his labor-power; the capitalist receives, in exchange for his means of subsistence, labor, the productive activity of the laborer, the creative force by which the worker not only replaces what he consumes, but also gives to the accumulated labor a greater value than it previously possessed. (Ibid.)

This difference between labor and labor-power was unique to labor: there was no other commodity where 'commodity' and 'commodity-power' could be distinguished. Therefore other commodities used up in production simply transferred their value to the product, whereas labor was the source of additional value. Surplus value, when successfully converted into money by the sale of commodities produced by the worker, was in turn the source of profit.

The labor theory of value and the demise of capitalism This direct causal relationship between surplus value and profit meant there was also a direct causal relationship between what Marx called the rate of surplus-value and the rate of profit.

The rate of surplus value was the ratio of the surplus labor-time performed by a worker to the time needed to reproduce the value of labor-power. In our example above, this ratio is 1 to 1, or 100 percent: six hours of surplus labor to six hours of what Marx called necessary labor.

Marx defined the rate of profit as the ratio of surplus (which he denoted by the symbol *s*) to the sum of the inputs needed to generate the surplus. Two types of inputs were needed: necessary labor, and the means of production (depreciation of fixed capital plus raw materials, intermediate goods, etc.). Marx called necessary labor variable capital (for which he used the symbol *v*), because it could increase value, and he called the means of production used up constant capital (for which he used the symbol *c*), because it could not increase value.

Taking the example of weaving which Marx used extensively, during one working day a weaver might use 1,000 yards of yarn and wear out one spindle. The yarn might have taken twelve hours of (direct and indirect) labor to make, and the spindle the same. Thus the sum of the direct labor-time of the worker, plus the labor-time embodied in the yarn and the spindle, is thirty-six hours: twelve hours' labor by the weaver, twelve for the yarn, and twelve for the spindle. The ratio of the surplus to $c + v$ is 6:30 for a rate of profit of 20 percent.

Marx assumed that the rate of surplus value – the ratio of *s* to *v* – was constant, both across industries and across time.[10] Simultaneously, he argued that the competitive forces of capitalism would lead to capitalists replacing direct labor with machinery, so that for any given production process, *c* would get bigger with time. With *s*/*v* constant, this would decrease the ratio of *s* to the sum of *c* and *v*, thus reducing the rate of profit.

Capitalists would thus find that, regardless of their best efforts, the rate of profit was falling.[11] Capitalists would respond to this by trying to drive down the wage rate, which would lead to revolt by the politically aware working class, thus leading to a socialist revolution.[12]

Well, it was a nice theory. The problem was that, even if you accepted the premise that labor was the only source of value, the theory still had major logical problems. Chief among these was what became known as the transformation problem.

The transformation problem The transformation problem arises from the fact that capitalists are motivated not by the rate of surplus value, but by the rate of profit. If the rate of surplus value is constant across industries, and labor is the only source of surplus, then industries with a higher than average ratio of labor to capital should have a higher rate of profit. Yet if a

capitalist economy is competitive, this situation cannot apply in equilibrium, because higher rates of profit in labor-intensive industries should lead to firms moving out of capital-intensive industries into labor-intensive ones, in search of a higher rate of profit.

Marx was not an equilibrium theorist, but this problem was serious because his description of equilibrium was inconsistent. Somehow, he had to reconcile a constant rate of surplus value across industries with at least a tendency towards uniform rates of profit.

Marx's solution was to argue that capitalism was effectively a joint enterprise, so that capitalists earned a profit which was proportional to their investment, regardless of whether they invested in a labor-intensive or capital-intensive industry:

> Thus, although in selling their commodities the capitalists of the various spheres of production recover the value of the capital consumed in their production, they do not secure the surplus-value, and consequently the profit, created in their own sphere by the production of these commodities – So far as profits are concerned, the various capitalists are just so many stockholders in a stock company in which the shares of profit are uniformly divided per 100. (Marx 1894)

He provided a numerical example (ibid.) that purported to show that this was feasible. He first provided a table (Table 17.1) showing the production of surplus value by a number of industries with differing ratios of variable to constant capital (in modern terms, varying labor-to-capital ratios).

In this 'value' table, a higher ratio of labor to capital is associated with a higher rate of profit. Thus 'labor-intensive' industry III, with a labor-to-capital ratio of 2:3, earns the highest 'value' rate of profit of 40 percent, while 'capital-intensive' industry V, with a 1:20 ratio, makes a 'value' rate of profit of just 5 percent.

Then Marx provided a second table in which the same industries earned a uniform rate of profit, now in terms of price rather than value. In contrast to Table 17.1, now all industries earned the same rate of profit.

The numbers in this example appeared feasible. The sums are consistent: the sum of all prices in Table 17.2 equals the sum of the value created in Table 17.1; the sum of surplus value in Table 17.1 equals the sum

of the differences between input costs (500) and the price of output in Table 17.2 (610). But this apparent consistency masks numerous internal inconsistencies. The best proof of this was provided by the Sraffian economist Ian Steedman (this next section is unavoidably technical, and can be skipped at first reading).

TABLE 17.1 Marx's unadjusted value creation table, with the rate of profit dependent upon the variable-to-constant ratio in each sector

Capitals	Constant capital	Variable capital	Rate of surplus value (%)	Surplus value	Product	Rate of profit (%)
I	80	20	100	20	120	20
II	70	30	100	30	130	30
III	60	40	100	40	140	40
IV	85	15	100	15	115	15
V	95	5	100	5	105	5
Sum	390	110	100	110	610	22

TABLE 17.2 Marx's profit distribution table, with the rate of profit now uniform across sector

Capitals	Constant capital	Variable capital	Surplus value	Price	Rate of profit	Deviation
I	80	20	22	122	22%	2%
II	70	30	22	122	22%	−8%
III	60	40	22	122	22%	−18%
IV	85	15	22	122	22%	7%
V	95	5	22	122	22%	17%
Sum	390	110	110	610	22%	

Marxist economics after Sraffa

We have already seen in Chapter 6 the damage Sraffa's crucible did to the economic theory of price determination and income distribution. In an illustration of the comparatively non-ideological nature of Sraffian analysis,

Steedman showed that Sraffa's method could equally well critique Marxian economics.

The basis of Sraffa's system is the acknowledgment that commodities are produced using other commodities and labor. Unlike conventional economics – which has invented the fictional abstraction of 'factors of production' – Marx's system is consistent with Sraffa's 'production of commodities by means of commodities' analysis (indeed, Marx's economics was a major inspiration for Sraffa).

Steedman began with an illustrative numerical model of an economy with just three commodities: iron, corn and gold. Iron and labor were needed to produce all three commodities, but neither gold nor corn was needed to produce anything.[13] Table 17.3 shows the quantities of inputs and outputs in Steedman's hypothetical economy.

TABLE 17.3 Steedman's hypothetical economy

Industries	Inputs		Outputs		
	Iron	Labour	Iron	Gold	Corn
Iron	28	56	56		
Gold	16	16		48	
Corn	12	8			8
Totals	56	80	56	48	8

The numbers in this table represent arbitrary units: the iron units could be tons, the labor units hours, gold units ounces, and corn units bushels – and any other set of arbitrary units would do as well. However, since each input is measured in a completely different unit, the numbers add up only down the columns: they don't add across the rows.

To analyze the labor theory of value, Steedman first had to convert these into the 'labor-value' units which Marx used. For simplicity, he set the labor-value of one unit ('hour') of labor at 1. Converted into value terms, Table 17.4 then says that it takes 28 times whatever the 'labor-value' of a ton of iron is, plus 56, to produce 56 times whatever the 'labor-value' of a ton of iron is. A bit of simple algebra shows that one ton of iron has a labor-value of 2.

TABLE 17.4 Steedman's physical table in Marx's value terms

	c	v	s	Totals
Iron	56	14	42	112
Gold	32	4	12	48
Corn	24	2	6	32
Totals	112	20	60	192

Similar calculations show that the labor-value of an ounce of gold is 1, and the labor-value of a bushel of corn is 4.

The next stage in the analysis is to work out the value of the commodity labor-power. It might appear that this has already been done – didn't he set this equal to 1? No, because this represents the total amount of labor performed, and in Marx's theory, workers get paid less than this. They get paid, not for their contribution to output, but for the commodity labor-power, whose value is equal to the means of subsistence.

Steedman assumed that it took five bushels of corn to reproduce the labor used in this hypothetical economy. Therefore the total value of labor-power in the entire economy was equal to the labor-value of five bushels of wheat. Since a bushel of wheat has a labor-value of 4, this means that the value of labor-power across the whole economy was 20 (and therefore, one unit of labor had a labor-value of 1/4). The difference between this amount and the total labor performed – 80 hours of labor, which we have set to equal 80 units of labor-value – is surplus value. So *v*, in Marx's scheme, is 20, while *s* is 60, for a rate of surplus value of 300 percent.

These numbers now allow the physical input data in Table 31 to be converted into Marx's labor-value terms. Since Marx assumed that the rate of surplus value was the same across all industries, ¼ of the labor input in each industry represents *v*, while ¾ represents *s*. Taking the iron industry, of the 56 labor-value units of direct labor, 14 represent *v* and 42 represent *s*. Since Table 17.5 is now in consistent units (everything is measured in labor-value units), the table adds up both horizontally and vertically.

With this table constructed, we can now calculate the average rate of profit in Marx's terms – which is the ratio of total *s* to the sum of *c* and *v*, or 60/132 (this factors to 5/11, and is equal to a rate of profit of 455Ú11 percent). In equilibrium, this rate of profit will apply across all industries, since otherwise capitalists would be shifting their resources from one sector

to another. Steedman then multiplied the input values by 1 plus this uniform rate of profit to yield Marx's 'transformation' of values into prices.

TABLE 17.5 Steedman's prices table in Marx's terms

	Inputs c	v	Total	Profit rate	Markup (%)	Total price	Per unit price
Iron price of production	56	14	70	45	31.82	101.82	1.82
Gold price of production	32	4	36	45	16.36	52.36	1.09
Corn price of production	24	2	26	45	11.82	37.82	4.73
	112	20	132		60	192	

So far, so good. Just as with Marx's table, the sum of values equals the sum of prices, and the sum of profits equals the sum of surplus values. However, all is not as well as it seems.

Table 17.5 tells us that the price of the total output of the iron industry is 101.82 (let's call this dollars, even though in these models the price simply means the ratios in which commodities exchange). If we divide this by the physical output of 56 tons of iron, then this means the price per ton is $1.82. If the iron industry pays this price for its iron inputs in the next period, it will pay out $50.91. To hire the workers it needs, it has to buy sufficient corn: the amount works out to 3.5 bushels (this is the total amount of corn consumed by all workers – 5 bushels – multiplied by the fraction of the total workforce employed in the iron industry). This costs $16.55. The iron industry's total outlays are thus $67.46, and yet (if Marx's equilibrium price calculations are accurate), it can sell its output for $101.82, for a profit of $34.36. But this is $2.55 more than the profit in the previous period.

Clearly there is an inconsistency – or rather, at least one. The simplest is that Marx converted the output into price terms, but didn't convert the inputs. However, it's worse than this: even if you amend this error, you get nonsense results: what is supposed to be an equilibrium (and therefore stationary) turns out not to be stationary at all.

Steedman then shows that you don't have to 'transform' physical quantities into values, and values into prices: you can instead derive prices

directly from the physical data and the equilibrium assumption of a uniform rate of profit. The basis of this is that, in equilibrium, the prices have to enable each sector to just pay for its inputs and make the average rate of profit. Thus for the iron industry, the price of its 28 tons of iron inputs, plus the price of its 56 hours of labor, plus the standard markup, must just equal the price of its 56 tons of iron output. There are two similar equations for corn and gold, and one final relation linking the wage to the cost of the subsistence amount of corn. If the gold price is notionally set to $1, this yields the average rate of profit, wage, and prices of iron and corn (in terms of gold) shown in Table 17.6.

TABLE 17.6 Profit rate and prices calculated directly from output/wage data

Variable	Value; price in terms of gold
Rate of profit	52%
Iron price of production	1.71
Gold price of production	1
Corn price of production	4.3

Things don't look so good for Marx's tables now. First, the rate of profit and prices worked out directly from the data (in Table 17.6) differ from those derived by taking Marx's route through the concept of value (in Table 17.5). Worse, whereas Marx's numbers aren't consistent – they are supposed to describe an equilibrium situation, but don't – the numbers derived directly from the data are consistent.

Take iron, for example. The iron sector pays $1.71 per ton for its 28 tons of inputs, for a total of $47.88. It buys 3.5 bushels of wheat for $4.3 a bushel, for an outlay of $15.05. Total expenses of production are therefore $62.93. It then marks this up by the rate of profit to a total of $95.65. Except for the effect of rounding error, this equals the price of iron ($1.71) times the output (56 tons).[14]

Steedman concluded that, far from value determining prices, prices could not be accurately derived from values. Instead, prices could be worked out directly from the physical production data, and knowledge of the real wage: value calculations were both superfluous and misleading. He concluded that

[t]here is no problem of transforming values into prices, etc., to be solved. The 'transformation problem' is a 'non-problem,' a spurious problem which can only be thought to arise and to have significance when one is under the misapprehension that the rate of profit must be determined in terms of labor quantities. Once it is seen that there is no such necessity, the 'problem' simply evaporates. (Steedman 1977)

Though he did not put his conclusion in this way, Steedman was essentially saying that Marx cannot be right that labor is the only source of surplus. We are better off to forget the whole question of 'where does the surplus come from?' and instead simply accept that it exists, and analyze capitalism on that basis.

The inconsistencies Steedman establishes[15] undermined Marx's sequence of claims that labor is the only source of value, that value is the only source of profits, and that value determines price. Marx could also provide no reason why capitalism, possible the most internally competitive social system ever, should ultimately behave so cooperatively, with capitalists sharing in total social profit as 'just so many stockholders in a stock company in which the shares of profit are uniformly divided per 100.'

Thus, though Marx used the labor theory of value to both attack capitalism and predict its downfall, the theory did not even seem to provide a consistent model of capitalism itself – let alone a 'scientific' explanation of why capitalism would wilt and socialism blossom. It appeared that the great revolutionary challenger to capitalism had promised a bang, but delivered a whimper.

The Marxist response This was no great disappointment to his conservative critics, who happily pointed out the flaws in Marx's logic, and turned to developing economics as we know it today. But devoted Marxists tried valiantly to resurrect Marx's program of 'scientific socialism' by showing that, somehow, at some deep level, Marx's theory of value was internally consistent.

Many years before Steedman turned Sraffa's blowtorch onto Marx's economics, leading Marxist economists had applauded Sraffa's methodical critique of neoclassical economics. However, some of them could also see that Sraffa's dispassionate analysis posed serious problems for the labor theory of value. One of the most thoughtful of such responses came from

Ronald Meek in his scholarly *Studies in the Labor Theory of Value*. In a section headed 'From values to prices: was Marx's journey really necessary?,' Ronald Meek asked:

> Why did he think that anything had to be 'transformed' in order to arrive at the equilibrium prices characteristic of competitive capitalism? And if something did have to be 'transformed' in order to arrive at them, why did it have to be these mysterious, non-observable, Volume I 'values'? Personally, although I am no longer at all religious about such matters, I find myself leaning much more towards the 'neo-Ricardians' than towards their critics. I think that it is useful to talk in terms of a broad Ricardo-Marx-Sraffa tradition or stream of thought, in which the question of the relation between the social surplus and the rate of profit has always been (and still is) a central theme. (Meek 1972)

In other words, Meek was prepared to abandon the emphasis upon value, and instead develop Marx's analysis of capitalism – minus the insistence that labor is the only source of value, and that value determines profit and prices. Many other scholars followed Meek's lead, and abandoned strict Marxist economics, with its insistence upon value analysis.

However, a minority has persisted, and continue to argue that, somehow, value is an essential part of Marxist analysis. This minority's response to Steedman's critique is best summarized in the title of a paper by Anwar Shaikh: 'Neo-Ricardian economics: a wealth of algebra, a poverty of theory' (Shaikh 1982).

The implication is that, somehow, Marx's philosophy sidesteps the mathematical problems highlighted by Steedman, or it points out a step in the mathematical chain which Steedman missed. To date, no Marxist has been able to put forward an explanation of this rejoinder, which has commanded assent from the majority of Marxists: there are almost as many competing ways to try to avoid Steedman's critique as there are Marxist economists. However, they all assent that there is something in Marx's philosophy which counteracts Steedman's mathematical attack.

Over one century after Marx's flawed solution to the transformation problem was first published, and almost a quarter of a century after

Steedman's devastating critique, they are still at it. The latest attempts argue that, since Marx's theory was actually dynamic rather than static, the transformation problem should be solvable in a dynamic model.

Nice try, guys, but you really shouldn't bother. The labor theory of value is internally inconsistent, and perhaps even more flawed than conventional economic theory itself. And far from philosophy saving the labor theory of value from mathematical criticism, philosophy provides further compelling reasons for its rejection. One convincing proof of this was given by the Indian economist Arun Bose.

Arun Bose: Marx's 'capital axioms'

Bose was well aware of the criticism leveled at Steedman that his argument, while mathematically impeccable, was somehow philosophically lacking. Though he disparagingly referred to this as 'a theological tendency to go so strictly by what Marx said as to adhere to the rule: "where logic contradicts Marx's words, go by his words"' (Bose 1980), Bose nonetheless tried to avoid this judgment by looking for textual support in Marx. He called his interpretation of Marx the 'capital theory' approach, and argued that: 'as far as logic goes, there are "two Marxes," the Marx of the "labor value" approach, and the Marx of the "capital theory" approach,' and that the 'second Marx' should be supported in preference to the first (in scientific discussion) (ibid.).[16]

Bose, unlike Steedman, accepted the Marxian position that the concept of value was somehow essential. However, what he argued was that, if value was in some sense the essence of a commodity, then that essence could not be reduced solely to labor. Therefore labor alone was not the essence of value: instead, both labor and commodities were the essence of value. As Bose put it: 'labor is never the only or the main "source of value" in any system which is defined as capitalist on the basis of a reasonable set of axioms. Labor is not, immediately or ultimately, the only or main source of price, surplus or profit. Labor and commodities are the two sources of wealth, value, price, of surplus value and profit' (ibid.). His logic used a concept we saw earlier in Chapter 6: the reduction of commodity inputs to dated labor.[17]

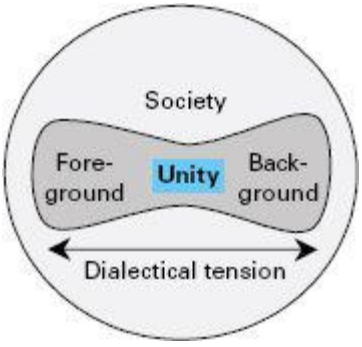
The manufacture of any commodity requires direct labor, machinery, intermediate goods, and raw materials. All the non-labor inputs had to have been produced at some time in the past: even unprocessed raw materials had to have been previously either mined or harvested. They in turn were made using some direct labor, and other commodity inputs (machinery, intermediate goods, raw materials). These again can be reduced to even earlier dated labor, and other commodity inputs.[18]

This process can go on indefinitely, with each step further reducing the commodity content. But no matter how far back you go, you can never eliminate this commodity residue. If you could, then there would be some commodities that can be created with absolutely no commodity inputs – or in other words, by magic.[19] Therefore if value is the essence of a commodity, then that essence consists of both labor and commodities – it cannot be derived solely from labor.

Bose's conclusion probably helped sway some more Marxists to abandon the faith. But generally, his argument was simply not acknowledged by Marxist economists. A similar fate has to date befallen the next argument, which establishes that the labor theory of value is inconsistent, not just with mathematical logic, or with any reductionist notion of the commodity, but with Marx's own philosophy.

The origin of surplus value (II)

As noted earlier, most Marxists believe that Marx reached the conclusion that labor was the source of value by a 'negative' proof, which eliminated any other possible contenders. This was true up until 1857, when he developed an alternative, and far superior, 'positive' proof. To understand this proof, we have to delve into Marx's 'dialectical' philosophy.[20]



17.1 A graphical representation of Marx's dialectics

In brief, dialectics is a philosophy of change. It begins from the proposition that any entity exists in a social environment (see Figure 17.1). The environment will emphasize some aspect of the entity, and necessarily places less emphasis upon all other aspects of the entity. However, the entity cannot exist without both the foreground aspects (the features the environment emphasizes) and background aspects (the ones it neglects). This sets up a tension within the entity, and possibly between the entity and the environment. This tension can transform the nature of the entity, and even the environment itself.

Marx first applied this logic to the concept of the commodity in 1857. He reasoned that the commodity was the unity of use-value and exchange-value. In a capitalist economy, the exchange-value of a commodity is brought to the foreground[21] while its use-value is pushed into the background. What this means in practice is that the use-value of a commodity is irrelevant to its price: its price is instead determined by its exchange-value. Yet the commodity can't exist without its use-value (something useless can't be a commodity), so that a dynamic tension is set up between use-value and exchange-value in capitalism.

Prior to this realization, Marx had concurred with Smith and Ricardo that use-value was irrelevant to economics. After it, the concept of use-value, in unison with exchange-value, became a unifying concept for his whole analysis of capitalism.

Marx's first exploration of this concept occurred when he was working on the 'rough draft' of *Capital* in 1857: 'Is not value to be conceived as the unity of use-value and exchange-value? In and for itself, is value as such the general form, in opposition to use-value and exchange-value as particular forms of it? Does this have significance in economics?' (Marx 1857).[22]

The manner in which he first puts the proposition, as questions to himself rather than didactic statements, and especially his comment 'Does this have significance in economics?', shows how novel the concept was to him. From this point on, Marx exclusively used this positive methodology, based on a general axiomatic analysis of the commodity, to explain the

source of surplus value. Since this point is appreciated by so few Marxists, it is worth citing several of Marx's many pronouncements on this issue.

I noted earlier that Marx mocked Ricardo for not having an explanation of why labor embodied differed from labor commanded. He notes that Smith fell for the fallacy that, under capitalism, a worker should be paid his full product. He continues:

Ricardo, by contrast, avoids this fallacy, but how? 'The value of labor, and the quantity of commodities which a specific quantity of labor can buy, are not identical.' Why not? 'Because the worker's product is not = to the worker's pay.' I.e. the identity does not exist, because a difference exists – Value of labor is not identical with wages of labor. Because they are different. Therefore they are not identical. This is a strange logic. There is basically no reason for this other than it is not so in practice. (Ibid.)

Marx then contrasts his easy ability to derive the source of surplus value with Ricardo's struggles to do the same: 'What the capitalist acquires through exchange is labor capacity; this is the exchange value which he pays for. Living labor is the use-value which this exchange value has for him, and out of this use-value springs the surplus value and the suspension of exchange as such' (ibid.).

There are many similar such statements, many of which were written in documents which were either not intended for publication or were never formally completed by Marx. But even in the most well-known passage where Marx derives the source of surplus value, in *Capital* I, this positive derivation takes precedence over the negative proof.

Marx began *Capital* by clearing intellectual cobwebs en route to uncovering the source of surplus, criticizing explanations based upon unequal exchange or increasing utility through exchange. He then restated the classical axiom that exchange involves the transfer of equivalents, and the conclusion that therefore exchange of itself cannot provide the answer. Yet at the same time circulation based on the exchange of equivalents must be the starting point from which the source of surplus value is deduced. Marx put the dilemma superbly:

The conversion of money into capital has to be explained on the basis of the laws that regulate the exchange of commodities, in such a way that the starting point is the exchange of equivalents. Our friend, Moneybags, who as yet is only an embryo capitalist, must buy his commodities at their value,

must sell them at their value, and yet at the end of the process must withdraw more value from circulation than he threw into it at starting. His development into a full-grown capitalist must take place, both within the sphere of circulation and without it. These are the conditions of the problem. (Marx 1867)

He began the solution of this dilemma with a direct and powerful application of the dialectic of the commodity. If the exchange-value of the commodity cannot be the source of surplus, then the dialectical opposite of value, use-value, is the only possible source:

The change of value that occurs in the case of money intended to be converted into capital must take place in the commodity bought by the first act, M-C, but not in its value, for equivalents are exchanged, and the commodity is paid for at its full value. We are, therefore, forced to the conclusion that the change originates in the use-value, as such, of the commodity, i.e. its consumption. In order to be able to extract value from the consumption of a commodity, our friend, Moneybags, must be so lucky as to find, within the sphere of circulation, in the market, a commodity, whose use-value possesses the peculiar property of being a source of value. (Ibid.)

Marx then used the quantitative difference between the exchange-value of labor-power and its use-value to uncover the source of surplus value in the transaction between worker and capitalist:

The past labor that is embodied in the labor power, and the living labor that it can call into action; the daily cost of maintaining it, and its daily expenditure in work, are two totally different things. The former determines the exchange-value of the labor power, the latter is its use-value. The fact that half a [working] day's labor is necessary to keep the laborer alive during 24 hours, does not in any way prevent him from working a whole day. Therefore, the value of labor power, and the value which that labor power creates in the labor process, are two entirely different magnitudes; and this difference of the two values was what the capitalist had in view, when he was purchasing the labor power. What really influenced him was the specific use-value which this commodity possesses of being a source not only of value, but of more value than it has itself. This is the special service that the capitalist expects from labor power, and in this transaction he acts in accordance with the 'eternal laws' of the exchange of commodities. The

seller of labor power, like the seller of any other commodity, realizes its exchange-value, and parts with its use-value. (Ibid.)

The one way in which Marx's 'negative' derivation survived was in the claim that labor-power was the only commodity with the property of being 'a source not only of value, but of more value than it has itself.' In *Capital* I, Marx appeared to successfully reach the conclusion that the means of production could not be a source of surplus value. However, he did so by contradicting a basic premise of his 'positive' proof, that the use-value and the exchange-value of a commodity are unrelated. Properly applied, his 'positive proof' contradicts the negative one by showing that all inputs to production are potential sources of surplus-value.

'Guilty of this or that inconsistency because of this or that compromise' In the course of his attempt to preserve the labor theory of value proposition that labor-power is the only source of surplus value, Marx advanced three propositions which fundamentally contravene his general approach to commodities: that, in the case of the means of production, the purchaser makes use of their exchange-value, not their use-value; that their use-value cannot exceed their exchange-value; and that the use-value of commodity inputs to production somehow reappears in the use-value of the commodities they help create.

Marx began with the simple assertion that the means of production can transfer no more than their exchange-value to the product. He next attempted to forge an equality between the exchange-value and the use-value of the means of production, by equating the depreciation of a machine to its productive capacity.

> Value exists only in articles of utility. If therefore an article loses its utility, it also loses its value. The reason why means of production do not lose their value, at the same time that they lose their use-value, is this: they lose in the labor process the original form of their use-value, only to assume in the product the form of a new use-value. Hence it follows that in the labor process the means of production transfer their value to the product only so far as along with their use-value they lose also their exchange-value. They give up to the product

> that value alone which they themselves lose as means of production. (Ibid.)

Don't worry if you found that paragraph hard to understand: it is replete with erroneous and ambiguous propositions. First, the two final sentences, which appear to link the transfer of value by the machine to its depreciation, are incorrect (see below). Secondly, the statement that the use-value of a machine reappears in the use-value of the product equates the use-value of the machine to the utility enjoyed by the 'consumers' who purchase the goods the machine produces. But the use-value of a machine is specific to the capitalist purchaser of the machine only. By arguing that the use-value of the machine reappears in the product, Marx is in fact contemplating the existence of abstract utility, with the 'usefulness' of the machinery being transmuted into the 'usefulness' of the commodities it produces. If anything, this is neoclassical economics, not Marx.

The ambiguous statement concerns the transfer of value by the means of production. Which of their two 'values' do machines transfer, their exchange-value or their use-value? If Marx meant that they transfer their use-value, then this sentence would be correct in terms of his analysis of commodities. But later he makes it clear that by this expression he meant that the means of production transfer not their use-value (which is the case with a worker) but their exchange-value. In the clearest illustration of the flaw in his logic, he states that over the life of a machine, 'its use-value has been completely consumed, and therefore its exchange-value completely transferred to the product' (ibid.: 197). This amounts to the assertion that in the case of machinery and raw materials, what is consumed by the purchaser is not their use-value, as with all other commodities, but their exchange-value.

This ambiguity reappears as Marx discusses the example of a machine which lasts only six days. He first states the correct proposition that the machine transfers its use-value to the product, but then equates this to its exchange-value. He says that if a machine lasts six days '[t]hen, on the average, it loses each day one sixth of its use-value, and therefore parts with one-sixth of its value to the daily product.' Initially he draws the correct if poorly stated inference that 'means of production never transfer more value to the product than they themselves lose during the labor-process by the

destruction of their own use-value.' However, the ambiguity between exchange-value and use-value is strong, and his conclusion takes the incorrect fork. Stating his conclusion rather more succinctly than his reasoning, he says:

> The maximum loss of value that they [machines] can suffer in the process, is plainly limited by the amount of the original value with which they came into the process, or in other words, by the labor-time necessary for their production. However useful a given kind of raw material, or a machine, or other means of production may be, though it may cost £150 – yet it cannot, under any circumstances, add to the value of the product more than £150. (Ibid.)

Essentially, Marx reached the result that the means of production cannot generate surplus value by confusing depreciation, or the loss of value by a machine, with value creation. The truisms that the maximum amount of value that a machine can lose is its exchange-value, and that a machine's exchange-value will fall to zero only when its use-value has been completely exhausted, were combined to conclude that the value a machine adds in production is equivalent to the exchange-value it loses in depreciation. With the value added by a machine equated to value lost, no net value is transferred to the product, and therefore only labor can be a source of surplus value.

While the argument may appear plausible, in reality it involves a confusion of two distinct attributes of a machine: its cost (exchange-value) and its usefulness (use-value). From a Marxist perspective, depreciation is the writing-off of the original exchange-value of a machine over its productive life. Consequently, the maximum depreciation that a machine can suffer is its exchange-value. As it wears out, both its residual value and its usefulness will diminish, and both will terminate at the same time. However, it does not follow that the usefulness (the value-creating capacity) of the machine is equal to its cost (its depreciation). Though a capitalist will 'write off' the latter completely only when the former has been extinguished, the two aspects are nonetheless completely different and unrelated. There is no reason why the value lost by the machine should be equivalent to the value added.

An analogy with labor highlights the fallacy involved in equating these two magnitudes. If workers receive a subsistence wage, and if the working day exhausts the capacity to labor, then it could be argued that in a day a worker 'depreciates' by an amount equivalent to the subsistence wage – the exchange-value of labor-power. However, this depreciation is not the limit of the amount of value that can be added by a worker in a day's labor – the use-value of labor. Value added is unrelated to and greater than value lost; if it were not, there could be no surplus.

But don't take my word for it. Take Marx's.

The origin of surplus value (III)

As noted above, Marx first developed his dialectical analysis of the commodity while working on the rough draft of *Capital*. He was initially so enthused with this approach that he explored it freely, with almost no regard for how it meshed with his previous analysis. While doing this, he made a statement that correctly applied this new logic and directly contradicted the old, by stating that a machine could add more value than it lost through depreciation.

Table 1 is typical of Marx's standard numerical examples of value productivity. In that table, surplus value is directly proportional to labor-power ('variable capital'), and the value of the total product is the sum of the value of the means of production, plus variable capital, plus surplus value. In this analysis, the contribution of non-labor inputs to the value of output is exactly equal to their depreciation. However, when referring to a similar table shortly after developing his use-value/exchange-value analysis, Marx comments: 'It also has to be postulated (which was not done above) that the use-value of the machine significantly [*sic*] greater than its value; i.e. that its devaluation in the service of production is not proportional to its increasing effect on production' (Marx 1857).

There then follows the example shown in Table 17.7.

Both firms employ the same amount of variable capital – four days' labor which is paid 40 'thalers' (a unit in the German currency of the time), the value of the labor-power purchased. However, the first firm ('Capital 1'), with older capital, produces surplus value of just 10, while the second, with newer capital, produces a surplus of 13.33. The 3.33 difference in the

surplus they generate is attributable to the difference in their machinery, and the fact that 'the use-value of the machine significantly greater than its value; i.e. – its devaluation in the service of production is not proportional to its increasing effect on production.'[23]

TABLE 17.7 Marx's example where the use-value of machinery exceeds its depreciation

Production	Paper	Press	Working days	Wage bill	Surplus	Output	Rate SV (%)	Profit (%)
Capital 1	30	30	4	40	10	30	25.0	10.0
Capital 2	100	60	4	40	13.33	100	33.3	6.7

Marx without the labor theory of value

Marx's dialectical analysis thus contradicts a central tenet of the labor theory of value, that labor is the only source of surplus value. Having reached the conclusion above, Marx suddenly found himself trapped, as he had argued (in his PhD thesis) that Hegel was, in a compromise with his own principles. The principle of the dialectical analysis of the commodity was powerful, and the conclusions that followed logically from it inescapable: the labor theory of value could be true only if the use-value of a machine was exactly equal to its exchange-value, and yet a basic tenet of this analysis was that use-value and exchange-value are incommensurable.[24]

If Marx had followed his newfound logic, the labor theory of value would have been history. But with the labor theory of value gone, so too would be the tendency for the rate of profit to fall, and with it the inevitability of socialism.

The tendency for the rate of profit to fall was predicated upon the propositions that (a), over time, the capital-to-labor ratio would rise, and that (b), this would cause the rate of profit to fall. But (b) was dependent upon labor being the only source of surplus value, so that a rising capital-to-labor ratio would mean a falling rate of profit. If surplus could instead be garnered from any input to production, not just labor, then an increase in the

capital-to-labor ratio would have no necessary implications for the rate of profit: it could fall, rise, or stay the same.

With no necessity for the rate of profit to fall, there was similarly no necessity for capitalism to give way to socialism. Yet Marx had prided himself upon being the 'scientific socialist,' the one who in contrast to 'utopian socialists,' who merely dreamed of a better world, would prove why socialism had to come about. Now he finds that his new logical tool, which is evidently so superior to his old, challenges the basis of his argument for the inevitability of socialism.

It is little wonder that Marx then tried to find a way to make his new logic appear consistent with the old. By the time of *Capital*, he had convinced himself that the two were consistent: that the new positive methodology concurred with the old on the issue of the value productivity of machinery. Marx succumbed to the same flaw that (in his PhD thesis) he once noted in Hegel:

> It is conceivable that a philosopher should be guilty of this or that inconsistency because of this or that compromise; he may himself be conscious of it. But what he is not conscious of is that in the last analysis this apparent compromise is made possible by the deficiency of his principles or an inadequate grasp of them. So if a philosopher really has compromised it is the job of his followers to use the inner core of his thought to illuminate his own superficial expression of it. In this way, what is a progress in conscience is also a progress in knowledge. This does not involve putting the conscience of the philosopher under suspicion, but rather construing the essential characteristics of his views, giving them a definite form and meaning, and thus at the same time going beyond them. (Karl Marx, 1839: notes to his doctoral dissertation, reprinted in McLellan 1971)

So Marx succeeded in compromising his theory in a way which hid 'the deficiency of his principles or an inadequate grasp of them.' But 'success' came at a cost. The new logic, of which Marx was so proud, was ignored by his successors. In part, Marx contributed to this by the obfuscation he undertook to make his positive method appear consistent with the old

negative one. But I can't detract from the impressive contribution 'Marxists' themselves have made to the misinterpretation of Marx.

The misinterpretation of Marx

Though much of this occurred after his death, Marx had one taste of how his theories would be misinterpreted by friend and foe alike. He wrote a caustic commentary on the German economist Adolph Wagner's gross misinterpretation of his arguments in *Capital*, yet ironically, Wagner's hostile misinterpretation became the accepted interpretation of Marx by his followers after his death.

Wagner argued that Marx had completely misunderstood the notion of use-value, and that use-value played no part in Marx's analysis. Marx acerbically commented that:

Rodbertus had written a letter to him – where he, Rodbertus, explains why 'there is only one kind of value,' use value – Wagner says: 'This is completely correct, and necessitates an alteration in the customary illogical "division" of 'value' into use-value and exchange-value' – and this same Wagner places me among the people according to whom 'use-value' is to be completely 'dismissed' 'from science.' (Marx 1971 [1879])

Marx then made an emphatic statement of the role that use-value played in his economics:

All this is 'driveling.' Only an obscurantist, who has not understood a word of *Capital*, can conclude: Because Marx, in a note to the first edition of *Capital*, overthrows all the German professorial twaddle on 'use-value' in general, and refers readers who want to know something about actual use-value to 'commercial guides,' – therefore, use-value does not play any role in his work. The obscurantist has overlooked that my analysis of the commodity does not stop at the dual mode in which the commodity is presented, [but] presses forward [so] that surplus value itself is derived from a 'specific' use-value of labor-power which belongs to it exclusively etc. etc., that hence with me use-value plays an important role completely different than [it did] in previous [political] economy. (Ibid.)

Marx's protestations were to no avail. Despite such a strident statement that use-value was an essential component of his analytic method, and despite the fact that this document was available to and read by early

twentieth-century Marxists, use-value and the 'positive' methodology of which it was an integral part were expunged from mainstream Marxism. Paul Sweezy stated in his influential *The Theory of Capitalist Development* that

'Every commodity,' Marx wrote, 'has a twofold aspect, that of use-value and exchange-value.' Use-value is an expression of a certain relation between the consumer and the object consumed. Political economy, on the other hand, is a social science of the relations between people. It follows that 'use-value as such lies outside the sphere of investigation of political economy.' (Sweezy 1942, citing Marx 1859)

Yet ironically, the statement Sweezy used to support the notion that use-value plays no role in Marx's analysis was the very one referred to by Marx (in the reference to the 'first edition of *Capital*,' by which he meant the 1859 work *A Contribution to the Critique of Political Economy*), when he labeled Wagner an 'obscurantist.' In Marx's own words, therefore, twentieth-century Marxism has completely misunderstood the philosophical core of Marx's analysis of capitalism.

A poverty of philosophy Bose's critique and Marx's dialectic of the commodity establish that philosophy can't save the labor theory of value from Steedman's critique. Philosophical analysis strengthens Steedman's case that the labor theory of value is logically flawed.

Instead, mathematics and Marx's philosophy confirm that surplus value – and hence profit – can be generated from any input to production. There is no one source of surplus: Adam Smith's apparently vague musings that animals and machines both contribute to the creation of new value were correct.

Whither Marxism?

Marxist economics is analytically far stronger once it is shorn of the labor theory of value. The use-value/exchange-value methodology, which was applied above only to the question of the source of surplus value, has application to a huge range of issues on which labor theory of value Marxism is either silent or pedestrian (see Groll 1980 and Keen 1993a, 1993b and 2000 for a discussion of some of these). Marxism becomes the pinnacle of classical economics, rather than its dead end.

However, I am as pessimistic about the chances of this 'new, improved Marxism' being adopted by today's Marxists as I am about the chances of neoclassical economists abandoning the concept of equilibrium.

Their resistance, as with neoclassical economists to the critiques outlined in this book, is due in large part to ideology.

The advantage Marxists have over economists is that at least they are upfront about having an ideology. Marxists are as consciously committed to the belief that capitalism should give way to a socialism as economists are to the often unconscious belief that, if only we could rid ourselves of government intervention in the market, we would currently reside in the best of all possible worlds.

The tendency for the rate of profit to fall is crucial to this belief in the inevitability of socialism, and it is one of the many concepts that evaporate once the labor theory of value is expunged. Marxist economists are likely to continue to cling to the labor theory of value, to hang on to the faith, in preference to embracing logic.

If my pessimism is well founded, then Marxist economics will continue its self-absorbed and impossible quest for a solution to the transformation problem, and will remain irrelevant to the future development of economics.

However, labor theory of value Marxism will continue to be the ideology of choice of the left, particularly in the Third World. The argument that labor is the only source of profit, and that capitalism is thus based upon the exploitation of the worker, is a simple, compelling analysis to the downtrodden in our obscenely unequal world. A specter may no longer be haunting Europe, but Marxism will continue to be the banner of the dispossessed for many a year to come.

However, if non-neoclassical and non-Marxist economists can ignore the hullabaloo generated by the remaining band of adherents to the labor theory of value, and instead extract from Marx his rich philosophical foundation for the analysis of capitalism, then Marx's dialectical theory of value may yet play a role in the reform of economic theory. At present, however, the various non-neoclassical schools of thought have no coherent theory of value as an alternative to the neoclassical school's flawed subjective theory of value. But even though they lack the central organizing concept of a theory of value, these alternative schools of thought contain the

promise of an economic theory that may actually be relevant to the analysis and management of a capitalist economy.

18 | THERE ARE ALTERNATIVES[1]

Why there is still hope for a better economics

Maggie Thatcher's second-best-known comment, in defence of following monetarist economic policies, was 'There is NO alternative.' A similar attitude pervades economics: if not neoclassical economics, then what?

In fact, there are many alternative schools of thought within economics. In addition to Marxian economics, the main alternatives are:

- Austrian economics, which shares many of the features of neoclassical economics, save a slavish devotion to the concept of equilibrium.
- Post-Keynesian economics, which is highly critical of neoclassical economics, emphasizes the fundamental importance of uncertainty, and bases itself upon the theories of Keynes and Kalecki.
- Sraffian economics, based on Sraffa's concept of the production of commodities by means of commodities.
- Complexity theory and Econophysics, which apply concepts from nonlinear dynamics, chaos theory and physics to economic issues. And
- Evolutionary economics, which treats the economy as an evolving system along the lines of Darwin's theory of evolution.

None of these is at present strong enough or complete enough to declare itself a contender for the title of 'the' economic theory of the twenty-first century. However, they all have strengths in areas where neoclassical economics is fundamentally flawed, and there is also a substantial degree of overlap and cross-fertilization between schools. It is possible that this century could finally see the development of a dominant economic theory

which actually has some relevance to the dynamics of a modern capitalist economy.

I would probably be regarded as partisan to the post-Keynesian approach. However, I can see varying degrees of merit in all five schools of thought, and I can imagine that a twenty-first-century economics could be a melange of all five.

In this chapter I give a very brief overview of each school, emphasizing the ways in which they are superior to neoclassical economics, but also noting when they share its weaknesses, or have problems of their own. This will necessarily be an inadequate survey – doing a proper survey would necessitate another book. But as I commented earlier, it is essential to at least outline the alternatives, to debunk the myth that there is no alternative.

Austrian economics

The Austrian school (so called because its main early protagonists – Menger, Hayek and von Mises – were Austrian, though it is now mainly an American tradition) is a close relative of mainstream economics. It developed at much the same time, shared the same intellectual parents, and is comfortable with virtually every aspect of neoclassical economics save one: its obsession with equilibrium. This one divergence results in a theory which is markedly different from its dominant but wayward cousin.

Strengths Far from arguing that capitalism is the best social system because of the conditions which pertain in equilibrium, Austrian economists argue that capitalism is the best social system because of how it responds to disequilibrium.

The Austrians make the sensible observation that equilibrium is an intellectual abstraction which is unlikely ever to occur in the real world. All real-world economic situations will thus be disequilibrium ones, some of which enable entrepreneurs to make above-normal profits. By seeking out these profit opportunities, entrepreneurs make capitalism a dynamic, adaptive social system.

The Austrians therefore have an evolutionary perspective on capitalism, and argue that capitalism is evolutionarily superior to other social systems, such as feudalism and socialism.

The Austrians emphasize the importance of uncertainty in analyzing capitalism, whereas neoclassical economists, as we have seen, either ignore uncertainty, or trivialize it by equating it to probabilistic risk. This again gives Austrians an ideological reason to prefer capitalism to any other social system, since they argue that the disaggregated nature of capitalist society makes it more adaptable to uncertainty than other, more centralized systems.

The entrepreneur is the key actor in the Austrian vision of capitalism. It is the entrepreneur who attempts to profit from disequilibrium situations, thus innovating and adding to the diversity and strength of the capitalist system. The entrepreneurs are those who boldly act in the face of uncertainty, and though many will fail, some will succeed – thus strengthening the economic system via an evolutionary process.

The Austrians thus demonstrate that the economic fixation with equilibrium is unnecessary: it is possible to be an ideological supporter of capitalism even if you believe that equilibrium is irrelevant.

However, as a near-relative of neoclassical economics, this school shares a number of its disabilities.

Weaknesses First, the Austrians accept the economic argument that production is characterized by diminishing returns. As a corollary of this, they also accept the marginal productivity theory of income distribution – though they temper this by arguing that disequilibrium allows for entrepreneurs to make supernormal profits.

As was shown in Chapters 3, 5 and 6, these economic notions are fundamentally unsound. To the extent that Austrian economics relies upon these same concepts, it is also unsound.

A simple illustration of this arises from the Austrian theory of production. The economic model argues that an increase in the quantity of a factor of production – such as capital – will decrease its marginal product, and thus reduce its income.

The Austrians instead argue that a cheapening of capital – via a fall in the rate of interest – will lead to a more 'roundabout' approach to production, meaning that less direct labor and more indirect capital will be applied to its production.

Sraffa's critique of the neoclassical theory of production, detailed in Chapter 6, is equally applicable to this Austrian theory. By providing a way to measure capital inputs in terms of wage units, Sraffa showed that the economic concept of a quantity of capital was dependent on the rate of profit: the same logic shows that it is impossible to define one way of producing a commodity as 'more roundabout' than another independently of the rate of profit.

Consider two ways of making wine: process A, which involves the application of 1 wage unit now, 8 units last year, and 1 unit 8 years earlier; and process B, which involves 1 unit now and 1 unit 20 years ago. At a low rate of profit, process A might be more roundabout than process B; at a higher rate of profit, the order could reverse; and it could reverse again for a higher rate of profit. Therefore, the Austrian notion of roundaboutness is as internally inconsistent as the neoclassical concept of the marginal productivity of capital.

Secondly, even more so than conventional economics, Austrian economics has a faith in the self-adjusting properties of the capitalist economy, with Say's Law providing much of that confidence. As was argued in Chapter 9, Say's Law is invalid in a production economy with growth.

Thirdly, while it is in general an evolutionary approach to economics, at least one branch of Austrian economics, associated with Murray Rothbard, has a quite non-evolutionary attitude towards both the existence of the state and the role of money. The market economy may have evolved, but it seems the state was simply imposed from outside as an alien artifact upon our landscape. This is certainly one way to consider the growth of the welfare state; but an equally tenable argument is that the welfare state evolved as a response to the failures of the pure market system during the Great Depression.

Similarly, while they believe that the money supply should be determined endogenously – by either handing over money creation to private banks, or by returning to the gold standard – they argue that the current system of state money means that the money supply is entirely exogenous, and under the control of the state authorities. They then attribute much of the cyclical behavior of the economy to government meddling with the money supply and the rate of interest.

The post-Keynesian school, on the other hand, argues that though it may appear that the state controls the money supply, the complex chain of causation in the finance sector actually works backwards. Rather than the state directly controlling the money supply via its control over the issue of new currency and the extent to which it lets banks leverage their holdings of currency, private banks and other credit-generating institutions largely force the state's hand. Thus the money supply is largely endogenously determined by the market economy, rather than imposed upon it exogenously by the state.

The empirical record certainly supports post-Keynesians rather than Austrians on this point. Statistical evidence about the leads and lags between the state-determined component of money supply and broad credit shows that the latter 'leads' the former (Kydland and Prescott 1990). If the Austrians were correct, state money creation would instead precede private credit creation.

Maggie Thatcher's embrace of monetarism also provides an evocative counter-example. Despite her toughness, and her adherence to Milton Friedman's mantra that controlling the money supply would control inflation, even Thatcher's England was forced to abandon monetary targeting – setting some goal for the rate of growth of the money supply in order to force down the rate of inflation – because it could never meet the targets. If the 'Iron Lady' couldn't control the money supply, then no one could: evidence enough that the post-Keynesians are closer to the mark than the Austrians.

This non-evolutionary weakness in Austrian economics is a sign of a wider problem. The philosopher Chris Sciabarra, a specialist on the Austrian school and Ayn Rand, identifies an inconsistency between Hayek's notion of 'spontaneous order' – which corresponds to evolutionary development – and 'designed order' – where change is imposed from outside the market by the state. While such a distinction makes for good polemic writing against state intervention, it ignores the extent to which the state's own behavior might be reactive to the market, and thus, to some extent, also a form of spontaneous order. As Sciabarra puts it:

> There are more fundamental problems with Hayek's social theory. By positing such a sharp distinction between spontaneous order

> and designed order, Hayek has not provided us with any explanation of the emergence of those institutions which are agents of constructivism [designed order]. To what extent is the state itself a spontaneous, emergent product of social evolution? To what extent does the state define the parameters of the extended order which Hayek celebrates? What are the actual interrelationships between the spontaneous order of the market and the designed institutions of the state? The reader of Hayek's works will strain to find developed answers to any of these important questions. (Sciabarra 1995)

Finally, though Austrians eschew equilibrium analysis, and regard it as an unattainable state, their preference for capitalism as a social system is partly dependent on the belief that it will remain close to equilibrium. If, instead, capitalism is endogenously unstable, then it may remain substantially distant from equilibrium situations all the time. This weakens Austrian economics, to the extent that its support for capitalism emanates from conditions which are assumed to apply in equilibrium.

The Austrian scorecard Overall, I regard the Austrian school as too close to its neoclassical cousin to make a major contribution to a reformed economics. However, it does have some contributions to make, and for ideological reasons it is likely to be far stronger in the future – regardless of what I or other non-orthodox economists might think of it.

The Austrian emphasis upon innovation, and the role of the entrepreneur, are valid concepts which capture the way in which a market economy adapts. This aspect of capitalism is to some extent underrated by the other non-neoclassical schools, except for evolutionary economics. This aspect of Austrian thought could be valuable to a reformed twenty-first-century economics.

However, it is far more likely that the 'pure and simple' Rothbardian stream of Austrian economics will play a large role in twenty-first-century economics. This is because the Rothbardian approach provides an alternative way to ideologically support a capitalist economy as the best possible social system, whereas all other non-orthodox schools are to some degree critical of the concept of unfettered capitalism. If neoclassical

economics becomes untenable for any reason, the Austrians are well placed to provide an alternative religion for believers in the primacy of the market over all other forms of social organization.

The one barrier which stands in the way of today's neoclassical economist transmuting into tomorrow's Austrian is the Austrian insistence that there is little, if any, role for mathematics in economic analysis. Because the Austrians believe that all real-world data are generated in a situation of disequilibrium, and because they take seriously the aggregation problems noted in Chapters 2 and 4, Austrians deny that mathematical aggregate analysis has any validity. Faced with a choice between ideology and their beloved equilibrium mathematics, most economists would probably prefer to keep the mathematics. The one way out for neoclassicals would be to embrace the Austrian celebration of capitalism as a dynamic, disequilibrium system, and then model this using chaos and complexity theory. But this leads to the dilemma that such models almost always display 'far from equilibrium' behavior, which undermines the validity of beliefs about capitalism and welfare that depend on the economy not straying too far from equilibrium.

Post-Keynesian economics

This school of thought developed in reaction to the 'bastardization' of Keynes's economics in the so-called Keynesian–neoclassical synthesis. Regarding themselves as the true carriers of Keynes's message, they emphasized the importance of uncertainty in economic analysis, and the profound difference between the monetary economy in which we live, and the barter economy which neoclassical economics regards as an adequate proxy for the real world. As Arestis et al. (1999) put it, the main unifying themes in post-Keynesian economics are 'a concern for history, uncertainty, distributional issues, and the importance of political and economic institutions in determining the level of activity in an economy.'

Strengths The emphasis upon uncertainty as a fundamental aspect of the real world – one which cannot be approximated by risk – makes the post-Keynesian approach to economics far more realistic than the neoclassical.

Realism, in fact, is a central methodological emphasis of this school. Though there is no agreed post-Keynesian methodology to rival the

hedonistic calculus of the neoclassicals, post-Keynesians are united by their belief that an economic model has to be realistic.

One essential aspect of this is the insistence that a monetary economy is fundamentally different from one in which commodities simply exchange against other commodities. The issues of credit creation, the nature of money, the role of debt, etc., are far more pressing to post-Keynesians than they are to neoclassicals.

Macroeconomics is also a far more important concern. In fact, post-Keynesians reverse the neoclassical pecking order, to argue that whatever microeconomics is developed must be consistent with the observed behavior of the macroeconomy. A microeconomic model which is inconsistent with such things as business cycles, sustained unemployment, commonplace excess capacity, and the importance of credit is to post-Keynesians an invalid model.

Their preferred model of the firm emphasizes monopoly and quasi-monopoly power, decreasing costs of production with increased scale, markup pricing, the competitive need to sustain excess capacity, and the link between macroeconomic conditions and the firm's investment decisions.

Post-Keynesians are also not hung up on the need to show whether capitalism is a better or worse social system than any other. They are relatively agnostic on the question of what might constitute a better society.

This comparative independence from ideology means that post-Keynesians do not feel compelled, as neoclassicals do, to show that capitalism generates the best welfare outcome for the majority of the people. They are therefore much more comfortable with acknowledging the existence of social class in their models – something which leading neoclassical economists admit they might have to consider, given the failure of their individualistic approach to explain human behavior.

Weaknesses One of this school's great strengths is also a weakness. Unlike the neoclassical or Marxian schools, they do not have a 'theory of value' – they have nothing to compare to the theory of utility maximization, or even the labor theory of value.

This is certainly a strength when one considers how these theories of value have led these rival schools up intellectual garden paths. However, at

the same time it means that post-Keynesians lack a methodological consistency: they are more united by what they oppose than by what they have in common – though there are many common threads.

This lack of a theory of value makes it difficult for post-Keynesians to explain why their approach is superior to fledgling students of economics, who have yet to confront any of the intellectual conundrums which afflict neoclassical economics (and they also have difficulty communicating with radical students who are attracted to Marxism). One must normally become disenchanted with mainstream or Marxian economics before one can become a post-Keynesian. That is perhaps too tortuous a path to rely upon, if this school ever hopes to gain the ascendancy in economics.

A final problem is that, despite their rejection of neoclassical economics, they tend to also use static logic in their analysis – even though their building blocks might be, for instance, markup pricing rather than 'marginal cost equals marginal revenue.' This lack of appreciation of how different dynamic analysis is from static is not universal among post-Keynesian authors, but it is widespread enough to be a problem. However, it must be said that younger members of the post-Keynesian school are much more comfortable with dynamic analysis than are some of its older members.

The post-Keynesian scorecard Despite the lack of an agreed methodological foundation, the post-Keynesians are easily the most coherent alternative school of economic thought today. They are also likely to gain substantial credence in the event of a financial crisis, given their explicitly monetary approach to economics.

Sraffian economics

No prizes for guessing which economist provided the major inspiration to this particular group of economists. Sraffa's *Production of Commodities by Means of Commodities* became the icon for these economists. As well as applying it to critique other schools – notably neoclassical economics and Marxism – they attempted to turn it into a means to analyze the real economy.

Strengths There is no doubt that Sraffa's analysis constituted the most detailed and careful analysis of the mechanics of production in the history of economics. Not for him any simple abstractions, such as the neoclassicals' 'factors of production,' or even Marx's 'industry sectors': his model analyzed the interrelations of production at the level of the individual commodity.

This study turned up many subtleties that escaped other schools of economics: the dependence of the 'quantity of capital' on the profit rate, rather than vice versa, the phenomenon of reswitching, etc. No other school of economics matches the Sraffians on this insistence of assumption-free rigor. Well, almost assumption-free rigor.

Weaknesses The one assumption Sraffians do make is that the economy can be analyzed using static tools. As a result, even though the proper treatment of time was an essential component of Sraffa's critique of neoclassical economics, modern Sraffian economics makes no use of time or dynamics. Ian Steedman gave the pithiest explanation of why Sraffians analysis ignores dynamics. It is because '"static" analysis does not "ignore" time. To the contrary, that analysis allows enough time for changes in prime costs, markups, etc., to have their full effects' (Steedman 1992).

This proposition can be true only if the long-run position of an economy is an equilibrium one: if, in other words, the economy has just one equilibrium, and it is stable. As Chapter 9 showed, this is highly unlikely to be the case. A market economy is likely to have multiple equilibria, and they are all likely to be unstable. The Sraffian position is thus ignorant of modern dynamic analysis.

Sraffians also have one other flaw: they pay too little attention to Piero Sraffa.

Some post-Keynesians are fond of pointing out how pedantic Sraffa was, and therefore how important was the subtitle to his magnum opus. Sitting beneath the title of *The Production of Commodities by Means of Commodities* was the caveat 'Prelude to a critique of economic theory.'

In other words, these economists argue that Sraffa's method was intended solely to provide a means to critique other economic theories: it was never meant to provide a basis for an economic theory in itself.

Sraffa's 1926 paper provides support for this position. When discussing how the firm should be modeled, Sraffa put great stress upon the issues of importance to 'business men': the necessity and expense of marketing a non-homogeneous product to a market of non-homogeneous consumers, the cost and dangers of credit as a major force limiting firm size, etc. The concepts Sraffa discusses here can be considered only with extreme difficulty in the framework of his 1960 book (check the web link Alternatives/Sraffa for a relevant extract from Sraffa's 1926 paper).

The Sraffian scorecard Though the Sraffian school was fairly influential up until 2000, there have been few developments in it since, certainly in comparison to the growth in post-Keynesian economics since that date.

Complexity theory and Econophysics

Complexity theory is not so much a school of thought in economics as a group of economists who apply what is popularly known as 'chaos theory' to economic issues. Since the first edition of this book, there has also been an enormous growth in the number of physicists taking an active interest in economics and finance, and this new school of 'Econophysics' has largely subsumed the complexity theory approach.

The concept of chaos itself was first discovered in 1899 by the French mathematician Henri Poincaré. However, knowledge of it languished until the mid-1960s because it could not be fully explored until after the invention of computers. Chaotic models of necessity cannot be understood simply by writing down the equations which represent them: instead, they must be simulated, and their properties analyzed numerically. This was simply not possible before the advent of computers.

An essential aspect of complexity is the existence of nonlinear relationships between elements of a system, and the apparent ability of complex systems to 'self-organize.' The Lorenz model, noted in Chapter 8, has both these attributes: the nonlinear relations between displacement and temperature lead to behavior which on the surface is 'chaotic,' but behind which lies the beautiful organizing force of Lorenz's 'strange attractor.' Complexity theorists argue that the economy demonstrates similar attributes, and these are what give rise to the cycles which are a self-evident aspect of real-world economies.

Econophysics substantially adds to the contribution made by the early proponents of complexity in economics – such as Richard Goodwin (Goodwin 1990, 1991), Benoit Mandelbrot (Mandelbrot 1971, 2005), Hans-Walter Lorenz (Lorenz 1987a, 1987b, 1989), Paul Ormerod (Ormerod 1997, 2001, 2004); Ormerod and Heineike (2009), Carl Chiarella (Chiarella and Flaschel 2000, Chiarella, Dieci et al. 2002, Chiarella et al. 2003) and myself, among many others – by bringing both the techniques and the empirical mindset of physicists to bear upon economic data.

Over the last century, physicists have developed a vast array of techniques to interpret the equally vast range of physical processes encountered in everything from fluid dynamics to the behavior of subatomic particles. Their approach has been fundamentally empirical, and devoid of any a priori assumption that physical processes occur in equilibrium – and the concept of equilibrium itself is far more richly specified.

These techniques have enabled Econophysicists to make substantial progress in understanding how finance markets in particular actually operate, with a range of models that accurately capture the 'fat tails' that bedevil asset price data and lie well outside the predictive capacity of neoclassically inspired models.

Strengths It is extremely difficult to work in complexity theory and not understand dynamics. Though some neoclassical dabblers occasionally attempt to use equilibrium thinking in so-called chaotic models, in the main practitioners in this camp are extremely well versed in dynamics.

They are also normally very competent in mathematics; far more so, not only than other alternative schools of economics, but also than most neoclassicals. Many complexity theorists in economics started out doing PhDs in physics, biology, or mathematics itself, and later delve into economics out of curiosity.

This alone has meant that complexity theorists have had a significant impact upon the profession. While they rarely indulge in direct attacks upon neoclassical economics per se, neoclassical economists are aware that they are quite capable of doing so if provoked. This technical superiority over neoclassical economists has taken the mathematical 'big stick' out of the hands of neoclassicals. This has been taken to another level by

Econophysicists, whose training in mathematics and computing is far more rigorous and extensive than that undertaken by economists.

Weaknesses Though many complexity economists are inclined to a post-Keynesian perspective on economics, in general they lack a full appreciation of the history of economic thought. For this reason, they will often generate models which combine incompatible streams in economics. Concepts such as IS-LM and rational expectations often crop up in complexity or Econophysics models of the economy, with the authors rarely being aware of the origins of these 'tools.'

While Econophysics has developed a very rich and empirically based analysis of financial markets to date, and their statistical analysis here – involving concepts like Power Law distributions and Tsallis-statistics – is far more accurate than neoclassical models, success here has led to neglect of the 'econo' part of the developing discipline's name: at present it could more accurately be called 'Finaphysics' than 'Econophysics.'

Econophysicists also occasionally succumb to the temptation to introduce one of the strongest weapons in their arsenal, which I believe has no place in economics: 'conservation laws.' These apply where some fundamental aspect of a system – such as, for example, the amount of mass and energy in the universe – is not altered by the physical processes that apply to it, though its distribution and nature may alter. This condition that 'the change in the amount of X equals zero' has been the source of many of the greatest advances in physics, including the derivation of the theory of relativity.

No such equivalent concepts exist in economics, which is more akin to biology than physics in this respect: biological populations fluctuate, and there is no law requiring the mass of biological entities to remain constant, for example. Consequently economics belongs to the class of dynamical systems known as 'dissipative,' rather than 'conservative.'

A concern that conservation laws were being introduced into areas where they did not belong – for example, the analysis of money (Patriarca, Chakraborti et al. 2004; Ding, Xi et al. 2006) or the distribution of wealth – led me to contribute to a paper that was critical of recent developments in Econophysics (Gallegati, Keen et al. 2006). However, over time I expect

developments like these to dissipate, given the innately empirical focus of physicists.

The complexity scorecard Complexity theory and Econophysics are among the 'glamour' areas of science in general today, and this affects economics, even given its relative isolation from the scientific mainstream. The techniques which complexity modelers in economics employ are thus 'refertilizing' economics with concepts from other disciplines. The economic fixation upon equilibrium appears quaint to these mathematically literate economists, and this alone may significantly undermine the hold which static thinking has on economics.

If statics were to die, then inevitably so too would neoclassical economics, since its way of thinking is unsustainable in dynamics. So Econophysics may be a harbinger of real change in economics, after sixty years of effective ossification.

Evolutionary economics

Evolutionary economics draws its inspiration from the theory of evolution. In this, it has much in common with the majority of physical sciences, which in recent years have started to apply the concept of evolution – so much so that it has been proposed that Darwinism is the 'universal' basis of science (Nightingale 2000).[2]

In all sciences, the basic building blocks of the evolutionary way of thinking are diversity, the environment, and adaptation. Diversity gives a range of possible 'solutions' to the challenges thrown up by the environment. The environment interacts with these diverse forms to favor some over others – and the environment itself may be altered by feedback between it and these newly emergent species (Levins and Lewontin 1986). Adaptation occurs at the systemic level through the differential survival of some of these diverse forms: while some die out, others prosper, and thus their characteristics are passed on more strongly to subsequent generations.

The economic equivalents of diversity are the heterogeneity of consumers, and the variety of commodities; the equivalents of adaptation are new product development, and the consequent endogenous alteration of consumer tastes; the equivalent of the environment is the economy itself,

which is endogenously created by the actions of myriad individuals, social groups and corporations.

Strengths It is undeniable that the economy is an evolutionary system – with the one embellishment that change in economics is often purposeful, as opposed to the random nature of variation in the environment (though of course, purposive change can fail to achieve its intended ends).[3] This self-evident fact was the basis for Thorstein Veblen's query, over a century ago, of 'Why is economics not an evolutionary science?'

Manifestly it should be, and this alone should be a major factor in the rise of evolutionary thinking in economics.

Weaknesses One problem with evolutionary systems is that, effectively, everything can change. Economists, on the other hand, have been wedded to the notion of '*ceteris paribus*' ('all other things remaining equal') as a way of being able to impose some order on the apparent chaos of the market.

Ceteris paribus is of course an illusion, but the illusion often seems preferable to reality when it appears that fully acknowledging reality forces one to abandon structure.

This, of course, is not correct: evolutionary modeling still has structure, as shown by the advances made in genetics and many other areas where evolutionary thinking rules. However, economists are thrown back upon analogy here, since in economic systems there is no comparable entity to the gene, nor to the processes of biological interaction.

The difficulty evolutionary economists face is developing analytic tools which are consistent with evolution, and yet which still enable meaningful statements to be made about economic issues. Generally these have to include computer simulation, but unfortunately economists receive no training in computer programming. Fortunately, many students arrive at university with these skills already, and programming tools for evolutionary modeling – such as NetLogo (ccl.northwestern.edu/netlogo/) and Repast (repast.sourceforge.net/) – are far more accessible than their predecessors of even a decade ago.

The evolutionary scorecard Evolutionary economics is still in its infancy, and a lot of its time is spent defining basic philosophical concepts at one extreme, and developing computer-based evolutionary economic models at the other. If it can coalesce into a coherent school of thought with effective analytic tools, then it could at last make economics what, one century ago, Veblen knew it should be: an evolutionary science.

W(h)ither economics?

We are now well into the economic crisis that I anticipated in the first edition of this book in 2000, and which I (and a handful of other non-neoclassical economists) had actively warned of since late 2005. The public backlash against neoclassical economics that I expected this would cause has also occurred, with one-time supporters like *The Times* of London's economic columnist Anatole Kaletsky now openly attacking it:

> These are just a few examples of the creative thinking that has started again in economics after 20 years of stagnation. But the academic establishment, discredited though it is by the present crisis, will fight hard against new ideas. The outcome of this battle does not just matter to academic economists. Without a better understanding of economics, financial crises will keep recurring and faith in capitalism and free markets will surely erode. Changes in regulation are not sufficient after this financial crisis – it is time for a revolution in economic thought. (Kaletsky 2009)

Now that the need for 'a revolution in economic thought' is more widely acknowledged, the question is, how to achieve it?

I have no faith in the capacity of academic economics to reform itself. The historic record on this front is evidence enough: Keynes's challenge was assimilated and emasculated within a year of it being made by Hicks's IS-LM misinterpretation, and within thirty years all semblance of the change Keynes wished to cause had been eliminated. The misinterpretation of Fisher's debt-deflation hypothesis dismembered the one other substantive challenge to the neoclassical equilibrium, non-monetary mindset. Consequently, the neoclassical orthodoxy that dominated academic

economics prior to this crisis was even more extreme, virulent and intolerant of alternative approaches than that which Keynes and Fisher tried to challenge during the Great Depression.

Though there have been some signs of contrition and realization that the core of neoclassical economics may not be the perfect jewel they once believed it was, the overwhelming reaction of neoclassical economists to this crisis has been to maintain business as usual. I attended the 2011 American Economic Association meeting in Denver this year, at which there was a session on 'the 50th Anniversary of Rational Expectations.' What should have been a wake was in fact a celebration, and when one of its proponents was asked what economics would be like in fifty years, he was adamant that 'rational expectations' would still be at the heart of macroeconomic modeling.

Not if I can help it. If that fate does eventuate, then there will be another financial crisis right around the corner, and another rebel will have to try to bring about real change where I – and my colleagues in reform Edward Fullbrook, Paul Ormerod, Michael Hudson and many others – will have failed.

Change, if it is to come now rather than later, will have to be driven by outside influences: by journalists and influential commentators like Kaletsky who now realize how barren the neoclassical approach is; by a public far better informed via the Internet about the weaknesses of conventional economic thinking than was the public of the 1930s; by intellectuals from other disciplines who have long questioned the merits of neoclassical theory and can no longer be rebuffed when the global economy wallows in a crisis that neoclassical economics completely failed to anticipate; and by new students who, again via the Internet, now know that there are other ways to think about economics.

There are some encouraging signs today, though only time will tell if these lead to the change that economics desperately needs:

- The PAECON ('Protest against Autistic ECONomics') movement that began in France with the rebellion of a group of young economics students has since spawned an international movement, with both a network that unites the many academic opponents of neoclassical economics (www.paecon.net/PAEmovementindex1.htm), a publicly

accessible journal, the *Real-World Economics Review* (www.paecon.net/PAEReview/), and an active blog (rwer.wordpress.com/).

- George Soros has put some of his substantial wealth behind the Institute for New Economic Thinking (INET, ineteconomics.org/), in an effort to redress the effective exclusion of non-neoclassical researchers from official funding channels like, to be parochial, the Australian Research Council (which, to be personal, has turned down my applications for funding to develop models of debt deflation nine times since 1996).
- The 'blogosphere,' a phenomenon that has arisen since the first edition of *Debunking Economics* was published, now allows a plethora of commentators to take pot-shots at conventional thinking on economics. I list my favorites (in no particular order) below; while I don't agree with everything published by these commentators, I agree with a lot, and they are doing serious good in letting people know that economics need serious reform:

- Yves Smith at *Naked Capitalism*: www.nakedcapitalism.com/
- David Hirst at *Planet Wall Street* (website currently not available)
- Dan Denning at www.dailyreckoning.com.au/author/dan/
- Max Keiser at *The Keiser Report*: maxkeiser.com/
- Mish Shedlock at *MISH'S Global Economic Trend Analysis*: globaleconomicanalysis.blogspot.com/
- Chris Martenson: www.chrismartenson.com/
- Doug Noland at *The PrudentBear*: www.prudentbear.com/index.php/commentary/creditbubblebulletin
- Harry Dent at www.hsdent.com/
- Edward Harrison at *Credit Writedowns*: www.creditwritedowns.com/
- *Zero Hedge*: www.zerohedge.com/
- *The Automatic Earth*: theautomaticearth.blogspot.com/
- *The Levy Institute*'s program: www.levyinstitute.org/ and its blog *The Multiplier Effect*: www.multiplier-effect.org/
- The University of Missouri Kansas City Economics Department's blog *New Economic Perspectives*: neweconomicperspectives.blogspot.com/

- The *Institute for New Economic Thinking*'s (ineteconomics.org/) blog *The Money View*: ineteconomics.org/blog
- Bill Mitchell's *Billy Blog*: bilbo.economicoutlook.net/blog/
- Michael Hudson – one of the Bezemer 12 who foresaw and warned of the 2007 financial crisis, and a leading contributor to the academic literature on the origins on money – at michael-hudson.com/
- And my own *Steve Keen's Debtwatch*: www.debtdeflation.com/blogs/

Much more than this is needed, however.

Within universities, I would like to see other departments start to offer courses on economics using their methodologies, rather than that of neoclassical economics. Here I believe it is possible to use the ideology of neoclassical economics against it. Neoclassical economists are vehement opponents of monopolies, and yet in the past economics departments have jealously and destructively protected their monopoly on the word 'economics.' However, the empirical failure of neoclassical economics in predicting the Great Recession, and the paucity of alternative approaches within economics departments, is a good reason to remove this monopoly from them. I would be especially pleased to see engineering departments start to offer courses on a Systems Engineering Approach to Economics.

New students of economics can also do their bit. Don't let lecturers get away with teaching the same old stuff during the Great Recession that they taught before. Challenge them about why they exclude money and debt from their macro models, why they pretend to model dynamic processes using comparative statics, and so on. Make a nuisance of yourself – and organize with your fellow students to get a voice in designing the curriculum. This is how I began on my path thirty-eight years ago, and it is even more necessary now than it was then – and fortunately, there are much better resources to guide you about what an alternative curriculum should include.

Go beyond the standard curriculum too, to learn the skills you will need to be a twenty-first-century economist, rather than a not-yet-extinct fossil from the nineteenth century. Do basic courses in mathematics (calculus, algebra and differential equations), computer programming, history and sociology, rather than the additional fare neoclassical economists prescribe.

If you're really lucky, and you have an engineering department that teaches system dynamics (see en.wikipedia.org/wiki/System_dynamics), do those courses. Download and become familiar with programs like QED (www.debtdeflation.com/blogs/qed/), Vensim (vensim.com/), NetLogo, and build your own dynamic models, working from the leads I've given in this book.

Ultimately, I have faith in humanity's ultimate capacity to develop a realistic theoretical perspective on how a complex monetary market economy functions, and to leave behind the neat, plausible and wrong creation that is neoclassical economics.

Whether my faith on this front proves justified or delusional is not up to me, but to you.

NOTES

Preface

1 This is actually an application of the 'theory of the second best' (Lancaster and Lipsey 1956). Briefly, Lancaster and Lipsey showed that any single step towards what economics describes as the ideal situation could reduce welfare, if more than one step was required to move from the present situation to the ideal.

Chapter 1

1 Though somewhat later than I had anticipated, since the continued growth of the Subprime Bubble (and Federal Reserve interventions) had papered over the DotCom downturn.

2 There will be resistance aplenty too from government departments, and the bureaucracies of central banks, where promotion has come to those who have held the economic faith.

Chapter 2

1 The other sub-group calls itself 'New Classical.' As I explain in Chapter 10, neither of these subgroups bears any resemblance to either Keynes or the Classical School of economic thought. But their battle – publicized in the press as a battle between 'Keynesians' and the rest – has confused many members of the public into believing that the dominant school of thought in economics at the time of the crisis was 'Keynesian economics.' Nothing could be farther from the truth – if Krugman, Woodford and other self-described 'New Keynesians' are Keynesian, then because I can say 'quack,' I am a duck.

2 Bill White, the research director at the Bank of International Settlements, was a notable exception here since he was a proponent of the

non-neoclassical 'Financial Instability Hypothesis.'

3 The rate of interest the Federal Reserve charges when it loans to a commercial bank.

4 See media.ft.com/cms/3e3b6ca8-7a08-11de-b86f-00144feabdc0.pdf.

5 mpra.ub.uni-muenchen.de/15892/1/MPRA_paper_15892.pdf.

6 If you're a neoclassical economist, you're probably offended by this statement and regard it as a parody; if you're a professional from another discipline – say, engineering – who has not had any previous exposure to economic theory, you probably regard this as hyperbole. In either case, I'd suggest that you hold judgment until you finish this book.

7 The 'Money and Banking' course at Bernanke's alma mater, where he gave this speech, is a case in point. See www.anababus.net/teach/syllabusECO342.pdf.

8 An interesting instance of this is the observation by Mark Thoma on the Economist's View blog on 'What's wrong with modern macroeconomics: comments' (economistsview.typepad.com/economistsview/2009/11/whats-wrong-with-modern-macroeconomics-comments.html), which shows that he was unaware of significant papers that show the foundations of neoclassical theory are unsound – research I discuss in the next chapter: 'One thing I learned from it is that I need to read the old papers by Sonnenschein (1972), Mantel (1974), and Debreu (1974) since these papers appear to undermine representative agent models. According to this work, you cannot learn anything about the uniqueness of an equilibrium, whether an equilibrium is stable, or how agents arrive at equilibrium by looking at individual behavior (more precisely, there is no simple relationship between individual behavior and the properties of aggregated variables – someone added the axiom of revealed preference doesn't even survive aggregating two heterogeneous agents).'

9 It is also because complexity theory tends to be incompatible with neoclassical economics, since a common property of complex systems is that they have unstable equilibria: see Chapter 9.

10 Nonlinear 'difference' equations also generate chaos, but economics courses normally cover only linear difference equations.

11 By way of balance, I also know that some of what I say about Marxism will be offensive to Marxist economists.

12 Curiously, academic neoclassical economists don't follow this philosophy themselves: they really believe that they are promoting the common good by developing and teaching neoclassical economics.

13 In the first edition, since my target audience didn't have access to academic journals, I decided to make references to academic papers uncluttered by not giving page references. Since I am now catering for an audience that does have access to those journals, all new references in this edition have page references for quotations.

14 Except in one footnote, where the equation concerns meteorology rather than economics, and can easily be skipped. Occasionally, when some proposition in the text is best stated in mathematical form, I have used words rather than mathematical symbols.

Chapter 3

1 However it is also also empirically impossible, as I discuss in the addendum.

2 Most of Bentham's endeavors in this regard related to devising a scale of punishments that he regarded as just sufficient to discourage the commission of crime.

3 Jevons, one of the three co-founders of neoclassical economics, was justly skeptical that mathematics could treat all behavior. He argued that 'economy does not treat of all human motives. There are motives nearly always present with us, arising from conscience, compassion, or from some moral or religious source, which economy cannot and does not pretend to treat. These will remain to us as outstanding and disturbing forces; they must be treated, if at all, by other appropriate branches of knowledge' (Jevons 1866). However, subsequent economists have applied this theory to all behavior, including interpersonal relations.

4 Cardinal refers to the ability to attach a precise quantity, whereas ordinal refers to the ability to rank things in size order, without necessarily being able to ascribe a numeric value to each.

5 As I point out later, the mathematician John von Neumann developed a way that a cardinal measure of utility could be derived, but this was ignored by neoclassical economists (Von Neumann and Morgenstern 1953: 17–29).

6 At its base (where, using my 'bananas and biscuits' example, zero bananas and zero biscuits were consumed), its height was zero. Then as you walked in the bananas direction only (eating bananas but no biscuits), the mountain rose, but at an ever-diminishing rate – it was its steepest at its base, because the very first units consumed gave the greatest 'marginal utility.' The same thing applied in the biscuits direction, while there was some path in the combined 'biscuits and bananas' direction that was the steepest of all to begin with.

7 The Wikipedia entry on contours explains how isobars are derived, and actually mentions indifference curves as an example of them in economics: en.wikipedia.org/wiki/Contour_line.

8 Economists assume that consumers spend all their income. They treat savings, in effect, as a form of consumption – only what is 'consumed' are goods in the future.

9 This is different to indirectly altering the consumer's effective income via a change in prices.

10 Hippasus apparently used the geometry of pentagrams to prove the existence of irrational numbers. The proof by contradiction that the square root of two is irrational, though mathematical, is very easy to understand. See footnote 1.

11 They ignore the role of credit in the economy, an issue that looms very large in my later critique of macroeconomics.

12 The remark may well be apocryphal – see en.wikipedia.org/wiki/Let_them_eat_cake; but the sentiment of the wealthy disregarding the fate of the poor certainly played a major role in ushering in the French Revolution.

13 Or returns two values for one input.

14 Mas-Colell's assumption of a 'benevolent central authority' that 'redistributes wealth in order to maximize social welfare' is probably derived from this ridiculous paper by Samuelson, since he references it as a paper 'For further discussion' (Mas-Colell, Whinston et al. 1995: 118).

15 Say gives a typical statement (reproduced on the Web at Hedonism/Say) of this approach to utility, which denies the ability of anyone to judge or measure the utility any other individual garners from a particular commodity.

16 Kirman's paper is an eloquent and well-argued instance of the phenomenon that those who have constructed the 'high theory' of economics are far less confident about its relevance than more ordinary economists.

17 The 'plus one' rule covers the case of buying no units of one commodity. This isn't an issue in my discrete interpretation of Sippel's experiment.

18 If this sounds extraordinary to you, consider that 10 multiplied by itself 8 times is equal to 100 million.

Chapter 4

1 In fact, the advanced courses also ignore these more difficult critiques, which means that students who do them, if anything, are even more ignorant than undergraduates.

2 This last paper – 'Debunking the theory of the firm – a chronology' – is freely downloadable from www.paecon.net/PAEReview/issue53/KeenStandish53.pdf.

3 Economists are likely to deflect these critiques by arguing that the theory has moved well beyond the simplistic models taught to undergraduates. However, at the very least economists should stop teaching these models. Secondly, economists still see the model of perfect competition as describing the ideal economy. This chapter argues that this ideal is in fact a farce.

4 Astute readers would already see a problem here with the model of perfect competition.

5 These numbers come from a mathematical function (a cubic – of the form a+bx+cx2+dx3 – to be precise), whereas most neoclassical textbooks, if they use numerical examples at all (most don't, and simply use drawings instead), simply 'pluck them out of the air.'

6 If this example seems silly to you – surely you would only use workers and machines in the ideal ratio, so if you have just one worker, he works one jackhammer while the other ninety-nine are left idle? – then congratulations, you're right! The whole idea that firms vary the ratio of fixed to variable factors as economists assume they do is a nonsense that we tackle in the next chapter.

7 Average fixed costs start off very high – in this example, at $10,000 per unit for the first unit produced – and fall uniformly from then on. Variable costs per unit may fall for a while too as productivity rises, but eventually they start increasing as output rises, and marginal productivity falls. The combination of falling fixed costs per unit of output, and rising variable costs, means that average costs are 'u-shaped': they fall while the firm experiences rising marginal productivity, flatten out as diminishing marginal productivity kicks in, and finally rise when marginal productivity has diminished so much that each additional unit costs more to produce than the average to date.

8 In fact, on the two occasions that Smith used this phrase, it was in relation to income distribution, and whether local producers would ship production offshore, not how the market mechanism operated. But the metaphor had a compelling impact on the development of economic theory about the market.

9 There are two other models, known as the Cournot-Nash and Bertrand models, in which firms do react to what they think other firms will do, which also reach the outcome that price equals marginal cost. Though they don't make the same mathematical error as the Marshallian model does, they have other problems that we discuss in Keen and Standish (2010). In a third edition of *Debunking Economics*, I might add an addendum on this – since I'm sure it will be the refuge of those who wish to cling to the neoclassical model – but I've left it out of this edition to avoid boring the rest of my audience to death.

10 This assumption is inconsistent with the assumption of a 'short run,' during which some factor of production cannot be changed, which is essential to get the phenomenon of diminishing marginal productivity, which in turn generates a rising marginal cost (see Chapter 4). Firms already inside the industry are assumed to be unable to alter their capital equipment at all, but in the same time period firms not currently in the industry can build a factory and move in on the market? Hello? This logic is about as watertight as the script of the average TV soap opera. However, this assumption plays no part in the standard mathematical model of perfect competition, which focuses simply on the impact of the number of firms currently in the industry.

11 If this argument doesn't convince you, good – because this is the point at which the economic argument starts to take on that 'flat earth' feeling.

12 There is no specification of time in the standard neoclassical model, so this could be, for example, 135 units per minute.

13 The Cournot and Bertrand theories erroneously argue that the level that maximizes firms' profits is identified by the firms behaving collusively, like a pseudo-monopoly. Instead, in Keen and Standish (2010), we show that the so-called collusive output level is simply the level the firms would produce if they behaved as simple profit maximizers.

14 Sraffa's paper 'The law of returns under competitive conditions' (Sraffa 1926) critiqued a forerunner to this idea: that economies of scale could be external to the firm, but internal to the industry. This would mean that as an industry expanded, all firms benefited from lower costs, but none benefited any more than any other. Sraffa argued that few, if any, economies of scale would fit in this category: instead most would be internal to a firm, and thus advantage the big firm more than the small.

15 The revised formula, in this 1,000-firm example, is that the firm should make the gap between its marginal revenue and marginal cost equal to 999/1000th of the gap between market price and its marginal cost. The number of firms can be safely ignored and the output level chosen will still be approximately right, whereas the neoclassical formula remains precisely wrong.

16 The then editor of the *Journal of Economics Education*, Bill Becker, was himself keen to have the paper published, and submitted it to eminent referees to try to improve its chances of being accepted.

17 In addition, to compare competitive firms to monopolies at all scales of output, then – for reasons outlined in Chapter 5 – the marginal cost curve must be drawn horizontally.

Chapter 5

1 Smith and Ricardo allowed exceptions to this rule; neoclassical economics in effect made these exceptions the rule.

2 Money is simply a measuring stick in this analysis, and the monetary unit could as easily be pigs as dollars.

3 As we'll see shortly, this generates just as many problems as aggregation of demand curves did.

4 Significant fixtures like factory buildings are an exception here, though machinery within the factory is not. Kornai's observations about the spare capacity in production, which I note later, supplement Sraffa's critique on this point.

5 Kornai used this last constraint to develop the concept of 'Hard and soft budget constraints.' This has great relevance to the position of banks during the Great Recession (see Kornai, Maskin et al. 2003: 1123–6; Kornai 1986), but is less relevant here.

6 Marketing expenses cannot be added in to 'rescue' the doctrine, since the true purpose of marketing is to alter the firm's demand curve, and this only makes sense if firms produce differentiated products – something the theory of perfect competition explicitly rules out.

7 The 'cost curve' for any one firm or industry is the product of interactions between all industries, an issue that is ignored in the neoclassical treatment of a single market. This issue is discussed in Chapter 6.

8 Economics ignores the issue of ecological sustainability, though clearly it must be considered by a reformed economics.

9 Friedman argued that the result that businessmen do not make their decisions on the basis of marginal cost and marginal revenue was 'largely irrelevant' (Friedman 1953: 15).

10 Though empirical work suggests that, in practice, there is little sign of any negative relationship between the quantity sold and the price – and hence little evidence of a 'demand curve' (Lee 1996).

Chapter 7

1 While this case is most easily made with equations, I'll stick to words here.

2 The same case can be made with respect to the change in the wages bill, but I focus just on profit times capital to keep the argument simple.

3 The ratio of a change in capital to a change in capital is 1.

4 This will apply only when the capital-to-labour ratio is the same in all industries – which is effectively the same as saying there is only one

industry.

5 Of course, this argument has already been eliminated by the 'benevolent central authority' assumption derived from the Sonnenschein-Mantel-Debreu conditions.

6 The rule in this example is that 10 quarters of wheat had to exchange for 1 ton of iron, or 2 pigs. These are relative price ratios in which commodities exchange – rather than absolute prices in terms of money.

7 At least, not until a 'von Neumann machine' – a machine that can both produce output and reproduce itself – is invented.

8 This is often all economists know of Sraffa's critique, and they dismiss it immediately by saying that it wrongly ignores the issue of marginal productivity. In fact, there is much more to Sraffa's critique, and Bhaduri's critique establishes the invalidity of the assertion that the rate of profit equals the marginal productivity of capital.

9 When output is measured in terms of a 'standard commodity,' and when the wage is normalized so that when the rate of profit r is zero, the wage w equals 1.

10 This correspondence is not exact, but it can be made accurate to any level short of 100 percent by continuing the process of reduction for long enough.

11 Approximately because of the irreducible commodity residue left from the reduction process.

12 I'm enough of a wine buff to realize that this example is practically impossible, but it will do as an illustration.

Chapter 8

1 I have hardened my opinion on this front since the first edition, when I was willing to describe economics as a science, though a rather 'pathological' one.

2 I first heard this joke in a public debate between my then professor of economics and a physicist. I now appreciate the irony that physicists are turning their attention to economics – and in general being horrified by neoclassical economic theory.

3 I am grateful to my student Marchessa Dy for suggesting this very evocative analogy.

4 The analysis below is a brief summary of Imre Lakatos's concept of competing 'scientific research programs.' The philosophy of science is today dominated by more 'postmodernist' concepts. I will leave exploration of these newer strands to the interested reader to pursue.

5 This reference to physics is now seriously dated, since this empirical observation has now been corroborated – see the Wikipedia item on the 'Accelerating Universe' for a brief discussion.

6 Ironically, Austrian economics, an alternative school of thought that is very closely related to neoclassical economics, differs by singing the praises of capitalism as a disequilibrium system (see Chapter 18).

7 Equilibrium in turn has been endowed with essential welfare properties, with a 'Pareto optimal equilibrium' being a situation in which no one can be made any better off without making someone else worse off.

Chapter 9

1 If you have ever taught a child to ride a bike, you would know that this lesson is the most difficult one to grasp – that a moving bike balances itself, without the need for training wheels or other props which would keep it upright when it was stationary.

2 This analogy is apt in more ways than one. The art of balancing a stationary bike requires great skill, and anyone who has mastered it is likely to 'show it off' at every opportunity, regardless of how impractical it might be. Similarly, economists who have mastered the difficult mental gymnastics involved in equilibrium analysis take every opportunity to parade their prowess – regardless of how irrelevant this skill might be to the art of managing a real economy.

3 Only the gold market in London even approaches this structure, and even that is a market at which only one commodity is traded, rather than 'all commodities' (O'Hara 1995).

4 The main theorem is the Perron-Frobenius theorem on the eigenvalues of a positive matrix. See en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem for an explanation.

5 Debreu used a notation that allowed for negative prices and negative quantities. However, this was a convenience only, and has no impact on the

analysis in this section.

6 My discussion of the instability of general equilibrium above was with respect to a production economy, where the nature of the input-output matrix makes stability impossible. There is no input–output matrix in an exchange-only economy because there is no production!

7 The actual equations were: 'the rate of change of x with respect to time equals the constant a multiplied by (y–z); the rate of change of y with respect to time equals x multiplied by (b–z) minus y; the rate of change of z with respect to time equals (x multiplied by y) minus (c multiplied by z).'

8 I use chapter and section references for Marx, rather than page numbers, since his work is now freely accessible via the Internet from the site www.marxists.org/archive/marx/.

9 The two equations are linked, because workers' wage demands depend on the rate of employment, while investment – which determines the rate of growth – depends on income distribution (a higher workers' share means lower profits, and hence lower investment).

10 For more details, see the Wikipedia entries en.wikipedia.org/wiki/Functional_flow_block_diagram, en.wikipedia.org/wiki/Transfer_function, http://en.wikipedia.org/wiki/State_space_(controls) and en.wikipedia.org/wiki/Systems_engineering.

Chapter 10

1 he became Fed chairman in February 2006, having briefly served as chairman of the president's Council of economic advisers before that.

2 more strictly, a market demand curve can have any shape that can be described by a polynomial equation. This rules out a curve that returns two or more prices for the same quantity, but allows curves that return the same price for many different quantities.

3 keynes lumped what we today term neoclassical economists with those we today call the classical economists. While they are distinctly different schools of thought, keynes was correct to group them together on this issue, since they concurred that a general deficiency of aggregate demand was impossible.

4 Surprisingly few books give this argument in full, given the extent to which it is a core belief in economics. Two that do are Baird (1981: ch. 3) and Crouch (1972: ch. 6).

5 his actual procedure was to argue that, when employment increased, demand for consumer goods would increase by less than the increase in employment, and that equilibrium would be achieved only if investment demand automatically took up the slack. This confusing argument is equivalent to the simpler case set out here.

6 as milgate observed, 'received opinion, that keynes's General Theory is a contribution to "disequilibrium" analysis, was stamped indelibly upon the collective consciousness of the economics profession at an early date – by critics and converts alike' (milgate 1987).

7 I explain how marx derived this result in Chapter 17.

8 For those of you for whom mcCarthyism is ancient history, see en.wikipedia.org/wiki/mcCarthyism. Though mcCarthy was out of the picture by the late 1950s, the influence of that period continued for many years.

9 I have found that many people find this confusing on the basis that, if debt has financed a purchase, wouldn't that already be recorded in Gdp? There are two reasons why this is not the case. First, part of spending is on pre-existing assets – which are not a component of Gdp. Secondly, in our demand-driven economy, the demand comes first – before the supply – and demand can be sourced either from previously earned income, or an increase in debt – where this debt reflects an increase in the money supply by the private banking system, as I explain in Chapters 12 and 14. The debt-financed demand for commodities does later generate production of more commodities, and this turns up in Gdp – but the debt precedes the supply. This relationship is thus best thought of in 'continuous time' terms: aggregate demand at a point in time equals income at that time, plus the change in debt at that time. aggregate supply (and the sale of existing assets) follows slightly later.

10 Sometimes they do, of course, but in order to clarify his argument Schumpeter considers the case where an entrepreneur does not have pre-existing money and must therefore borrow to finance his venture.

11 marx's theory of value is normally regarded as the labor theory of value, which is criticized in Chapter 13. I argue that his theory of value is

something quite different.

12 There is an interesting parallel in research into producing robots that can walk. The first attempts designed a robot that always kept its center of gravity directly above the foot in contact with the ground – resulting in a robot that was always in gravitational equilibrium, but which could walk only in straight lines with five seconds between steps. To enable fluid motion, the researchers found they had to put the center of gravity in continuous disequilibrium: then it could walk as naturally as we humans do. See world.honda.com/aSImo/history/e0.html and world.honda.com/aSImo/technology/walking_02.html for details.

13 Whether this growth can be sustained indefinitely is another matter altogether that I do not address in this book. On that front I regard *The Limits to Growth* (Meadows, Randers et al. 1972) as the definitive reference.

14 Though because keynes hadn't completely escaped from the neoclassical way of thinking, those concepts do occasionally occur in the *General Theory*, in a very muddled way – as the lengthy quote from the *General Theory* illustrates.

15 Blatt also provides an excellent mathematical treatment of investment under uncertainty – see Chapters 12 and 13 respectively of Blatt (1983) and Boyd and Blatt (1988).

16 hicks also had savings depending upon the level of output, but output was already determined by the first equation and therefore 'we do not need to bother about inserting Income here unless we choose' (hicks 1937).

17 Total employment is the sum of the number of workers needed to produce investment output and the number needed to produce consumption output, so if labour productivities differ between the two sectors then the breakdown has to be known before total employment is determined.

18 another paper in a mainstream journal makes some similar concessions (hicks 1979).

19 only hicks could see similarities between keynes's work and this bizarre model of a one-commodity economy (bread) which had a market in which prices were set on one day (monday) that then applied for the remainder of the week, and in which there was no model of how the bakery that made the bread was actually manufactured.

20 This is false, as a simple check of the table of contents of the *General Theory* can confirm: Chapter 19 is entitled 'Changes in money-wages.' In it, keynes concludes that flexible wages would not eliminate the prospect of deficient aggregate demand.

21 Walras's Law is invalid in a growing economy, as I explained earlier. This section considers when it can't be applied to eliminate one market from the analysis even in the no-growth realm to which it does apply.

22 many neoclassical macroeconomic models to this day are based on IS-Lm and have time-based equations – including one for the price level – in them that appear superficially dynamic. however, most of these models are solved by assuming that the price level (and everything else) converges to a long-run equilibrium over the medium term, which is a travesty of proper dynamic modeling.

23 mathematicians and system dynamics practitioners would find the first of these references very strange. It gives a detailed discussion of how to solve a nonlinear model where the solution involves approximating a matrix inversion – which can only derive the equilibrium for a model – yet makes no mention of standard numerical techniques for simulating systems of differential equations (like the runge-kutta or Levenberg-marquardt methods), which can return the actual time path of a model rather than simply its equilibrium. It's as if economists live in a parallel universe where techniques that are commonplace in real sciences haven't been invented yet.

24 The number in square brackets refers to the page numbers of the online reprint of Frisch's paper, available at www.frisch.uio.no/Frisch_Website/ppIp.pdf.

25 as so often happens in economics, the 'founding father' responsible for this view also contemplated the alternative possibility, that fluctuations were endogenous to the economy, as Schumpeter argued at the time. however, since this was more difficult to model, he left it for others to do later: 'The idea of erratic shocks represents one very essential aspect of the impulse problem in economic cycle analysis, but probably it does not contain the whole explanation […] In mathematical language one could perhaps say that one introduces here the idea of an auto-maintained oscillation […] It would be possible to put the functioning of this whole instrument into equations under more or less simplified assumptions about the construction and functioning of the valve, etc. I even think this will be a

useful task for a further analysis of economic oscillations, but I do not intend to take up this mathematical formulation here.' (Frisch 1933: 33–5). Unfortunately, his successors stuck with his easier-to-model exogenous shocks analogy, leaving his sensible suggestion to model endogenous fluctuations to wither on the vine.

26 There's at least one phd in producing such a simulation model – I hope some brave student takes that task on one day (brave because it would be a difficult task that would make him highly unpopular with neoclassical economists).

27 Joan robinson, who played a leading role in the Cambridge controversies outlined in Chapter 8, coined the term 'Bastard keynesianism' to describe the neoclassical interpretation of keynes (robinson 1981).

28 Lucas was far from the only exponent of this microeconomic takeover of macroeconomics – others who made a significant contribution to the microeconomic hatchet job on macroeconomics include muth, Wallace, kydland, prescott, Sargent, rapping, and latterly Smets and Woodford.

29 of course, an economy without growth hasn't existed, but Friedman extended this belief in the economy tending to full-employment equilibrium over to his model with growth, and he had the same views about the actual economy.

30 Now you know where the 'helicopter Ben' moniker that is applied to Ben Bernanke actually comes from! I would regard this as unfair to Bernanke, were it not for his fawning speech at Friedman's ninetieth birthday, noted later.

31 While inflation did ultimately fall, the policy was nowhere near as easy to implement as Friedman's analysis implied – the Federal reserve almost always failed to achieve its targets for money growth by large margins, the relationship between monetary aggregated and inflation was far weaker than Friedman implied, and unemployment grew far more than monetarists expected it would. Central banks ultimately abandoned money growth targeting, and moved instead to the 'Taylor rule' approach of targeting short-term interest rates. See desai (1981) and kaldor (1982) for critiques of the monetarist period.

32 See www.ukagriculture.com/production_cycles/pigs_production_cycle.cfm.

33 None of these made it through to the version of rational expectations that was incorporated into models of the macroeconomy.

34 'ergodic' is a frequently misunderstood term, especially within economics. It is properly defined by the Wiktionary (en.wiktionary.org/wiki/ergodic), and the Wikipedia entry on ergodic Theory (en.wikipedia.org/wiki/ergodic_theory) makes the important point that 'For the special class of ergodic systems, the time average is the same for almost all initial points: statistically speaking, the system that evolves for a long time "forgets" its initial state.' This is not the case for complex or chaotic models, which show 'sensitive dependence on initial conditions' (see en.wikipedia.org/wiki/Butterfly_effect and en.wikipedia.org/wiki/Chaos_theory).

35 I can think of no more apt term to describe the group that led the campaign to make macroeconomics a branch of neoclassical microeconomics. Certainly the neoclassical attitude to researchers who refused to use 'rational expectations' in their models approached the old mafia cliché of 'an offer you can't refuse': 'assume rational expectations, or your paper won't get published in a leading journal.'

36 This is based on the belief that output would be higher (and prices lower) under competition than under monopoly, which I showed to be false in Chapter 4.

37 a rule of thumb that asserts that the central bank can control inflation by increasing real interest rates roughly twice as much as any increase in inflation. See Box 10.1.

38 he noted that 'the first two equations of the model are patently false […] The aggregate demand equation ignores the existence of investment, and relies on an intertemporal substitution effect in response to the interest rate, which is hard to detect in the data on consumers. The inflation equation implies a purely forward-looking behavior of inflation, which again appears strongly at odds with the data' (Blanchard 2009: 215).

Chapter 11

1 The first draft of this chapter, completed in March 2000, began with the sentence 'The Internet stock market boom is the biggest speculative bubble in world history.' This was before the Nasdaq crash of 4 April 2000

– when, as luck would have it, I was actually in New York on holiday, and, as one then could, observed the action on a tour of the NYSE (and later in Times Square on the giant Nasdaq screen). For the publication itself, 'is' became 'was,' since the book was sent to the typesetters in November 2000, when the Nasdaq was down 50 percent from its peak, and the bubble was clearly over.

2 Though he avoided bankruptcy thanks to loans from his wealthy sister-in-law (they were never repaid, and she forgave them in her will; Barber 1997), and selling his house to Yale in return for life tenancy.

3 Fisher's theory was first published in another work in 1907; *The Theory of Interest* restated this theory in a form which Fisher hoped would be more accessible than was the 1907 book.

4 Muslim societies continue with the traditional definition, and therefore prohibit – with varying degrees of effectiveness – any loan contract in which the lender does not share in the risk of the project.

5 'In a country, such as Great Britain, where money is lent to government at three per cent and to private people upon a good security at four and four and a half, the present legal rate, five per cent, is perhaps as proper as any' (Smith 1838 [1776]: Book II, ch. 4).

6 The best record of the famous early panics is in Charles Mackay's *Extraordinary Popular Delusions and the Madness of Crowds*. The chronicler of our day is Charles P. Kindleberger.

7 Of course, many of the high-flying companies of 1929 were no longer in the index in 1955, so that anyone who held on to their 1929 share portfolio took far more than twenty-five years to get their money back, and most of the shares they held were worthless.

8 Barber notes that after Fisher came into great wealth when his filing invention was taken over by the Remington Rand Corporation, he was 'eager to add to his portfolio of common stocks and placed himself in some exposed positions in order to do so. At this time, his confidence in the soundness of the American economy was complete' (Barber 1997).

9 Barber observed that among the other reasons was the fact that 'In the 1930s, his insistence on the urgency of "quick fix" solutions generated frictions between Fisher and other professional economists' (ibid.).

10 Almost 90 percent of the over 1,200 citations of Fisher in academic journals from 1956 were references to his pre-Great Depression works

(Feher 1999).

11 Strictly speaking, this was supposed to be anything in which one could invest, but practically the theory was applied as if the investments were restricted to shares.

12 Since diversification reduces risk, all investments along this edge must be portfolios rather than individual shares. This concept is important in Sharpe's analysis of the valuation of a single investment, which I don't consider in this summary.

13 In words, this formula asserts that the expected return on a share will equal the risk-free rate (P), plus 'beta' times the difference between the overall market return and the risk-free rate. Beta itself is a measure of the ratio of the variability of a given share's return to the variability of the market index, and the degree of correlation between the share's return and the market index return.

14 There are three variations on this, known as the weak, semi-strong and strong forms of the EMH.

15 As I have explained, however, to Fisher's credit, his failure led to an epiphany that resulted in him renouncing neoclassical thinking, and making a major contribution to the alternative approach to economics that Minsky later developed into the Financial Instability Hypothesis.

Chapter 12

1 Bernanke went on to rephrase debt deflation using several concepts from neoclassical microeconomics – including information asymmetry, the impairment of banks' role as adjudicators of the quality of debtors, and so on. He also ultimately developed a cumbersome neoclassical explanation for nominal wage rigidity which gave debt a role, arguing that 'nonindexation of financial contracts, and the associated debt-deflation, might in some way have been a source of the slow adjustment of wages and other prices' (Bernanke 2000: 32–3). By 'nonindexation,' he meant the fact that debts are not adjusted because of inflation. This is one of many instances of Bernanke criticizing real-world practices because they don't conform to neoclassical theory. In fact, the only country ever to put neoclassical theory on debts into practice was Iceland – with disastrous consequences when its credit bubble burst.

2 For a start, Fisher's process began with over-indebtedness, and falling asset prices were one of the consequences of this.

3 There are numerous measures of the money supply, with varying definitions of each in different countries. The normal definitions start with currency; then the 'Monetary Base' or M0, which is currency plus the reserve accounts of private banks at the central bank; next is M1, which is currency plus check accounts but *does not include* reserve accounts; then M2, which includes M1 plus savings accounts, small (under $100,000) time deposits and individual money market deposit accounts, and finally M3 – which the US Federal Reserve no longer measures, but which is still tracked by Shadowstats – which includes M2 plus large time deposits and all money market funds.

4 It then grew at up to 2.2 percent per annum until October 1929 (the month of the stock market crash) and then turned sharply negative, falling at a rate of up to 6 percent per annum by October 1930. However, here it is quite likely that the Fed was being swamped by events, rather than being in control, as even Bernanke concedes was the case by 1931: 'As in the case of the United States, then, the story of the world monetary contraction can be summarized as "self-inflicted wounds" for the period through early 1931, and "forces beyond our control" for the two years that followed' (Bernanke 2000: 156).

5 'When prices are stable, one component of the cost [of holding money balances] is zero – namely, the annual cost – but the other component is not – namely, the cost of abstinence. This suggests that, perhaps, just as inflation produces a welfare loss, *deflation may produce a welfare gain*. Suppose therefore that we substitute a furnace for the helicopter. Let us introduce a government which imposes a tax on all individuals and burns up the proceeds, engaging in no other functions. *Let the tax be altered continuously to yield an amount that will produce a steady decline in the quantity of money at the rate of, say, 10 per cent a year*' (Friedman 1969: 16; emphases added). Friedman went on to recommend a lower rate of deflation of 5 percent for expediency reasons ('The rough estimates of the preceding section indicate that that would require for the U.S. a decline in prices at the rate of at least 5 percent per year, and perhaps decidedly more' – p. 46), but even this implied a rate of reduction of the money supply of 2

percent per annum – the same rate that he criticized the Fed for maintaining in the late 1920s.

6 These were June 1946 till January 1949, June 1950 till December 1951, 1957/58, June 1974 till June 1975, 1979–82 and December 2000 till January 2001.

7 See en.wikiquote.org/wiki/H._L._Mencken.

8 The minimum fraction that banks can hold is mandated by law, but banks can hold more than this, weakening the multiplier; and the public can decide to hang on to its cash during a financial crisis, which further weakens it. Bernanke considered both these factors in his analysis of why the Great Depression was so prolonged: 'In fractional-reserve banking systems, the quantity of inside money (M1) is a multiple of the quantity of outside money (the monetary base) […] the money multiplier depends on the public's preferred ratio of currency to deposits and the ratio of bank reserves to deposits […] sharp variations in the money multiplier […] were typically associated with banking panics, or at least problems in the banking system, during the Depression era. For example, the money multiplier in the United States began to decline precipitously following the "first banking crisis" identified by Friedman and Schwartz, in December 1930, and fell more or less continuously until the final banking crisis in March 1933, when it stabilized. Therefore, below we interpret changes in national money stocks arising from changes in the money multiplier as being caused primarily by problems in the domestic banking system' (Bernanke 2000: 125–6).

9 '[T]he reserves required to be maintained by the banking system are predetermined by the level of deposits existing two weeks earlier' (Holmes 1969: 73).

10 Such a sequence has a 1 in 4,000 chance of occurring.

11 The word 'debt' doesn't even appear in the Ireland paper, and while McKibbin and Stoeckel's model does incorporate borrowing, it plays no role in their analysis.

12 Samuel Johnson's aphorism, that something is 'like a dog's walking on his hind legs. It is not done well; but you are surprised to find it done at all,' is one of those phrases that was offensive in its origins – since Johnson used it to deride the idea of women preaching – but utterly apt in its usage today.

13 An update in February 2011 made no changes to the paper apart from adding an additional eleven works, only one of which – a 1975 paper by James Tobin – could even remotely be described as non-neoclassical.

14 I actually posted a comment to this effect on Krugman's blog when he announced that he had decided to read Minsky and had purchased this book.

15 A paper based on the model that I described in this chapter (Keen 2011) was rejected unrefereed by both the *AER* and the specialist *AER: Macroeconomics*, before being accepted by the *Journal of Economic Behavior and Organization*.

Chapter 13

1 The Revere Award recognized 'the three economists who first and most clearly anticipated and gave public warning of the Global Financial Collapse and whose work is most likely to prevent another GFC in the future.' More than 2,500 people – mainly economists – cast votes for a maximum of three out of the ninety-six candidates. I was the eventual winner with 1,152 of the 5,062 votes cast; Nouriel Roubini came second with 566 votes and Dean Baker third with 495 votes. See rwer.wordpress.com/2010/05/13/keen-roubini-and-baker-win-revere-award-for-economics-2/ for full details.

2 From Steve Keen (1995), 'Finance and economic breakdown: modeling Minsky's "Financial Instability Hypothesis"', *Journal of Post Keynesian Economics*, 17(4): 607–35. Copyright © 1995 by M. E. Sharpe, Inc. Used by permission.

3 Minsky made it into the *AER* on one other occasion, but only as a discussant of another paper at its annual conference.

4 The base model he used, known as the Hicks-Hansen-Samuelson multiplier-accelerator model, also derived its cycles from the economic error of equating an expression for actual savings with one for desired investment. See Keen (2000: 88–92).

5 This verbal model of perpetual cycles in employment and income distribution was first developed by Marx, and published in Section 1 of Chapter 25 of Volume 1 of *Capital* (Marx 1867). Marx finished his verbal model with the statement 'To put it mathematically: the rate of

accumulation is the independent, not the dependent, variable; the rate of wages, the dependent, not the independent, variable,' and it is believed that his attempt to learn calculus late in his life was motivated by the desire to express this model in mathematical form (Marx 1983 [1881]).

6 Fama and French give empirical support for this equation, which is rather ironic given their role in promoting the empirically invalid CAPM model of finance: 'These correlations confirm the impression that debt plays a key role in accommodating year-by-year variation in investment' (Fama and French 1999: 1954). In a draft version, they stated this even more clearly: 'Debt seems to be the residual variable in financing decisions. Investment increases debt, and higher earnings tend to reduce debt.'

7 I signed a contract that year with Edward Elgar Publishers to deliver a book entitled *Finance and Economic Breakdown* in 2002. That long-overdue book will hopefully be available in 2013.

8 This and later reports are downloadable from www.debtdeflation.com/blogs/pre-blog-debtwatch-reports/. I ceased writing the monthly report in April 2009, in order to devote more time to fundamental research. The blog posts, however, continued.

9 The authority here is Bill Black of the University of Missouri Kansas City, who as a public servant played a major role in enforcing the law against fraudsters in the aftermath to the Savings and Loans fiasco. See Black (2005a, 2005b); Galbraith and Black (2009).

10 When I use GDP in this context I am referring to GDP as estimated by the income measure, not the production measure.

11 The federal government's fiscal stimulus also played a major role – a topic I will consider in more detail in my next book.

12 A variable that is growing at 1 percent per annum will double in roughly seventy years, so a 3 percent rate of growth means that it will double roughly every twenty-three years.

13 It also partly reflects the impact of misguided neoclassically inspired government policies that are trying to return to 'business as usual' by encouraging private credit growth – an issue I will consider in much more detail in my next book.

14 I expect that history will judge the period from 1997 to 2009 as one continuous Ponzi scheme with two phases: the Internet Bubble and the

Subprime Bubble. A term will be needed to describe the period, and this is my nomination for it.

15 Comparing the two periods is feasible, though changes in statistical standards complicate matters. On the negative side, debt data from the 1920s (derived from the US Census) are annual, whereas those data are quarterly today, so the date of changes can't be pinpointed as well for the 1920s–1940s as for today. On the positive side, the measure of unemployment was far less distorted back then than it is today, after all the politically motivated massaging of definitions that has occurred since the mid-1970s to understate the level of unemployment in the OECD, and especially in the USA.

16 This is why the bankruptcy of Lehman Brothers was so disastrous: they had largely cornered the market for commercial paper, and when they went bankrupt this market collapsed – meaning that many ordinary firms could not pay their workers or suppliers.

Chapter 14

1 Discussed in Chapter 10.

2 The nominated policy failing this time would probably be the alleged deviation from the Taylor Rule after 2001 – the case Taylor himself is already making (see Box 10.1).

3 Seppecher's Java-based model is accessible at p.seppecher.free.fr/jamel/.

4 And they incur essentially no costs in doing so – the cost of 'producing' a dollar is much less than a dollar. This is the source of Graziani's third stricture that the system can't enable banks to exploit this opportunity for seigniorage.

5 Economists normally say 'agents' here rather than classes – given the microeconomic focus of neoclassical modeling, and the pejorative association that class was given by nineteenth-century politics. I use the term classes because social classes are an objective reality in capitalism, and because the SMD conditions, as Alan Kirman put it, suggest that 'If we are to progress further we may well be forced to theorise in terms of groups who have collectively coherent behaviour […] Thus demand and expenditure functions if they are to be set against reality must be defined at

some reasonably high level of aggregation. The idea that we should start at the level of the isolated individual is one which we may well have to abandon' (Kirman 1989: 138).

6 To register as a bank, and therefore to be able to print its own notes, 'free banking' banks still had to meet various regulatory requirements, and normally also purchase government bonds of an equivalent value to their initial printing of notes. In what follows, I'm taking these operations as given, and focusing just on the banking operations that followed incorporation.

7 My thanks to Peter Humphreys from the School of Accounting at UWS for advice on how to lay out this table in accordance with standard banking practice.

8 I have ignored interest on workers' deposit accounts simply to make the table less cluttered. They are included in my more technical description of this model in the paper 'Solving the paradox of monetary profits' (Keen 2010), which is downloadable from www.economics-ejournal.org/economics/journalarticles/2010-31.

9 See en.wikipedia.org/wiki/Time_constant for an exposition. These are normally expressed as fractions of a year – so that the assumption that workers turn their accounts over twenty-six times a year means that the time constant for workers' consumption is 1/26 – but to simplify the exposition I'm expressing them in times per year instead.

10 This point was disputed by early Circuitist literature, but this was an error of logic due to a confusion of stocks with flows (for a detailed exposition on this point, see Keen 2010: 10–12).

11 It is just a coincidence that this equals the equilibrium amount in the firms' deposit accounts – a different wage/profit share would return a different profit level.

12 A cash handout of $960 was sent to every Australian over eighteen who had a tax return for the previous year.

13 There is only a transient difference between the firm and worker bailouts on this front, while the bailout is being made. Workers' consumption is higher for the duration of the bailout if they receive the money – since they spend almost all of what they receive – but their incomes are slightly lower than when the firms get the bailout.

14 However, a more complete model is as likely to amplify these basic results as it is to attenuate them. For example, the injection of fiat money puts the banking sector's assets and liabilities out of balance, when an essential aspect of banking practice is that they are balanced. The firms bailout could thus force the banks to lend more rapidly to bring their assets back into line with their liabilities, thus amplifying the boost from the fiat money injection.

15 The equation is derived in Keen (2010: 17–19). The basic idea is as follows. The monetary value of demand equals wages plus profits, and as explained above this equals the money in the firms' deposit accounts, divided by the turnover period. The monetary value of supply is the price level times output, and output is labor times average labor productivity. The number of workers employed in turn equals the monetary value of wages divided by the wage rate. In this simple model, the monetary value of wages also depends on the balance in the firms' deposit accounts: it's equal to the amount in the firms' deposit accounts, divided by the turnover period, and multiplied by the share of surplus that goes to workers. Some cancelation yields the result that, in equilibrium, the price level will equal the wage level, divided by labor productivity and multiplied by the inverse of workers' share of surplus. A dynamic equation has prices converging to this level over time.

16 Fitting a nonlinear model to data is something mathematicians describe as a 'non-trival' exercise – which in lay-speak is something that takes eons to do and requires supercomputer processing power. I will do this for my next book with a far more complex model than the one shown here.

Chapter 15

1 A subscriber to my blog pointed out another reason: accepting the gamble involves wagering money that has taken you time and effort to earn, against the possibility of a chance gain. Most people sensibly value the effort they've put into earning something more highly than what they might get from a gamble.

2 I say 'almost' because the degree of uncertainty drops as the probability rises. If you were spinning a roulette wheel, and only one of its

thirty-eight slots would lose you money, there's far less uncertainty about the outcome of any one spin than there is with a coin toss.

3 These are obviously very similar to those used by Samuelson to derive the concept of revealed preference, but one interesting difference is that von Neumann was aware that at least the first of these was doubtful in practice. However, he argued that, if this were true, then it undermined both his approach and indifference curves:

'We have conceded that one may doubt whether a person can always decide which of two alternatives – with the utilities u, v – he prefers. But, whatever the merits of this doubt are, this possibility – i.e. the completeness of the system of (individual) preferences – must be assumed even for the purposes of the "indifference curve method." But if this property of u>v is assumed, then our use of the much less questionable [probabilistic method] yields the numerical utilities too!' (von Neumann and Morgenstern 1953: 28–9).

4 Chaos was first 'discovered' by Henri Poincaré in 1899, when he tried to find a solution to the 'many body problem' – the problem of gravitational attraction between a star and more than just one planet – and instead proved that there was no analytic solution; instead, the bodies would follow complex aperiodic paths (i.e. cycles occur which never exactly repeat themselves, unlike conventional cyclical functions like sine waves, etc.), which were later labelled 'chaotic.'

5 More complex data distributions are predicted by some more elaborate versions of the EMH, but the normal distribution is still the overall yardstick.

6 There are a number of econometric analyses that attempt to account for this. As Peters comments, they capture some of the local statistical features, but fail to capture the overall characteristics (Peters 1994).

7 Haugen is effectively a proponent of 'behavioral finance,' which has been gaining acceptance in applied and academic finance in recent years, though its adherents are still a minority compared to supporters of the Efficient Markets Hypothesis.

8 A remark that Yi-Cheng Zhang made in response to a question from Paul Ormerod as to how Econophysics came about, during a dinner at the first Econophysics conference in Bali in 2002.

9 I will cover these approaches to finance in more detail in my next book, *Finance and Economic Breakdown*.

10 The fact that unemployment to date has not reached Great Depression heights – owing in part to the under-reporting of unemployment in official statistics, as noted earlier – should be no comfort until this crisis is over and unemployment has returned to pre-crisis levels. Since the level of private debt is still enormous – 260 percent of GDP as of December 2010, 90 percent higher than the pre-Great Depression level – it is likely that this crisis has many years to run.

11 I am going somewhat beyond Kornai's logic in this paper, but in the spirit of his concepts of hard and soft budget constraints.

12 Debt that adds to the economy's productive capacity can expand this constraint over time, but Ponzi lending inflates asset prices without increasing the quantity or productivity of assets.

Chapter 16

1 This proof is very easy to understand, even if you don't think you're good at mathematics. If you assume that the square root of 2 is the ratio of two integers, then you can label these two as yet unknown integers *a* and *b*, and know that they have no common factors. Starting from the assumption that the square root of 2 equals *a* divided by *b*, you get rid of the square root by squaring both sides, so that *a* squared divided by *b* squared equals 2. This now tells you that *a* squared equals 2 times *b* squared, which is only possible if *a* is an even number – since if you square an odd number, you get another odd number. This now means that *a* has to equal two times some other number – call this *c*. Since *a* squared equals 2 times *b* squared, and *a* equal 2 times *c*, it also follows that 4 times *c* squared equals 2 times *b* squared. Divide both sides by 2 and you now find that *b* squared equals 2 times *c* squared. This means that *b* is also an even number – but this means that *a* and *b* have the common factor of 2. This contradicts your assumption that they have no common factors. Therefore the square root of 2 can't be the ratio of two integers, and it is therefore irrational.

2 This is that for changing all incomes and prices by the same factor to have no effect, 'all other nominal magnitudes [including] assets and liabilities that are expressed in nominal terms)' (Friedman 1969: 1) have to

be altered by the same factor as well – and even this ignores the fact that debt amortization makes the effect of interest rates nonlinear.

3 One example of this is the paper by Caplan (2000) which attempts to explain findings which show that experimental subjects do not conform to the neoclassical definition of rational. Rather than accepting that the neoclassical definition of rationality may be flawed, Caplan proposes that irrationality may be a 'good,' which people 'consume' like any other, and then represents a rationality–irrationality trade-off using indifference curves. This one article is not the final word on the neoclassical response to such findings. But I expect it to be far more readily adopted by the profession than any acknowledgment that the 'curse of dimensionality' makes rational behavior as economists define it simply impossible in the real world.

4 Though this branch of mathematics provides many tools which enable mathematicians to characterize the behavior of more complex and realistic models of the real world – including such things as differential equation models of the economy.

Chapter 17

1 Though economists from several other schools of thought still pay great attention to Marx's original writings on economics, and see Marx as the father of many important concepts in economic analysis.

2 This story may or may not be apocryphal. Check the website thebeadsite.com/FRO-MANG.html for one perspective, and www.crazyhorse.org/ for another.

3 Sraffa's critique of the concept of an upward-sloping demand curve, and the critiques of the market demand curve covered earlier, also undermine the neoclassical position and support the classical view.

4 The subject was a bone of contention from the time of Aristotle on. However, predecessors to the physiocrats were quite unsystematic about the determination of value and price.

5 This means that as output rose, costs of production fell. Smith was thus thinking in terms of a 'downward-sloping supply curve' – at least in the medium to long term – in contrast to the upward-sloping supply curve that is so central to economics today, which was debunked in Chapter 5.

6 All these examples were hypothetical, of course: Ricardo did not go out and measure the labor involved in producing the means of production in any industry, and then present his findings.

7 ‘Not only his labouring servants, but his labouring cattle, are productive labourers’ (Smith 1838 [1776]).

8 ‘By the invention of machinery […] a million of men may produce double, or treble the amount of riches, […] but they will on no account add anything to value’ (Ricardo 1817). Marx commented that ‘This is quite wrong. The value of the product of a million men does not depend solely on their labor but also on the value of the capital with which they work’ (Marx 1968 [1861]: Part II, p. 538).

9 Marx qualified this as ‘socially necessary labor-time,’ to take account of the possibility of out-of-equilibrium situations in which more labor-time might be lavished on a product than could be recouped by its sale.

10 There is no reason why the rate of surplus value should be constant over time in practice, and Joan Robinson used this as the basis of her critique of Marxian economics. She argued that an increase in c could cause a rise in s/v, the rate of surplus value, so that the rate of profit would not fall over time.

11 There were several counter-tendencies that could attenuate this, but ultimately Marx thought the tendency of the rate of profit to fall would prevail.

12 This is an extremely brief outline of a much more complicated argument. Its purpose is not to provide a detailed exposition of Marx’s theory of revolution, but to prepare the ground for critiques of the labor theory of value.

13 This is clearly unrealistic, but the logic is the same even if we incorporate the reality that corn would be needed to produce corn. Steedman’s example just made the numerical algebra easier to follow. He then continued his argument using symbolic linear algebra, to establish the generality of his analysis.

14 If I had worked with exact numbers rather than rounded them to two decimal places, the two calculations would have corresponded exactly. The value calculations, on the other hand, differ systematically, and by far more than can be attributed to rounding error.

15 Similar arguments had been made before, as early as at the end of the nineteenth century. Steedman simply provided the most comprehensive and definitive critique.

16 I dispute Bose's reading of Marx on this subject, but find the logic in his 'essence of value' analysis impeccable.

17 He also employed a set of axioms from which his conclusions were derived.

18 At each step in the reduction, one period's capital inputs are reduced to the previous period's direct labor and capital inputs, marked up by the equilibrium rate of profit.

19 Services such as a massage, which might appear to be a commodity-free good, involve commodities directly (massage bench, oil), and if even these are forgone (an oil-free massage while lying on bare ground), they involve it indirectly through the need for the masseur to eat to stay alive. The commodity 'massage' could therefore not be reproduced in the absence of commodity inputs, such as food.

20 Marx's philosophy was derived from Hegel's, with Marx arguing that he replaced Hegel's idealism with realism. Dialectics is popularly known as the trio of thesis-antithesis-synthesis, and though this concept is popularly associated with both Hegel and Marx, it in fact derives from another, lesser-known philosopher, Fichte. For an intelligent discussion of dialectical philosophy in general, and Marx's application of it in particular, see Wilde (1989).

21 In a different type of economy, use-value could well be brought to the foreground: commodities could be produced for the ruling elite at ostentatious expense, without regard to their cost of production. I well remember seeing a backscratcher in the Forbidden Palace in Beijing, made out of jade, gold, diamonds, emeralds and rubies.

22 This 'discovery' of the application of dialectical philosophy to economics occurred after Marx happened to re-read Hegel while he was drafting the *Grundrisse* (Oakley 1983; Mandel 1971).

23 Marx's discussion of this example still attributed the increased surplus-value to labor; however, the source of this difference was not any difference in the rate of surplus value with respect to labor employed, but to the postulate that the machine's use-value exceeded its exchange-value.

24 ‘Exchange-value and use-value [are] intrinsically incommensurable magnitudes’ (Marx 1867). Notice that Marx describes use-value as a magnitude in this circumstance. Outside production, when commodities are purchased to be consumed rather than being used to produce other commodities, their use-value will be qualitative, and therefore incommensurable with their exchange-values.

Chapter 18

1 This is a necessarily brief and personally opinionated survey of five very complex schools of thought. Readers who wish to delve deeper should consult the references given in this chapter. I have also omitted separate discussion of a notable school of economic thought, institutional economics, because I expect it to be subsumed under evolutionary economics.

2 Though evolutionary theorists themselves now argue that Darwin’s vision of the evolutionary process, in which ‘nature did not make jumps,’ is flawed, and that therefore Darwinism is an inappropriate label for modern evolutionary theory (Schwartz 2000).

3 And, ironically, some evolutionary theorists are now arguing that biological evolution may in some ways be purposive (McFadden 2001).

BIBLIOGRAPHY

60 Minutes (2008) 'Chat: Professor Steve Keen,' sixtyminutes.ninemsn.com.au/webchats/643288/chat-professor-steve-keen.

7.30 Report (2007) 'Web extra: extended interview with Assoc. Prof. Steve Keen,' www.abc.net.au/7.30/content/2007/s2006034.htm.

Albert, D. Z. (1992) *Quantum Mechanics and Experience*, Cambridge, MA: Harvard University Press.

Albert, D. Z. (1994) 'Bohm's alternative to quantum mechanics,' *Scientific American*, 270(5): 32–9.

American Review of Political Economy (2011), www.arpejournal.com, forthcoming.

Anderson, P. W. (1972) 'More is different,' *Science*, 177(4047): 393–6.

Arestis, P., S. P. Dunn and M. Sawyer (1999) 'Post Keynesian economics and its critics,' *Journal of Post Keynesian Economics*, 21: 527–49.

Arrow, K. J., M. D. Intriligator et al. (1982) *Handbook of Mathematical Economics*, Amsterdam: Elsevier.

Baird, W. C. (1981) *Elements of Macroeconomics*, New York: West St Paul.

Ballard, D. H. (2000) *An Introduction to Natural Computation*, Cambridge, MA: MIT Press.

Barber, W. J. (ed.) (1997) *The Works of Irving Fisher*, London: Pickering and Chatto.

Barbier, E. B. (ed.) (1993) *Economics and Ecology*, London: Chapman & Hall.

Barnett, W. (1999) 'A single-blind controlled competition among tests for nonlinearity and chaos,' *Journal of Econometrics*, 82: 157–92.

Barnett, W. A. (1979) 'Theoretical foundations for the Rotterdam model,' *Review of Economic Studies*, 46: 109–30.

Barnett, W. A., C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds) (2000) *Commerce, Complexity and Evolution*, New York: Cambridge University Press.
Barr, J. M., T. Tassier et al. (2008) 'Symposium on agent-based computational economics: introduction,' *Eastern Economic Journal*, 34(4): 421–2.
Battalio, R. C., J. H. Kagel, R. C. Winkler, E. B. Fisher, R. L. Bassmann and L. Krasner (1977) 'A test of consumer demand theory using observations of individual consumer purchases,' *Western Economic Journal*, 11: 411–28.
Bell, D. and I. Kristol (1981) *The Crisis in Economic Theory*, New York: Basic Books.
Bentham, J. (1787) *In Defence of Usury*, socserv2.socsci.mcmaster.ca:80/~econ/ugcm/3ll3/bentham/usury.
Bentham, J. (1948 [1780]) *The Principles of Morals and Legislation*, New York: Hafner Press.
Bernanke, B. S. (2000) *Essays on the Great Depression*, Princeton, NJ: Princeton University Press.
Bernanke, B. S. (2002a) *Deflation: Making Sure 'It' Doesn't Happen Here*, Washington, DC: Federal Reserve Board.
Bernanke, B. S. (2002b) 'Remarks by Governor Ben S. Bernanke at the Conference to Honor Milton Friedman,' *Conference to Honor Milton Friedman*, Chicago, IL: University of Chicago.
Bernanke, B. S. (2004a) 'Panel discussion: what have we learned since October 1979?' *Conference on Reflections on Monetary Policy 25 Years after October 1979*, St Louis, MI: Federal Reserve Bank of St Louis.
Bernanke, B. S. (2004b) 'The Great Moderation: remarks by Governor Ben S. Bernanke at the meetings of the Eastern Economic Association, Washington, DC, February 20, 2004,' *Eastern Economic Association*, Washington, DC: Federal Reserve Board.
Bernanke, B. S. (2010) 'On the implications of the financial crisis for economics,' *Conference Co-sponsored by the Center for Economic Policy Studies and the Bendheim Center for Finance, Princeton University*, Princeton, NJ: US Federal Reserve.
Bernanke, B. S. and M. Gertler (1989) 'Agency costs, net worth and business fluctuations,' *American Economic Review*, 79: 14–31.

Besley, T. and P. Hennessy (2009) *Letter to Her Majesty the Queen about the 'The Global Financial Crisis – Why Didn't Anybody Notice?'* London: London School of Economics.

Bezemer, D. J. (2009) *'No One Saw This Coming': Understanding Financial Crisis through Accounting Models*, Groningen: Faculty of Economics, University of Groningen.

Bezemer, D. J. (2010) 'Understanding financial crisis through accounting models,' *Accounting, Organizations and Society*, 35(7): 676–88.

Bezemer, D. J. (2011) 'The credit crisis and recession as a paradigm test,' *Journal of Economic Issues*, 45: 1–18.

Bhaduri, A. (1969) 'On the significance of recent controversies in capital theory: a Marxian view,' *Economic Journal*, 79: 532–9.

Bharadwaj, K. and B. Schefold (eds) (1990) *Essays on Pierro Sraffa: Critical Perspectives on the Revival of Classical Theory*, London: Unwin Hyman.

Biggs, M., T. Mayer et al. (2010) 'Credit and economic recovery: demystifying phoenix miracles,' SSRN eLibrary.

Bishop, R. L. (1948) 'Cost discontinuities, declining costs and marginal analysis,' *American Economic Review*, 38: 607–17.

Black, W. K. (2005a) '"Control frauds" as financial super-predators: how "pathogens" make financial markets inefficient,' *Journal of Socio-Economics*, 34(6): 734–55.

Black, W. K. (2005b) *The Best Way to Rob a Bank Is to Own One: How Corporate Executives and Politicians Looted the S&L Industry*, Austin: University of Texas Press.

Blanchard, O. J. (2008) 'The state of macro,' SSRN eLibrary.

Blanchard, O. J. (2009) 'The state of macro,' *Annual Review of Economics*, 1(1): 209–28.

Blatt, J. M. (1983) *Dynamic Economic Systems: A post-Keynesian approach*, Armonk, NY: M. E. Sharpe.

Blaug, M. (1998) 'Disturbing currents in modern economics,' *Challenge!*, 41(3): 11–34.

Blinder, A. S. (1982) 'Inventories and sticky prices: more on the microfoundations of macroeconomics,' *American Economic Review*, 72(3): 334–48.

Blinder, A. S. (1998) *Asking about Prices: A new approach to understanding price stickiness*, New York: Russell Sage Foundation.
Blodget, H. (2010) '10 Years after NASDAQ 5000, Henry Blodget reflects,' finance.yahoo.com/tech-ticker/article/440898/10-Years-After-NASDAQ-5000,-Henry-Blodget-Reflects.
Böhm-Bawerk, E. (1949 [1896]) *Karl Marx and the Close of His System*, ed. P. Sweezy, New York: Orion Editions.
Bond, N. W. (2000) 'Psychology: a science of many faces,' in N. W. Bond and K. M. McConkey (eds), *An Introduction to Psychological Science*, Sydney: McGraw-Hill.
Bose, A. (1980) *Marx on Exploitation and Inequality*, Delhi: Oxford University Press.
Bowles, H. and S. Gintis (1993) 'The revenge of Homo Economicus: contested exchange and the revival of political economy,' *Journal of Economic Perspectives*, 7(1): 83–102.
Boyd, I. and J. M. Blatt (1988) *Investment Confidence and Business Cycles*, Berlin: Springer.
Braun, M. (1993) *Differential Equations and Their Applications*, New York: Springer-Verlag.
Caldwell, B. (ed.) (1984) *Appraisal and Criticism in Economics: A Book of Readings*, London: Allen & Unwin.
Caldwell, B. J. and S. Boehm (eds) (1992) *Austrian Economics: Tensions and New Directions*, Boston, MA: Kluwer Academic.
Caplan, B. (2000) 'Rational irrationality: a framework for the neoclassical-behavioral debate,' *Eastern Economic Journal*, 26: 191–212.
Carpenter, S. B. and S. Demiralp (2010) *Money, Reserves, and the Transmission of Monetary Policy: Does the Money Multiplier Exist?*, Finance and Economics Discussion Series, Washington, DC: Federal Reserve Board.
Carter, J. R. and M. D. Irons (1991) 'Are economists different, and if so, why?' *Journal of Economic Perspectives*, 5(2): 171–7.
Chiarella, C. et al. (2003) 'Asset price dynamics among heterogeneous interacting agents,' *Computational Economics*, 22(2/3): 213–23.
Chiarella, C. and P. Flaschel (2000) *The Dynamics of Keynesian Monetary Growth*, Cambridge: Cambridge University Press.

Chiarella, C., R. Dieci et al. (2002) 'Speculative behaviour and complex asset price dynamics: a global analysis,' *Journal of Economic Behavior and Organization*, 49(2): 173–97.
Chipman, J. S. (1974) 'Homothetic preferences and aggregation,' *Journal of Economic Theory*, 8: 26–38.
Clapham, J. H. (1922a) 'Of empty economic boxes,' *Economic Journal*, 32: 303–14.
Clapham, J. H. (1922b) 'The economic boxes – a rejoinder,' *Economic Journal*, 32: 560–3.
Clark, J. B. (1898) 'The future of economic theory,' *Quarterly Journal of Economics*, 13: 1–14.
Clower, R. W. (1969) 'The Keynesian counter-revolution: a theoretical appraisal,' in R. W. Clower, *Monetary Theory*, Harmondsworth: Penguin.
Clower, R. W. and A. Leijonhufvud (1973) 'Say's Principle, what it means and doesn't mean: Part I,' *Intermountain Economic Review*, 4(2): 1–16.
Coddington, A. (1976) 'Keynesian economics: the search for first principles,' *Journal of Economic Literature*, 14(4): 1258–73.
Colander, D. (2011) 'Is the fundamental science of macroeconomics sound?' American Economic Association Annual Conference, Denver, CO.
Costanza, R. (1993) 'Ecological economic systems analysis: order and chaos,' in E. B. Barbier, *Economics and Ecology*, London: Chapman & Hall, pp. 29–45.
Cotis, J.-P. (2007) 'Editorial: achieving further rebalancing,' *OECD Economic Outlook*, 2007/1: 7–10.
Crouch, R. (1972) *Macroeconomics*, New York: Harcourt Brace Jovanovich.
Debreu, G. (1959) *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, New Haven, CT: Yale University Press.
Debreu, G. (1970) 'Economies with a finite set of equilibria,' *Econometrica*, 38: 387–92.
Debreu, G. (1974) 'Excess demand functions,' *Journal of Mathematical Economics*, 1: 15–21.
Desai, M. (1981) *Testing Monetarism*, London: Frances Pinter.

Dierker, E. (1972) 'Two remarks on the number of equilibria of an economy,' *Econometrica*, 40: 951–3.
Diewert, W. E. (1977) 'Generalized Slutsky conditions for aggregate consumer demand functions,' *Journal of Economic Theory*, 15(2): 353–62.
Dillard, D. (1984) 'Keynes and Marx: a centennial appraisal,' *Journal of Post Keynesian Economics*, 6(3): 421–32.
Ding, N., N. Xi et al. (2006) 'The economic mobility in money transfer models,' *Physica A: Statistical Mechanics and Its Applications*, 367: 415–24.
Dixon, R. (2000a) *A Formal Proof of Walras' Law*, www.ecom.unimelb.edu.au/ ecowww/rdixon/walproof.html.
Dixon, R. (2000b) *Walras' Law and Macroeconomics*, www.ecom.unimelb.edu.au/ecowww/rdixon/wlaw.html.
Dow, S. C. (1997) 'Endogenous money,' in G. C. Harcourt and P. A. Riach (eds), *A 'Second Edition' of the General Theory*, London, Routledge, pp. 61–78.
Downward, P. (1999) *Pricing Theory in Post Keynesian Economics: A realist approach*, Cheltenham: Edward Elgar.
Downward, P. and P. Reynolds (1999) 'The contemporary relevance of Post-Keynesian economics: editors' introduction,' *Economic Issues*, 4: 1–6.
Dumenil, G. and D. Levy (1985) 'The classicals and the neo-classicals: a rejoinder to Frank Hahn,' *Cambridge Journal of Economics*, 9: 327–45.
Earl, P. E. (1995) *Microeconomics for Business and Marketing*, Cheltenham: Edward Elgar.
Ehrenberg, A. S. C. (1975) *Data Reduction: Analysing and interpreting statistical data*, London: Wiley.
Einstein, A. (1961 [1916]) *Relativity: The Special and the General Theory*, New York: Random House.
Eiteman, W. J. (1947) 'Factors determining the location of the least cost point,' *American Economic Review*, 37: 910–18.
Eiteman, W. J. (1948) 'The least cost point, capacity and marginal analysis: a rejoinder,' *American Economic Review*, 38: 899–904.
Eiteman, W. J. and G. E. Guthrie (1952) 'The shape of the average cost curve,' *American Economic Review*, 42(5): 832–8.

Fama, E. F. (1970) 'Efficient capital markets: a review of theory and empirical work,' *Journal of Finance*, 25(2): 383–417.
Fama, E. F. and K. R. French (1999) 'The corporate cost of capital and the return on corporate investment,' *Journal of Finance*, 54(6): 1939–67.
Fama, E. F. and K. R. French (2004) 'The Capital Asset Pricing Model: theory and evidence,' *Journal of Economic Perspectives*, 18(3): 25–46.
Feher, D. C. (1999) *Debt Deflation: The Birth of a Concept and Its Development over Time*, Unpublished honors thesis, University of Western Sydney.
Financial Crisis Inquiry Commission (2011) *The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*.
Fisher, I. (1929) 'Transcript of an address of Professor Irving Fisher,' in W. J. Barber (ed.) (1997), *The Works of Irving Fisher*, vol. 10, London: Pickering and Chatto.
Fisher, I. (1930) *The Theory of Interest*, New York: Macmillan, reprinted in W. J. Barber (ed.) (1997), *The Works of Irving Fisher*, vol. 3, London: Pickering and Chatto.
Fisher, I. (1932) *Booms and Depressions: Some First Principles*, reprinted in W. J. Barber (ed.), *The Works of Irving Fisher*, vol. 10, London: Pickering and Chatto.
Fisher, I. (1933) 'The debt-deflation theory of great depressions,' *Econometrica*, 1: 337–55.
Frank, R. H., T. Gilovich and D. T. Regan (1993) 'Does studying economics inhibit cooperation?' *Journal of Economic Perspectives*, 7(2): 159–71.
Frank, R. H., T. Gilovich and D. T. Regan (1996) 'Do economists make bad citizens?' *Journal of Economic Perspectives*, 10(1): 187–92.
Franklin, J. and A. Daoud (1988) *Introduction to Proofs in Mathematics*, New York: Prentice-Hall.
Freeman, A. and G. Carchedi (1996) *Marx and Non-Equilibrium Economics*, Cheltenham: Edward Elgar.
Freeman, S. and F. E. Kydland (2000) 'Monetary aggregates and output,' *American Economic Review*, 90(5): 1125–35.
Friedman, M. (1953) 'The methodology of positive economics,' reprinted in B. Caldwell (ed.) (1984), *Appraisal and Criticism in Economics: A*

Book of Readings, London: Allen & Unwin.
Friedman, M. (1968) 'The role of monetary policy,' *American Economic Review*, 58(1): 1–17.
Friedman, M. (1969) 'The optimum quantity of money,' in *The Optimum Quantity of Money and Other Essays*, Chicago, IL: Macmillan, pp. 1–50.
Friedman, M. (1971) 'A monetary theory of nominal income,' *Journal of Political Economy*, 79(2): 323–37.
Frisch, R. (1933) 'Propagation problems and impulse problems in dynamic economics,' in *Economic Essays in Honour of Gustav Cassel*, London: George Allen & Unwin Ltd.
Fullbrook, E. (2010) 'Keen, Roubini and Baker win Revere Award for Economics,' *Real-World Economics Review* blog, New York.
Gabaix, X., P. Gopikrishnan et al. (2006) 'Institutional investors and stock market volatility,' *Quarterly Journal of Economics*, 121(2): 461–504.
Galbraith, J. K. (1997) *The Socially Concerned Today: The First Honorary Keith Davey Lecture*, Toronto: University of Toronto Press.
Galbraith, J. K. and W. K. Black (2009) 'Trust but verify,' in K. van den Heuvel (ed.), *Meltdown: How Greed and Corruption Shattered Our Financial System and How We Can Recover*, New York: Nation Books, pp. 244–7.
Gallegati, M., S. Keen et al. (2006) 'Worrying trends in econophysics,' *Physica A: Statistical Mechanics and Its Applications*, 370(1): 1–6.
Glassman, J. K. and K. A. Hassett (1999) *DOW 36,000: The New Strategy for Profiting from the Coming Rise in the Stock Market*, New York: Times Business.
Goodwin, R. (1967) 'A growth cycle,' in C. H. Feinstein (ed.), *Socialism, Capitalism and Economic Growth*, Cambridge: Cambridge University Press, pp. 54–8.
Goodwin, R. M. (1986) 'The economy as an evolutionary pulsator,' *Journal of Economic Behavior and Organization*, 7(4): 341–9.
Goodwin, R. M. (1990) *Chaotic Economic Dynamics*, Oxford: Oxford University Press.
Goodwin, R. M. (1991) 'New results in non-linear economic dynamics,' *Economic Systems Research*, 3(4): 426–7.

Gordon, R. J. (1976) 'Can econometric policy evaluations be salvaged? – a comment,' *Carnegie-Rochester Conference Series on Public Policy*, pp. 47–61.
Gorman, W. M. (1953) 'Community preference fields,' *Econometrica*, 21(1): 63–80.
Graziani, A. (1989) 'The theory of the monetary circuit,' *Thames Papers in Political Economy*, Spring, pp. 1–26.
Groll, S. (1980) 'The active role of "use value" in Marx's economics,' *History of Political Economy*, 12(3): 336–71.
Hahn, F. (1982) 'The neo-Ricardians,' *Cambridge Journal of Economics*, 6: 353–74.
Haines, W. W. (1948) 'Capacity production and the least cost point,' *American Economic Review*, 38: 617–24.
Harcourt, G. (1982) *The Social Science Imperialists*, ed. P. Kerr, London: Routledge & Kegan Paul.
Harford, T. (2005) *The Undercover Economist*, London: Oxford University Press.
Harrod, R. (1939) 'An essay in dynamic theory,' *Economic Journal*, 49: 14–33.
Haugen, R. A. (1999a) *The Beast on Wall Street*, New Jersey: Prentice-Hall.
Haugen, R. A. (1999b) *The New Finance*, New Jersey: Prentice-Hall.
Haugen, R. A. (1999c) *The Inefficient Stock Market*, New Jersey: Prentice-Hall.
Henwood, D. (1997) *Wall Street*, New York: Verso.
Hicks, J. R. (1935) 'Wages and interest: the dynamic problem,' *Economic Journal*, 45(179): 456–68.
Hicks, J. R. (1937) 'Mr. Keynes and the "Classics": a suggested interpretation,' *Econometrica*, 5(2): 147–59.
Hicks, J. R. (1949) 'Mr. Harrod's dynamic theory,' *Economica*, 16(62): 106–21.
Hicks, J. R. (1979) 'On Coddington's interpretation: a reply,' *Journal of Economic Literature*, 17(3): 989–95.
Hicks, J. R. (1980) 'IS–LM: an explanation,' *Journal of Post Keynesian Economics*, 3(2): 139–54.
Hirshleifer, J. (1993) 'The dark side of the force,' *Economic Inquiry*, 32: 1–10.

Hodgson, G. (1991) *After Marx and Sraffa*, New York: St Martin's Press.
Hodgson, G. M. (1998) *The Foundations of Evolutionary Economics*, Cheltenham: Edward Elgar.
Hodgson, G. M. (1999) *Evolution and Institutions: On Evolutionary Economics and the Evolution of Economics*, Cheltenham: Edward Elgar.
Hodgson, G. M., W. J. Samuels and M. R. Tool (eds) (1994) *The Elgar Companion to Institutional and Evolutionary Economics*, Aldershot: Edward Elgar.
Hogan, C. J., R. P. Kirshner and N. B. Suntzeff (1999) 'Surveying space–time with supernovae,' *Scientific American*, 280(1): 28–33.
Holmes, A. R. (1969) 'Operational constraints on the stabilization of money supply growth,' in F. E. Morris (ed.), *Controlling Monetary Aggregates*, Nantucket Island, MA: Federal Reserve Bank of Boston, pp. 65–77.
Hudson, M. (2000) 'The mathematical economics of compound interest: a 4,000-year overview,' *Journal of Economic Studies*, 27(4/5): 344–63.
Institute of Chartered Financial Analysts (eds) (1991) *The Founders of Modern Finance: Their Prize-winning Concepts and 1990 Nobel Lectures*, Charlottesville, VA: Institute of Chartered Financial Analysts.
Ireland, P. N. (2011) 'A new Keynesian perspective on the Great Recession,' *Journal of Money, Credit and Banking*, 43(1): 31–54.
Jevons, W. S. (1866) 'Brief account of a general mathematical theory of political economy,' *Journal of the Royal Statistical Society, London*, 29: 282–7, www.marxists.org/reference/subject/economics/jevons/mathem.htm.
Jevons, W. S. (1888) *The Theory of Political Economy*, Internet: Library of Economics and Liberty, www.econlib.org/library/YPDBooks/Jevons/jvnPE4.html.
Jones, N. L. (1989) *God and the Moneylenders*, Oxford: Basil Blackwell.
Jorgenson, D. W. (1960) 'A dual stability theorem,' *Econometrica*, 28: 892–9.
Jorgenson, D. W. (1961) 'Stability of a dynamic input-output system,' *Review of Economic Studies*, 28: 105–16.
Jorgenson, D. W. (1963) 'Stability of a dynamic input-output system: a reply,' *Review of Economic Studies*, 30: 148–9.

Kaldor, N. (1982) *The Scourge of Monetarism*, Oxford: Oxford University Press.
Kaletsky, A. (2009) 'Now is the time for a revolution in economic thought,' *The Times*, London, 9 February.
Kariya, T. (1993) *Quantitative Methods for Portfolio Analysis: MTV Model Approach*, Dordrecht: Kluwer Academic Publishers.
Kates, S. (1998) *Say's Law and the Keynesian Revolution*, Cheltenham: Edward Elgar.
Kates, S. (2003) 'Economic management and the Keynesian revolution: the policy consequences of the disappearance of Say's Law,' in *Two Hundred Years of Say's Law*, Cheltenham: Edward Elgar.
Keen, S. (1993a) 'Use-value, exchange-value, and the demise of Marx's labor theory of value,' *Journal of the History of Economic Thought*, 15: 107–21.
Keen, S. (1993b) 'The misinterpretation of Marx's theory of value,' *Journal of the History of Economic Thought*, 15: 282–300.
Keen, S. (1995) 'Finance and economic breakdown: modeling Minsky's "Financial Instability Hypothesis,"' *Journal of Post Keynesian Economics*, 17(4): 607–35.
Keen, S. (1996) 'The chaos of finance: the chaotic and Marxian foundations of Minsky's "Financial Instability Hypothesis,"' *Economies et Sociétés*, 30(2/3): 55–82.
Keen, S. (1998) 'Answers (and questions) for Sraffians (and Kaleckians),' *Review of Political Economy*, 10: 73–87.
Keen, S. (2000) 'The nonlinear dynamics of debt-deflation,' in W. A. Barnett, C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds), *Commerce, Complexity and Evolution*, New York: Cambridge University Press.
Keen, S. (2001a) *Debunking Economics: The naked emperor of the social sciences*, Annandale, Sydney/London: Pluto Press Australia/Zed Books.
Keen, S. (2001b) 'Minsky's thesis: Keynesian or Marxian?' in R. Bellofiore and P. Ferri (eds), *The Economic Legacy of Hyman Minsky*, vol. 1: *Financial Keynesianism and Market Instability*, Cheltenham: Edward Elgar, pp. 106–20.
Keen, S. (2003) 'Standing on the toes of pygmies: why econophysics must be careful of the economic foundations on which it builds,' *Physica A:*

Statistical Mechanics and Its Applications, 324(1/2): 108–16.
Keen, S. (2004) 'Deregulator: Judgment Day for microeconomics,' *Utilities Policy*, 12: 109–25.
Keen, S. (2005) 'Why economics must abandon its theory of the firm,' in M. Salzano and A. Kirman (eds), *Economics: Complex Windows*, New Economic Windows series, Milan and New York: Springer, pp. 65–88.
Keen, S. (2006) Steve Keen's Monthly Debt Report, 'The recession we can't avoid?' *Steve Keen's Debtwatch*, Sydney, 1: 21, November.
Keen, S. (2008) 'Keynes's "revolving fund of finance" and transactions in the circuit,' in R. Wray and M. Forstater (eds), *Keynes and Macroeconomics after 70 Years*, Cheltenham: Edward Elgar, pp. 259–78.
Keen, S. (2009a) 'A pluralist approach to microeconomics,' in J. Reardon (ed.), *The Handbook of Pluralist Economics Education*, London: Routledge, pp. 120–49.
Keen, S. (2009b) 'The dynamics of the monetary circuit,' in S. Rossi and J.-F. Ponsot (eds), *The Political Economy of Monetary Circuits: Tradition and Change*, London: Palgrave Macmillan, pp. 161–87.
Keen, S. (2010) 'Solving the paradox of monetary profits,' *Economics: The Open-Access, Open Assessment E-Journal*, 4(2010-31).
Keen, S. (2011) 'A monetary Minsky model of the Great Moderation and the Great Recession,' *Journal of Economic Behavior and Organization*, forthcoming.
Keen, S. and E. Fullbrook (2004) 'Improbable, incorrect or impossible? The persuasive but flawed mathematics of microeconomics,' in *A Guide to What's Wrong with Economics*, London: Anthem Press, pp. 209–22.
Keen, S. and R. Standish (2006) 'Profit maximization, industry structure, and competition: a critique of neoclassical theory,' *Physica A: Statistical Mechanics and Its Applications*, 370(1): 81–5.
Keen, S. and R. Standish (2010) 'Debunking the theory of the firm – a chronology,' *Real-World Economics Review*, 54(54): 56–94.
Kehoe, T. J. and E. C. Prescott (2002) 'Great Depressions of the 20th century,' *Review of Economic Dynamics*, 5(1): 1–18.
Keynes, J. M. (1925) *Essays in Persuasion*, London: Macmillan for the Royal Economic Society.

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money*, London: Macmillan.
Keynes, J. M. (1937) 'The General Theory of Employment,' *Quarterly Journal of Economics*, 51(2): 209–23.
Keynes, J. M. (1971 [1923]) 'A tract on monetary reform,' in *The Collected Works of John Maynard Keynes*, vol. IV, London: Macmillan.
Keynes, J. M. (1972 [1925]) 'A short view of Russia', in J. M. Keynes, *Essays in Persuasion*, Cambridge: Cambridge University Press.
Kirman, A. (1989) 'The intrinsic limits of modern economic theory: the emperor has no clothes,' *Economic Journal*, 99(395): 126–39.
Kirman, A. P. (1992) 'Whom or what does the representative individual represent?' *Journal of Economic Perspectives*, 6(2): 117–36.
Kirzner, I. M. (1996) *Essays on Capital and Interest*, Cheltenham: Edward Elgar.
Klein, L. R. (1950) *Economic Fluctuations in the United States 1921–1941*, New York: John Wiley & Sons.
Klein, P. A. (1994) 'The reassessment of institutionalist mainstream relations,' *Journal of Economic Issues*, 28: 197–207.
Koo, R. (2009) *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession*, New York: John Wiley & Sons.
Kornai, J. (1979) 'Resource-constrained versus demand-constrained systems,' *Econometrica*, 47(4): 801–19.
Kornai, J. (1986) 'The soft budget constraint,' *Kyklos*, 39(1): 3–30.
Kornai, J. (1990) *Economics of Shortage*, Amsterdam: North-Holland.
Kornai, J., E. Maskin et al. (2003) 'Understanding the soft budget constraint,' *Journal of Economic Literature*, 41(4): 1095–136.
Kregel, J. A. (1983) 'Post-Keynesian theory: an overview,' *Journal of Economic Education*, 14(4): 32–43.
Kreps, D. M. (1990) *A Course in Microeconomic Theory*, Princeton, NJ: Princeton University Press.
Krugman, P. (1996) 'What economists can learn from evolutionary theorists,' web.mit.edu/krugman/www/evolute.html.
Krugman, P. (2009a) 'A Dark Age of macroeconomics (wonkish),' in *The Conscience of a Liberal*, New York: New York Times.
Krugman, P. (2009b) 'How did economists get it so wrong?' in *The Conscience of a Liberal*, New York: New York Times.

Krugman, P. and G. B. Eggertsson (2010) *Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo approach*, 2nd draft 14 February 2011, New York: Federal Reserve Bank of New York and Princeton University, www.princeton.edu/~pkrugman/debt_deleveraging_ge_pk.pdf.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago, IL: University of Chicago Press.

Kydland, F. E. and E. C. Prescott (1990) 'Business cycles: real facts and a monetary myth,' *Federal Reserve Bank of Minneapolis Quarterly Review*, 14(2): 3–18.

Kydland, F. E. and E. C. Prescott (1991) 'The econometrics of the general equilibrium approach to business cycles,' *Scandinavian Journal of Economics*, 93(2): 161–78.

Labaton, S. (1999) 'Congress passes wide-ranging bill easing bank laws,' *New York Times*, 5 November, p. 2.

Lakatos, I. (1978) *The Methodology of Scientific Research Programs*, ed. J. Worrall and G. Currie, New York: Cambridge University Press.

Lancaster, K. and R. G. Lipsey (1956) 'The general theory of the second best,' *Review of Economic Studies*, 24: 11–32, reprinted in K. Lancaster (1996), *Trade, Markets and Welfare*, Cheltenham: Edward Elgar.

Langlois, C. (1989) 'Markup pricing versus marginalism: a controversy revisited,' *Journal of Post Keynesian Economics*, 12: 127–51.

Lavoie, M. (1994) 'A Post Keynesian approach to consumer choice,' *Journal of Post Keynesian Economics*, 16: 539–62.

Lee, F. (1996) 'Pricing and the business enterprise,' in C. W. Whalen (ed.), *Political Economy for the 21st century*, Armonk, NY: M. E. Sharpe.

Lee, F. (1998) *Post Keynesian Price Theory*, Cambridge: Cambridge University Press.

Leeson, R. (1991) 'The validity of the expectations-augmented Phillips curve model,' *Economic Papers*, 10: 92–6.

Leeson, R. (1994) 'A. W. H. Phillips M.B.E. (Military Division),' *Economic Journal*, 104(424): 605–18.

Leeson, R. (1997) 'The trade-off interpretation of Phillips's Dynamic Stabilization Exercise,' *Economica*, 64(253): 155–71.

Leeson, R. (1998) 'The origins of Keynesian discomfiture,' *Journal of Post Keynesian Economics*, 20: 597–619.

Leeson, R. (2000) *A. W. H. Phillips: Collected works in contemporary perspective*, Cambridge, New York and Melbourne: Cambridge University Press.
Leijonhufvud, A. (1968) *On Keynesian Economics and the Economics of Keynes: A Study in Monetary Theory*, New York: Oxford University Press.
Leijonhufvud, A. (1973) 'Life among the Econ,' *Western Economic Journal*, 11(3): 327–37, republished in C. P. Clotfelter (ed.) (1996), *On the Third Hand: Humor in the Dismal Science*, Ann Arbor: University of Michigan Press, pp. 24–35.
Leijonhufvud, A. (1986) 'What would Keynes have thought of rational expectations?' in J. L. Butkiewicz, K. J. Koford and J. B. Miller (eds), *Keynes' Economic Legacy: Contemporary Economic Theories*, New York: Praeger, pp. 25–52.
Leijonhufvud, A. (1991 [1967]) 'Keynes and the Keynesians: a suggested interpretation,' in E. Phelps (ed.), *Recent Developments in Macroeconomics*, Aldershot: Edward Elgar.
Levins, R. and R. C. Lewontin (1986) *The Dialectical Biologist*, London: Harvard University Press.
Levitt, S. D. and S. J. Dubner (2009) *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*, New York: Harper Perennial.
Li, T.-Y. and J. A. Yorke (1975) 'Period three implies chaos,' *American Mathematical Monthly*, 82(10): 985–92.
Lindsey, D. E., A. Orphanides et al. (2005) 'The reform of October 1979: how it happened and why,' *Federal Reserve Bank of St Louis Review*, 87(2): 187–235.
Littlefield, S. (ed.) (1990) *Austrian Economics*, Aldershot: Edward Elgar.
Ljungqvist, L. and T. J. Sargent (2004) *Recursive Macroeconomic Theory*, 2nd edn, Cambridge, MA: MIT Press.
Lorenz, H.-W. (1987a) 'Goodwin's nonlinear accelerator and chaotic motion,' *Zeitschrift fur Nationalokonomie* [Journal of Economics], 47(4): 413–18.
Lorenz, H.-W. (1987b) 'Strange attractors in a multisector business cycle model,' *Journal of Economic Behavior and Organization*, 8(3): 397–411.

Lorenz, H.-W. (1989) *Nonlinear Dynamical Economics and Chaotic Motion (Lecture Notes in Economics and Mathematical Systems)*, Berlin: Springer.
Lucas, R. E., Jr (1972) 'Econometric testing of the Natural Rate Hypothesis,' *The Econometrics of Price Determination Conference, October 30–31 1970*, Washington, DC: Board of Governors of the Federal Reserve System and Social Science Research Council, pp. 50–9.
Lucas, R. E., Jr (1976) 'Econometric policy evaluation: a critique,' *Carnegie-Rochester Conference Series on Public Policy*, 1: 19–46.
Lucas, R. E., Jr (2003) 'Macroeconomic priorities,' *American Economic Review*, 93(1): 1–14.
Lucas, R. E., Jr (2004) 'Keynote address to the 2003 HOPE Conference: my Keynesian education,' *History of Political Economy*, 36: 12–24.
Mackay, C. (1841) *Extraordinary Popular Delusions and the Madness of Crowds*, New York: Crown Trade Paperbacks, www.litrix.com/madraven/madne001.htm.
Mandel, E. (1971) *The Formation of the Economic Thought of Karl Marx*, London: NLB.
Mandelbrot, B. (1971) 'Linear regression with non-normal error terms: a comment,' *Review of Economics and Statistics*, 53(2): 205–6.
Mandelbrot, B. (2005) 'The inescapable need for fractal tools in finance,' *Annals of Finance*, 1(2): 193–5.
Mandelbrot, B. B. and R. L. Hudson (2004) *The (Mis)behaviour of Markets: A fractal view of risk, ruin and reward*, London: Profile.
Mankiw, N. G. (2008) *Principles of Microeconomics, 5E*, Boston, MA: South-Western College Publishers.
Mantel, R. R. (1974) 'On the characterisation of aggregate excess demand,' *Journal of Economic Theory*, 7: 348–53.
Mantel, R. R. (1976) 'Homothetic preferences and community excess demand functions,' *Journal of Economic Theory*, 12: 197–201.
Marks, R. (2000) 'Evolved perception and the validation of simulation models,' in W. A. Barnett, C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds), *Commerce, Complexity and Evolution*, New York: Cambridge University Press.

Marshall, A. (1920 [1890]) *Principles of Economics*, Internet: Library of Economics and Liberty, www.econlib.org/library/Marshall/marP0.html.
Martin, S. (2000) *Advanced Industrial Economics*, Oxford: Basil Blackwell.
Marwell, G. and R. Ames (1981) 'Economists free-ride, does anyone else?' *Journal of Public Economics*, 15(3): 295–310.
Marx, K. (1847) *Wage Labor and Capital*, Moscow: Progress Press.
Marx, K. (1857) *Grundrisse*, Harmondsworth: Penguin.
Marx, K. (1859) *A Contribution to the Critique of Political Economy*, Moscow: Progress Press.
Marx, K. (1867) *Capital*, vol. 1, Moscow: Progress Press.
Marx, K. (1885) *Capital*, vol. 2, Moscow: Progress Press.
Marx, K. (1894) *Capital*, vol. 3, Moscow: Progress Press.
Marx, K. (1951 [1865]) 'Wages, price and profit,' in Marx-Engels-Lenin Institute (ed.), *Marx-Engels Selected Works*, vol. I, Moscow: Foreign Languages Publishing House.
Marx, K. (1968 [1861]) *Theories of Surplus Value*, Parts I, II and III, Moscow: Progress Press.
Marx, K. (1971 [1879]) 'Marginal notes on A. Wagner,' in D. McLennan (ed.), *Karl Marx: Early Texts*, Oxford: Basil Blackwell.
Marx, K. (1983 [1881]) *Marx's Mathematical Manuscripts*, London, New Park Publications.
Mas-Colell, A. (1977) 'On the equilibrium price set of an exchange economy,' *Journal of Mathematical Economics*, 4: 117–26.
Mas-Colell, A. (1986) 'Notes on price and quantity tatonnement dynamics,' in H. F. Sonnenschein (ed.), *Lecture Notes in Economics and Mathematical Systems*, Berlin: Springer-Verlag.
Mas-Colell, A., M. D. Whinston et al. (1995) *Microeconomic Theory*, New York: Oxford University Press.
May, R. M. and G. F. Oster (1976) 'Bifurcations and dynamic complexity in simple ecological models,' *American Naturalist*, 110(974): 573–99.
McCauley, J. (2004) *Dynamics of Markets: Econophysics and Finance*, Cambridge: Cambridge University Press.
McCauley, J. (2006) 'Response to "Worrying trends in econophysics,"' *Physica A*, 371: 601–9.

McCullough, B. D. and C. G. Renfro (2000) 'Some numerical aspects of nonlinear estimation,' *Journal of Economic and Social Measurement*, 26(1): 63–77.
McDonough, T. and J. Eisenhauer (1995) 'Sir Robert Giffen and the great potato famine: a discussion of the role of a legend in neoclassical economics,' *Journal of Economic Issues*, 29: 747–59.
McFadden, D., A. Mas-Colell and R. R. Mantel (1974) 'A characterisation of community excess demand functions,' *Journal of Economic Theory*, 9: 361–74.
McFadden, J. (2001) *Quantum Evolution: How Physics' Weirdest Theory Explains Life's Biggest Mystery*, New York: W. W. Norton.
McKibbin, W. J. and A. Stoeckel (2009) 'Modelling the global financial crisis,' *Oxford Review of Economic Policy*, 25(4): 581–607.
McKinsey Global Institute (2010) *Debt and Deleveraging: The Global Credit Bubble and Its Economic Consequences*, http://www.mckinsey.com/mgi/publications/books/.
McLellan, D. (1971) *Karl Marx: Early Texts*, Oxford: Blackwell.
McLennan, W. (1996) *Standards for Labour Force Statistics*, Australian Bureau of Statistics, Canberra: Australian Government Publishing Service.
McManus, M., (1963) 'Notes on Jorgenson's model,' *Review of Economic Studies*, 30: 141–7.
Meadows, D. H., J. Randers et al. (1972) *The Limits to Growth*, New York: Signet.
Means, G. C. (1972) 'The administered-price thesis reconsidered,' *American Economic Review*, 62: 292–306.
Meek, R. (1972) *The Economics of Physiocracy*, London: George Allen & Unwin.
Milgate, M. (1987) 'Keynes's General Theory,' in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave: A Dictionary of Economics*, London: Macmillan.
Minsky, H. (1957) 'Monetary systems and accelerator models,' *American Economic Review*, 67: 859–83.
Minsky, H. (1970) 'Financial instability revisited: the economics of disaster,' reprinted in H. Minsky (1982 [1963]), *Can 'It' Happen*

Again?: Essays on instability and finance, Armonk, NY: M. E. Sharpe, pp. 117–61.
Minsky, H. (1971) 'The allocation of social risk: discussion,' *American Economic Review*, 61(2): 389–90.
Minsky, H. (1975) *John Maynard Keynes*, New York: Columbia University Press.
Minsky, H. (1977) 'The financial instability hypothesis: an interpretation of Keynes and an alternative to "standard" theory,' *Nebraska Journal of Economics and Business*, reprinted in H. Minsky (1982 [1963]), *Can 'It' Happen Again?: Essays on instability and finance*, Armonk, NY: M. E. Sharpe, pp. 59–70.
Minsky, H. (1980) 'Capitalist financial processes and the instability of capitalism,' *Journal of Economic Issues*, 14, reprinted in H. Minsky (1982 [1963]), *Can 'It' Happen Again?: Essays on instability and finance*, Armonk, NY: M. E. Sharpe, pp. 71–89.
Minsky, H. (1982) *Inflation, Recession and Economic Policy*, Brighton: Wheatsheaf.
Minsky, H. (1982 [1963]) *Can 'It' Happen Again?: Essays on instability and finance*, Armonk, NY: M. E. Sharpe.
Minsky, H. (1986) *Stabilizing an Unstable Economy*, Twentieth Century Fund Report series, New Haven, CT and London: Yale University Press.
Mirowski, P. (1984) 'Physics and the "marginalist revolution,"' *Cambridge Journal of Economics*, 8: 361–79.
Mirowski, P. (1989) *More Heat than Light: Economics as social physics: Physics as nature's economics*, Cambridge: Cambridge University Press.
Mohun, S. (1994) *Debates in Value Theory*, New York: St Martin's Press.
Moore, B. J. (1979) 'The endogenous money stock,' *Journal of Post Keynesian Economics*, 2(1): 49–70.
Moore, B. J. (1983) 'Unpacking the Post Keynesian black box: bank lending and the money supply,' *Journal of Post Keynesian Economics*, 5(4): 537–56.
Moore, B. J. (1988a) 'The endogenous money supply,' *Journal of Post Keynesian Economics*, 10(3): 372–85.

Moore, B. J. (1988b) *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge: Cambridge University Press.
Moore, B. J. (1997) 'Reconciliation of the supply and demand for endogenous money,' *Journal of Post Keynesian Economics*, 19(3): 423–8.
Moore, B. J. (2001) 'Some reflections on endogenous money,' in L.-P. Rochon and M. Vernengo (eds), *Credit, Interest Rates and the Open Economy: Essays on horizontalism*, Cheltenham: Edward Elgar, pp. 11–30.
Musgrave, A. (1981) '"Unreal assumptions" in economic theory: the untwisted,' *Kyklos*, 34: 377–87, reprinted in B. Caldwell (1984), *Appraisal and Criticism in Economics: A Book of Readings*, London: Allen & Unwin.
Muth, J. F. (1961) 'Rational expectations and the theory of price movements,' *Econometrica*, 29(3): 315–35.
Nightingale, J. (2000) 'Universal Darwinism and social research: the case of economics,' in W. A. Barnett, C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds), *Commerce, Complexity and Evolution*, New York: Cambridge University Press.
Nikaido, H. (1996) *Prices, Cycles and Growth*, Cambridge, MA: MIT Press.
Oakley, A. (1983) *The Making of Marx's Critical Theory*, London: Routledge and Kegan Paul.
Obama, B. (2009) 'Obama's remarks on the economy,' *New York Times*, 14 April.
O'Brien, Y.-Y. J. C. (2007) 'Reserve requirement systems in OECD countries,' SSRN eLibrary.
Oda, S. H., K. Miura, K. Ueda and Y. Baba (2000) 'The application of cellular automata and agent models to network externalities in consumers' theory: a generalization of life game,' in W. A. Barnett, C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds), *Commerce, Complexity and Evolution*, New York: Cambridge University Press.
O'Hara, M. (1995) *Market Microstructure Theory*, Cambridge: Blackwell.
Ormerod, P. (1997) *The Death of Economics*, 2nd edn, New York: John Wiley & Sons.

Ormerod, P. (2001) *Butterfly Economics: A New General Theory of Social and Economic Behavior*, London: Basic Books.
Ormerod, P. (2004) 'Neoclassical economic theory: a special and not a general case,' in E. Fullbrook (ed.), *A Guide to What's Wrong with Economics*, London: Anthem Press, pp. 41–6.
Ormerod, P. and A. Heineike (2009) 'Global recessions as a cascade phenomenon with interacting agents,' *Journal of Economic Interaction and Coordination*, 4(1): 15–26.
Ott, E. (1993) *Chaos in Dynamical Systems*, New York: Cambridge University Press.
Palley, T. I. (1996) *Post Keynesian Economics: Debt, Distribution and the Macro Economy*, London: Macmillan.
Patriarca, M., A. Chakraborti et al. (2004) 'Gibbs versus non-Gibbs distributions in money dynamics,' *Physica A: Statistical Mechanics and Its Applications*, 340(1–3): 334–9.
Paulson, H. M. (2010) *On the Brink: Inside the race to stop the collapse of the global financial system*, New York: Business Plus.
Perino, M. (2010) *The Hellhound of Wall Street: How Ferdinand Pecora's investigation of the Great Crash forever changed American finance*, New York: Penguin Press.
Peters, E. E. (1994) *Fractal Market Analysis*, New York: John Wiley & Sons.
Peters, E. E. (1996) *Chaos and Order in the Capital Markets*, 2nd edn, New York: John Wiley & Sons.
Phillips, A. W. (1950) 'Mechanical models in economic dynamics,' *Economica*, 17(67): 283–305.
Phillips, A. W. (1954) 'Stabilisation policy in a closed economy,' *Economic Journal*, 64(254): 290–323.
Phillips, A. W. (1968) 'Models for the control of economic fluctuations,' in *Scientific Growth Systems, Mathematical Model Building in Economics and Industry*, London, Griffin, pp. 159–65.
Pierce, A. (2008) 'The Queen asks why no one saw the credit crunch coming,' *Daily Telegraph*, London.
Pigou, A. C. (1922) 'Empty economic boxes – a reply,' *Economic Journal*, 36: 459–65.

Pigou, A. C. (1927) 'The law of diminishing and increasing cost,' *Economic Journal*, 41: 188–97.
Pigou, A. C. (1928) 'An analysis of supply,' *Economic Journal*, 42: 238–57.
Poincaré, H. (1956 [1905]) 'Principles of mathematical physics,' *Scientific Monthly*, 82(4): 165–75.
Prescott, E. C. (1999) 'Some observations on the Great Depression,' *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(1): 25–31.
Rassuli, A. and K. M. Rassuli (1988) 'The realism of Post Keynesian economics: a marketing perspective,' *Journal of Post Keynesian Economics*, 10: 455–73.
Renfro, C. G. (2009) *Building and Using a Small Macroeconometric Model: Klein Model I as an Example*, a MODLER Workbook, Philadelphia, PA: MODLER Information Technologies Press.
Ricardo, D. (1817) *On the Principles of Political Economy and Taxation*, Indianapolis, IN: Library of Economics and Liberty, www.econlib.org/library/Ricardo/ricP.html.
Robbins, L. (1928) 'The representative firm,' *Economic Journal*, 42: 387–404.
Robbins, L. (1952 [1932]) *An Essay on the Nature and Significance of Economic Science*, 2nd edn, London: Macmillan.
Robertson, D. H. (1924) 'Those empty boxes,' *Economic Journal*, 34: 16–31.
Robertson, D. H. (1930) 'The trees of the forest,' *Economic Journal*, 44: 80–9.
Robinson, J. (1971a) 'Continuity and the "rate of return,"' *Economic Journal*, 81(321): 120–2.
Robinson, J. (1971b) 'The existence of aggregate production functions: comment,' *Econometrica*, 39(2): 405.
Robinson, J. (1972) 'The second crisis of economic theory,' *American Economic Review*, 62(2): 1–10.
Robinson, J. (1975) 'The unimportance of reswitching,' *Quarterly Journal of Economics*, 89(1): 32–9.
Robinson, J. (1981) *What Are the Questions?: And other essays*, Armonk, NY: M. E. Sharpe.

Roosevelt, F. D. (1933) First Inaugural Address of Franklin D. Roosevelt, Washington, DC.
Rosdolsky, R. (1977) *The Making of Marx's Capital*, London: Pluto Press.
Rosser, J. B. (1999) 'On the complexities of complex economic dynamics,' *Journal of Economic Perspectives*, 13(4): 169–92.
Roth, T. P. (1989) *The Present State of Consumer Theory*, Lanham, MD: University Press of America.
Rothbard, M. (1970) *Power and the Market*, Kansas City: Sheed Andrews and McMeel Inc..
Rothbard, M. (1972 [1963]) *America's Great Depression*, Kansas City: Sheed & Ward.
Rotheim, R. J. (1999) 'Post Keynesian economics and realist philosophy,' *Journal of Post Keynesian Economics*, 22: 71–103.
Salvadori, N. and I. Steedman (1988) 'No reswitching? No switching!' *Cambridge Journal of Economics*, 12: 481–6.
Samuelson, P. A. (1938a) 'A note on the pure theory of consumer's behaviour,' *Economica*, 5(17): 61–71.
Samuelson, P. A. (1938b) 'A note on the pure theory of consumer's behaviour: an addendum,' *Economica*, 5(19): 353–4.
Samuelson, P. A. (1948) *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.
Samuelson, P. A. (1956) 'Social indifference curves,' *Quarterly Journal of Economics*, 70(1): 1–22.
Samuelson, P. A. (1966) 'A summing up,' *Quarterly Journal of Economics*, 80(4): 568–83.
Samuelson, P. A. and W. D. Nordhaus (2010) *Microeconomics*, New York: McGraw-Hill Irwin.
Samuelson, P. A. and W. D. Nordhaus (1991) 'A personal view on crises and economic cycles,' in M. Feldstein (ed.), *The Risk of Economic Crisis. A National Bureau of Economic Research Conference Report*, Chicago, IL, and London: University of Chicago Press, pp. 167–70.
Samuelson, P. A. and W. D. Nordhaus (1998) 'Report card on Sraffa at 100,' *European Journal of the History of Economic Thought*, 5: 458–67.
Sargan, J. D. (1958) 'The instability of the Leontief dynamic model,' *Econometrica*, 26: 381–92.

Sargent, T. J. and N. Wallace (1976) 'Rational expectations and the theory of economic policy,' *Journal of Monetary Economics*, 2(2): 169–83.
Sato, K. (1979) 'A note on capital and output aggregation in a general equilibrium model of production,' *Econometrica*, 47: 1559–68.
Sawyer, M. C. (ed.) (1988) *Post-Keynesian Economics*, Aldershot: Edward Elgar.
Say, J. B. (1967 [1821]) *Letters to Mr Malthus on several subjects of political economy and on the cause of the stagnation of commerce to which is added a Catechism of Political Economy or familiar conversations on the manner in which wealth is produced, distributed and consumed in society*, New York: Augustus M. Kelly.
Schumpeter, J. A. (1934) *The Theory of Economic Development: An inquiry into profits, capital, credit, interest and the business cycle*, Cambridge, MA: Harvard University Press.
Schwartz, J. H. (2000) *Sudden Origins: Fossils, Genes, and the Emergence of Species*, New York: John Wiley & Sons.
Sciabarra, C. (1995) *Marx, Hayek and Utopia*, New York: State University of New York Press.
Sent, E. M. (1997) 'Sargent versus Simon: bounded rationality unbound,' *Cambridge Journal of Economics*, 21: 323–8.
Seppecher, P. (2010) 'Dysfonctionnement bancaire, bulle du crédit et instabilité macroéconomique dans une économie monétaire dynamique et complexe' [Dysfunctional banking system, credit bubble and macroeconomic instability in a complex, dynamic, monetary economy – with English summary], *Revue Economique*, 61(3): 441–9.
Shafer, W. and H. Sonnenschein (1982) 'Market demand and excess demand functions,' in K. J. Arrow and M. D. Intriligator (eds), *Handbook of Mathematical Economics*, vol. II, Amsterdam: Elsevier.
Shafer, W. J. (1977) 'Revealed preference and aggregation,' *Econometrica*, 45: 1173–82.
Shaikh, A. (1982) 'Neo-Ricardian economics: a wealth of algebra, a poverty of theory,' *Review of Radical Political Economics*, 14: 67–83.
Sharpe, W. F. (1964) 'Capital asset prices: a theory of market equilibrium under conditions of risk,' *Journal of Finance*, 19(3): 425–42.
Sharpe, W. F. (1970) *Portfolio Theory and Capital Markets*, New York: McGraw-Hill.

Shepherd, W. G. (1984) '"Contestability" vs. competition,' *American Economic Review*, 74: 572–87.
Shove, G. F. (1930) 'The representative firm and increasing returns,' *Economic Journal*, 44: 93–116.
Silvestre, J. (1993) 'The market-power foundations of macroeconomic policy,' *Journal of Economic Literature*, 31(1): 105–41.
Simon, H. A. (1996) *The Sciences of the Artificial*, Cambridge, MA: MIT Press.
Sippel, R. (1997) 'An experiment on the pure theory of consumer's behaviour,' *Economic Journal*, 107(444): 1431–44.
Smith, A. (1838 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Edinburgh: Adam and Charles Black.
Solow, R. M. (1956) 'A contribution to the theory of economic growth,' *Quarterly Journal of Economics*, 70(1): 65–94.
Solow, R. M. (2001) 'From neoclassical growth theory to new classical macroeconomics,' in J. H. Drèze (ed.), *Advances in Macroeconomic Theory*, New York: Palgrave.
Solow, R. M. (2003) 'Dumb and dumber in macroeconomics,' *Festschrift for Joe Stiglitz*, New York: Columbia University.
Solow, R. M. (2007) 'The last 50 years in growth theory and the next 10,' *Oxford Review of Economic Policy*, 23(1): 3–14.
Sonnenschein, H. (1972) 'Market excess demand functions,' *Econometrica*, 40(3): 549–63.
Sonnenschein, H. (1973a) 'Do Walras' identity and continuity characterize the class of community excess demand functions,' *Journal of Economic Theory*, 6: 345–54.
Sonnenschein, H. (1973b) 'The utility hypothesis and demand theory,' *Western Economic Journal*, 11: 404–10.
Sornette, D. (2003) *Why Stock Markets Crash: Critical events in complex financial systems*, Princeton, NJ: Princeton University Press.
Sornette, D. (2011) 'Financial Crisis Observatory,' www.er.ethz.ch/fco.
Sraffa, P. (1926) 'The laws of returns under competitive conditions,' *Economic Journal*, 36(144): 535–50.
Sraffa, P. (1930) 'The trees of the forest – a criticism,' *Economic Journal*, 44: 89–92.

Sraffa, P. (1960) *The Production of Commodities by Means of Commodities: Prelude to a critique of economic theory*, Cambridge: Cambridge University Press.
Steedman, I. (1977) *Marx after Sraffa*, London: NLB.
Steedman, I. (1992) 'Questions for Kaleckians,' *Review of Political Economy*, 4: 125–51.
Stevens, G. (2008) 'Interesting times,' *Reserve Bank of Australia Bulletin*, December, pp. 7–12.
Stigler, G. J. (1957) 'Perfect competition, historically contemplated,' *Journal of Political Economy*, 65(1): 1–17.
Stiglitz, J. (1993) 'Post Walrasian and post Marxian economics,' *Journal of Economic Perspectives*, 7 (1): 109–14.
Stiglitz, J. (1998) 'The confidence game: how Washington worsened Asia's crash,' *New Republic Online*, 9 September.
Stiglitz, J. (2000) 'What I learned at the world economic crisis,' *New Republic*, 17–24 April, pp. 56–60.
Strange, S. (1997) *Casino Capitalism*, Manchester: Manchester University Press.
Swan, T. W. (2002) 'Economic growth,' *Economic Record*, 78(243): 375–80.
Sweezy, P. M. (1942) *The Theory of Capitalist Development*, New York: Oxford University Press.
Taleb, N. (2007) *The Black Swan: The Impact of the Highly Improbable*, New York: Random House.
Taslim, F. and A. Chowdhury (1995) *Macroeconomic Analysis for Australian Students*, Sydney: Edward Elgar.
Taylor, J. B. (1993) 'Discretion versus policy rules in practice,' *Carnegie-Rochester Conference Series on Public Policy*, 39: 195–214.
Taylor, J. B. (2007) 'The explanatory power of monetary policy rules,' *Business Economics*, 42(4): 8–15.
Tisdell, C. (1995) 'Evolutionary economics and research and development,' in S. Dowrick (ed.), *Economic Approaches to Innovation*, Aldershot: Edward Elgar.
Valentine, T., G. Ford et al. (2011) *Fundamentals of Financial Markets and Institutions in Australia*, Sydney: Pearson Australia.
Varian, H. R. (1984) *Microeconomic Analysis*, New York: W. W. Norton.

Varian, H. R. (1992) *Microeconomic Analysis*, new edn, New York: W. W. Norton.
Veblen, T. (1898) 'Why is economics not an evolutionary science?' *Quarterly Journal of Economics*, 12(4): 373–97.
Veblen, T. (1899) 'The preconceptions of economic science,' *Quarterly Journal of Economics*, 13: 121–50.
Veblen, T. (1909) 'The limitations of marginal utility,' *Journal of Political Economy*, 17: 620–36.
Veblen, T. (1990 [1919]) 'The place of science in modern civilization,' in W. J. Samuels (ed.), *The Place of Science in Modern Civilization*, New Brunswick, NJ: Transaction Publishers.
Von Kurt, F. (1945) 'The discovery of incommensurability by Hippasus of Metapontum,' *Annals of Mathematics*, 46(2): 242–64.
Von Neumann, J. and O. Morgenstern (1953) *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
Walras, L. (1954 [1874]) *Elements of Pure Economics*, London: George Allen & Unwin.
Walters, B. and D. Young (1997) 'On the coherence of Post Keynesian economics,' *Scottish Journal of Political Economy*, 44: 329–49.
Whaples, R. (1995) 'Changes in attitudes among college economics students about the fairness of the market,' *Journal of Economic Education*, 26: 308–13.
Wilber, C. K. (1996) 'Ethics and economics,' in C. J. Whalen (ed.), *Political Economy for the 21st century*, Armonk, NY: M. E. Sharpe.
Wilde, L. (1989) *Marx and Contradiction*, Aldershot: Avebury.
Witt, U. (ed.) (1992) *Explaining Process and Change: Approaches to Evolutionary Economics*, Ann Arbor: University of Michigan Press.
Woodford, M. (2010) 'Simple analytics of the Government Expenditure Multiplier,' NBER Working Paper Series, Cambridge, MA: National Bureau of Economic Research.
Yao, X. and P. Darwen (2000) 'Genetic algorithms and evolutionary games,' in W. A. Barnett, C. Chiarella, S. Keen, R. Marks and H. Schnabl (eds), *Commerce, Complexity and Evolution*, New York: Cambridge University Press.
Yezer, A. M., R. S. Goldfarb and P. J. Poppen (1996) 'Does studying economics discourage cooperation? Watch what we do, not what we say

or how we play,' *Journal of Economic Perspectives*, 10(1): 177–86.

INDEX

academia, purge in
adaptation
Adaptive Expectations
agents; as basis of analysis
aggregate demand; composition of; debt-financed
aggregate supply
Aggregate Supply–Aggregate Demand model (AS-AD)
aggregation, Sraffa's views of
agriculture
Albert, David
algebraic equations
Allais Paradox
alternative economics
American Economic Review (AER)
Anderson, Philip
Arrow, K. J.
assumptions; ancillary; and logic; domain assumptions; heuristic assumptions; importance of, to economists; negligibility assumptions; taxonomy of *see also ceteris paribus* assumption
auctioneer, Walras model of
Australia, cash handout in
Austrian school of economics; weaknesses of
Automatic Earth, The
Axioms of Revealed Preference

Bagehot, Walter
bailout: of banks; of firms; of individuals
bankers, as separate class
bankruptcy

banks; bailout of; controlling lending of; creation of new money; distinct from firms; electronic transfer banking; profit from debt; reserves of; role of
barter economy
Basel Rules
Behavioral Finance
Bentham, Jeremy; 'In defence of usury'
Bernanke, Ben; as chairman of Federal Reserve; *Essays on the Great Depression*
Bertrand model
Bezemer, Dirk; 'No one saw this coming'
Bhaduri, A.
bicycle riding, learning of
Big Bang theory
Big Government
Biggs, M.
Black Swan events
Blanchard, Olivier
Blatt, John
Blaug, Mark
Bliley, Thomas J.
Blinder, Alan
Blodget, Henry
Bohm, David
Bose, Arun
Brahe, Tycho de
budget constraints
budget deficits
Buffet, Warren
business cycle; in USA; real
Butterfly Effect

calculus; use of
Cambridge controversies
capacity; capacity utilization

capital: as amorphous mass; circulating; definition of; different meanings of; fixed; measurement of; theory of
capital assets pricing model (CAPM); failure of
capitalism: as joint enterprise; demise of, imagined; monetary model of; response to disequilibrium; value as social system
Carpenter, Seth, 'Money, reserves, and the transmission of monetary policy'
ceteris paribus assumption
chain reaction of depression
chain rule
chaos; concept of; exponential sensitivity of situations
chaos theory
Chiarella, Carl
circuit of capital
Clark, J. B.
classes: in capitalism; power of
classical economics
closed loop models
Clower, Robert
Coase, Ronald
cobweb model
Coddington, A.
commodity and commodity-power
commodity production: mystery of; Marx's analysis of; theories of
comparative statics method
competition *see also* perfect competition
completeness
complexity theory
Computational Theory, laws of
computers, training in
conjectural variation
conservation laws
constant technology, assumption of
constructionism
consumer, individual, in economic theory
consumer demand; theory of

consumer sector, and debt
consumption
continuity
conventions, formation of
convexity
Copenhagen school of quantum theory
costs: fixed; variable
Cotis, Jean-Philippe
Cournot and Bertrand competition
Cournot-Nash model
creative accounting
credit; analysis of; creation of; nature of; pure credit economy; role of
credit crunch
Credit Impulse
curves *see* demand curves, Engel curves, indifference curves *and* supply curves
cycles; cyclical nature of capitalism *see also* business cycle

Davidson, Paul
Debreu, Gerard; *Theory of Value*
debt; absence of, from macroeconomic models; acceleration of; analysis of; and aggregate demand; banks profit from; correlated with unemployment; creation of; empirical dynamics of; further endebtment of debtors; growth of; limitation of; private (dominant role of); rate of change of; reduction of; relation to national income (GDP) *see* debt to GDP ratio; repayment of (non-repayment of); role of; zero target *see also* Jubilee policy
debt bubble
debt deflation; hypothesis; in USA; modeling of
debt-financed speculation
debt to GDP ratio
Debtwatch
deflation
deleveraging
demand, affected by rising income *see also* Law of Demand
demand constraints

demand curves; downward-sloping; for labor; Hicksian compensated; individual, determining of
demand management
Demiralp, S.
Denning, Dan
Dent, Harry
depreciation
depression; alleged impossibility of
deregulation; and crisis
diagrams, use of
dialectical philosophy, of Marx
dictator, requirement for
differential equations
dimensionality, curse of
disequilibrium; capital's response to
distribution of income; neoclassical theory of
diversity
divine right of kings
division of labor
dollar, value of
Dorgan, Byron L.
DotCom bubble
double-entry bookkeeping
Dow, Sheila
Dow Jones Industrial Average, movements of
dynamic analysis
Dynamic Stochastic General Equilibrium model (DGSE); uselessness of

earthquakes, theory of
Econlit online service
econometric models, large-scale
economics: as pre-science; as science; different from true sciences; dumbing-down of textbooks; future of; loss of interest in; methodology of; need for revolution in; neoclassical *see* neoclassical economics; paradigm shifts in; teaching of
economies of scale

econophysics
efficiency of markets
efficient markets hypothesis; contradicted; failure of
Ehrenberg, Andrew
Einstein, Albert; Theory of Relativity
emergent behavior
emergent properties, concept of
employment, theory of
Engel curves
entrepreneur, role of; as key actor
entropy
equilibrium; conflated with economic utopia; dismissed by Keynes; general (unreal assumptions of); in exchange rate of commodities; in Marx; in neoclassical theory; instability of; irrelevance of; partial equilibrium method; self-equilibrating economy; value of shares; Walrasian model of
Escher drawings
euphoric economy
European Union (EU), and economic growth
evolutionary school of economics
exchange, in primitive societies
exchange-value
expectation; formation of; rational
expected value, concept of
experiments, importance of
externalities

‘F-twist’
factories, built with excess capacity
factors of production; fixed; flows of; variable
failure to see crisis coming
false equalities
Fama, Eugene
Federal Reserve Board; mismanagement of money supply
feudalism
finance sector; reform of

financial assets: analysis of; pricing of
Financial Crisis Inquiry Commission
Financial Crisis Observatory
Financial Instability Hypothesis *see* Minsky's Financial Instability Hypothesis
firm, theory of (neoclassical)
Fisher, Irving; bankruptcy of; debt deflation paper; *The Theory of Interest*
flat world analogy
Fokker-Planck model
fractal geometry
Fractal Markets Hypothesis
Fractional Reserve Banking
Freakonomics
freedom, and labor
French, Kenneth
freshwater *v*. saltwater macroeconomics
Friedman, Milton; 'assumptions don't matter'; helicopter image; paper on methodology
full employment
Fullbrook, Edward
Fullwiler, Scott

Galbraith, J. K.
Galileo Galilei
Game Theory
Garicano, Luis
Giffen goods
Glass-Steagall Act, repeal of
glut; alleged impossibility of
Gödel's proof
Godley, Wynne
gold standard; theory of
Goodwin, Richard; growth cycle model
Gordon, R. J.
Gorman, William
government spending, as stabilizer of economy

Gramm, Phil
Graziani, Augusto
Great Depression; misunderstanding of
Great Moderation
Great Recession; combating of; misunderstanding of; neoclassical response to
Greater Fool principle
Greenspan, Alan
growth theory; neoclassical

Handbook of Mathematical Economics
Harrison, Edward
Harrod, R.
Haugen, Robert
Hayek, Friedrich von
Hegel, G. W. F., Marx's critique of
Henon Map
heterogeneity
Hicks, John; 'Mr Keynes and the Classics' *see also* IS-LM model
Hicksian compensated demand curve
Hippases of Metapontum
Hirst, David
Holmes, Alan
homothetic conditions
Horizontalism
households: convert goods into labor/capital; debt of
housing bubble
Hudson, Michael
Humpty Dumpty
hypotheses, assumptions of

imperfections in model
income effects
income of capital
independence
indifference curves; elimination of

individual behavior, theory of
individual rational investor
Inefficient Markets Hypothesis
inertness of economics
infinitesimal amounts, not zeros
inflation; causes of; relation to unemployment
inflationary universe, theory of
innovation
instability; of capitalism; of systems *see also* equilibrium, instability of *and* Minsky's Financial Instability Hypothesis
Institute for New Economic Thinking (INET)
instrumentalism
interest; medieval prohibition of; rates of (ceiling for; legal; setting of)
interests, sum of
international trade, theory of
Internet; use of
Internet Bubble
investment; badly apportioned; cloud of possible investments; impact of expectations on
investment counselling
Investment Opportunity Curve (IOC)
invisible hand
Ireland, P. N.
irrational behavior
irrational numbers
irrationality; introduced into models; of managers; of markets
IS-LM model; Lucas critique of; demolition of; origins of

Japan
Jevons, W. S.
Journal of Post Keynesian Economics (JPKE)
Jubilee policy
Jubilee shares

Kaletsky, Anatole
Kates, Steve

Keiser, Max
Keynes, J. M.; and uncertainty; critique of marginal efficiency of capital; critique of Say's Law; *General Theory of Employment, Interest and Money*; models of convention formation; on expectations; view of stock market
Keynesian economics; American; new; overthrow of *see also* post-Keynesianism
Keynesianism
Kindleberger, Charles
Kirman, Alan
Koo, Richard, *The Holy Grail of Macroeconomics*
Kornai, Janos
Krugman, Paul; 'Debt, deleveraging, and the liquidity trap'
Kumar, Dharma
Kydland, Finn

labor: as only source of profit; surplus labor; value of
labor embodied/labor commanded
labor market; theory of
labor theory of value
Lakatos, I.
land: as factor of production; as source of value
Law of Demand; applicability of; non-applicability of
Leeson, Robert
Leijonhufvud, Axel
leisure time, and income
lemmings, fluctuation of populations of
Leontief input–output matrix
leverage, responsible levels of
Levy Institute
liability structures of financial institutions
liquidity; of firms
liquidity preference
liquidity trap
Ljungqvist, L.
loan-to-value ratio (LVR)

logical contradiction
long run average cost curve
Long Term Capital Management Crisis
Lorenz, E. N.
Lorenz, Hans-Walter
Lucas, Robert; critique of IS-LM model

Maastricht Treaty, impact on Europe
machinery; depreciation of; role of; value of
macroeconomics; and anticipation; and reductionist fallacy; as applied microeconomics; as Keynesian invention; methodology of; modeling of
Malthus, T.
Mandelbrot, Benoit
Mankiw, Gregory
Mantel, R. R.
marginal cost; curve of; diminishing; equated with marginal revenue
marginal efficiency of capital; Keynes's critique of
marginal productivity; of capital; of labor; theory of
marginal returns, diminishing, law of
marginal revenue
marginal workers
market: artificial, creation of; competitive; free; imperfection of
market demand curve; downward-sloping; not downward-sloping
market economy; complexity of
marketing, as cost of distribution
Markowitz model
Marshall, Alfred; theory of the firm
Martenson, Chris
Marx, Karl; analysis of commodity; as 'minor post-Ricardian'; *Capital*; capital axioms of; circuits analysis; critique of Ricardo; critique of Say's Law; dialectical philosophy of; labor theory of value *see also* labor theory of value; misinterpretation of; model of growth; on origins of surplus value; relevance of; tables provided by; theory of value; treatment of equilibrium; without labor theory of value
Marxist economics; after Sraffa; future of
Mas-Colell, Andreu; *Microeconomic Theory*

Massachusetts Institute of Technology (MIT)
mathematics; fallacious propositions; humility in; in economics (failure of; limits to); of fractal geometry; Pythagorean; use of formulae
mathematization of physics
Mayer, T.
McCauley, Joe
McKibbin, W. J.
means of production
Meek, Ronald, *Studies in the Labor Theory of Value*
Mencken, Henry
Menger, Carl
meteorology, modelling of
methodology of macroeconomics
microeconomics
Minsky, Hyman; *Can 'It' Happen Again?*; *John Maynard Keynes*; modeling of; *Stabilizing an Unstable Economy*
Minsky's Financial Instability Hypothesis
Mirowski, P.
Mises, Ludwig von
Mississippi Land Scheme
Mitchell, Bill
monetarism
Monetary Circuit
monetary model of capitalism
monetary policy
money; as non-commodity; base money; creation of, model of; endogenous; specificities of; theory of *see also* printing of money
money illusion
Money Multiplier; fallacy of
money neutrality
money supply: determination of; increase of; management of
MONIAC simulation model of economy
monopoly; economists' hostility to; laws against
monopsony
Moore, Basil
more is different

multiverses
Musgrave, Alan
Muth, J. F.

Nasdaq; crash of
Natural Rate Hypothesis
neoclassical economics; destabilizing effect of; as progressive school; backlash against; barter paradigm; failure of; flaws in; intellectual weakness of; rejection of; response to Sraffa; theory of production; unscientific discipline; view of trade unions
Neoclassical Counter-Revolution
NetLogo program
Neumann, John von, with Oskar Morgenstern, *The Theory of Games and Economic Behavior*
neutral goods
New Economic Perspectives
Newton, Isaac, assumptions of
Noland, Doug
non-satiation
nonlinear differential equations
nonlinear relationships between variables

Obama, Barack; bank bailout decision
omitted variables
options pricing models
order, spontaneous/designed
Organization for Economic Cooperation and Development (OECD), 'small global forecasting' model
Ormerod, Paul

PAECON movement
pain, avoidance of
Patinkin, Don
Pecora, Ferdinand
perfect competition; durability of; equals monopoly; war over
personal satisfaction
Phillips, A. W.

Phillips curve; short-run
physiocrats
piano playing
Pick, A.
Pigou, A. C.
planetary behavior, theory of
pleasure, pursuit of
Poincaré, Henri
policy ineffectiveness proposition
politicians, influenced by neoclassical economics
polynomialism
Ponzi finance
Popper, Karl
positive economics
post-Keynesianism
potato famine in Ireland
Power Law model
Prescott, Edward
price; changing, impact on consumer demand; controls imposed on; determination of; positive; theory of ('additive')
pricing, of financial assets
printing of money
probability
product rule
production; Austrian theory of; economists' assumptions about; neoclassical theory of; reswitching of; with a surplus; with explicit labor; with no surplus
productivity: diminishing, causes rising price; does not determine wages; falls as output rises
profit; definition of; falling tendency of; Marxian calculation of; maximization of (short-run); normal; rate of (determination of; falling); source of; super-normal
proof by contradiction
Property Income Limited Leverage (PILL)
proto-energetics
psychic income

Pythagorean mathematics

QE1
qualitative easing
quantum uncertainty
Quesnay, François
Quesnay Economic Dynamics (QED)

random numbers
Rapping, Leonard
rational behavior
rational expectation *see* expectation, rational
rational numbers
rational private behavior
rationality; definition of; in economics
realism
recession
reductionism; fallacy of; reconstituted; strong
Regulation Q
regulatory capture
rent; theory of
Repast program
representative agent *passim*
Reserve Requirement
reswitching
returns to scale
Ricardo, David; Marx's critique of; *Principles of Political Economy and Taxation*; theory of rent
risk; and return; measurement of; risk-averse behavior
Robbins, L.
Robinson, Joan
Rodbertus, J. K.
Roosevelt, Franklin D.
Rothbard, Murray
Roubini, Nouriel
roundaboutness

Russia

Samuelson, Paul; ‘social indifference curves’; *Foundation of Economic Analysis*
Sargent, Thomas
Savings and Loans crisis
Say, Jean Baptiste
Say’s Law; Marx’s critique of
scarcity
Schiff, Peter
Schrödinger’s cat
Schumpeter, J. A.
Sciabarra, Chris
scientific realism
second-order linear differential equations
self-employment
self-interest
self-organization of systems
Shadowstats.com
Shaikh, Anwar, ‘Neo-Ricardian economics’
Sharpe, W. F.
Shedlock, Mish
Shiller, Robert
shocks, exogenous
short run
short-term funding, dependence on
Simulink system dynamics program
Sippel, Reinhard
size, importance of
Smith, Adam; *The Wealth of Nations*
Smith, Edgar, *Common Stocks as Long Term Investment*
Smith, Yves
social conflict, not taken into account
social security
social welfare; maximization of
social well-being

socialism; perceived inevitability of; scientific
socialist economies, shortages in
society, as sum of individuals
soft budget constraint
Solow, Robert
Sonnenschein, H. F.
Sonnenschein-Mantel-Debreu (SMD) conditions
Sornette, Didier
Soros, George
South Sea Bubble
speculation; debt-financed; in housing; on stock market
Sraffa, Piero; alternative outlook of; and law of diminishing marginal returns; 'The law of returns under competitive conditions'; neoclassical response to; on aggregation; *The Production of Commodities by Means of Commodities*; summing up of
Sraffian economics
stability: negative; of markets
stagflation
standard deviation measurement
state, role in creating money
static analysis
Steedman, Ian
Stevens, Glenn
Stigler, George
Stiglitz, Joseph
stock market: behavior of; Keynes's view of; volatility of
Stoekel, A.
straight lines, in economic analysis
subprime bubble
substitution effects
Summers, Larry
supply and demand; analysis of; backward-bending, in labor market; independence of; laws of; of labor; setting of price
supply curves; non-existence of; of labor; theory of
surplus, production of
surplus value, origins of

Swan, Trevor
Sweezy, Paul, *The Theory of Capitalist Development*
Systems Engineering Approach

Taleb, Nassim
tatonnement (groping)
Taylor Rule
Taylor, John
Thatcher, Margaret
thermodynamics
third agent
Thornton, Henry
time: concept of; economists' treatment of; factor omitted; importance of
time constants
time discount; rate of
time value: of goods; of money
total differential of profit
trade unions; economists' view of; laws against; non-unionization of economists
transformation problem
transitivity
Tsallis-statistics
Tulip Mania
Turing, Alan
turnover: of money; periods of

uncertainty; Keynes and
Undercover Economist, *The*
unemployment; definitions of; relation to wage increases *see also* inflation, relation to unemployment
unexplored conditions
unions *see* trade unions
United States of America (USA); debt deflation in; economic boom in; economic future of; economy of (insolvency of); idle industrial capacity in; industrialization in; leveraging in; money supply in; power of finance sector in; shadow banking in; unemployment definition in

use-value; in Marx
usury
util
utilitarianism
utility; cardinal; concept of; expected; immeasurable; marginal (diminishing); maximization of; ordinal; perceived; ratio of; social; subjective

value; relation to price; source of; theory of; transition to price *see also* labor theory of value
value at risk formulas
Varian, Hal; *Microeconomic Analysis*
Veblen, Thorstein
velocity of money
Vensim program
Vissim program
Volcker, Paul

wages; determination of (theory of); minimum wages; relation to unemployment
Wagner, Adolph
Wallace, Neil
Walras, L.; *Elements of Pure Economics*; view of equilibrium
Walras's Law; fallacy of
weather, modeling of
Weintraub, Sidney
welfare state
well-being, social
Wellstone, Paul
Wheelwright, Ted
Williams, John
Woodford, M.
work, as 'pain' *see also* labor
working hours of labor

Zero Hedge
Zhang, Cheng