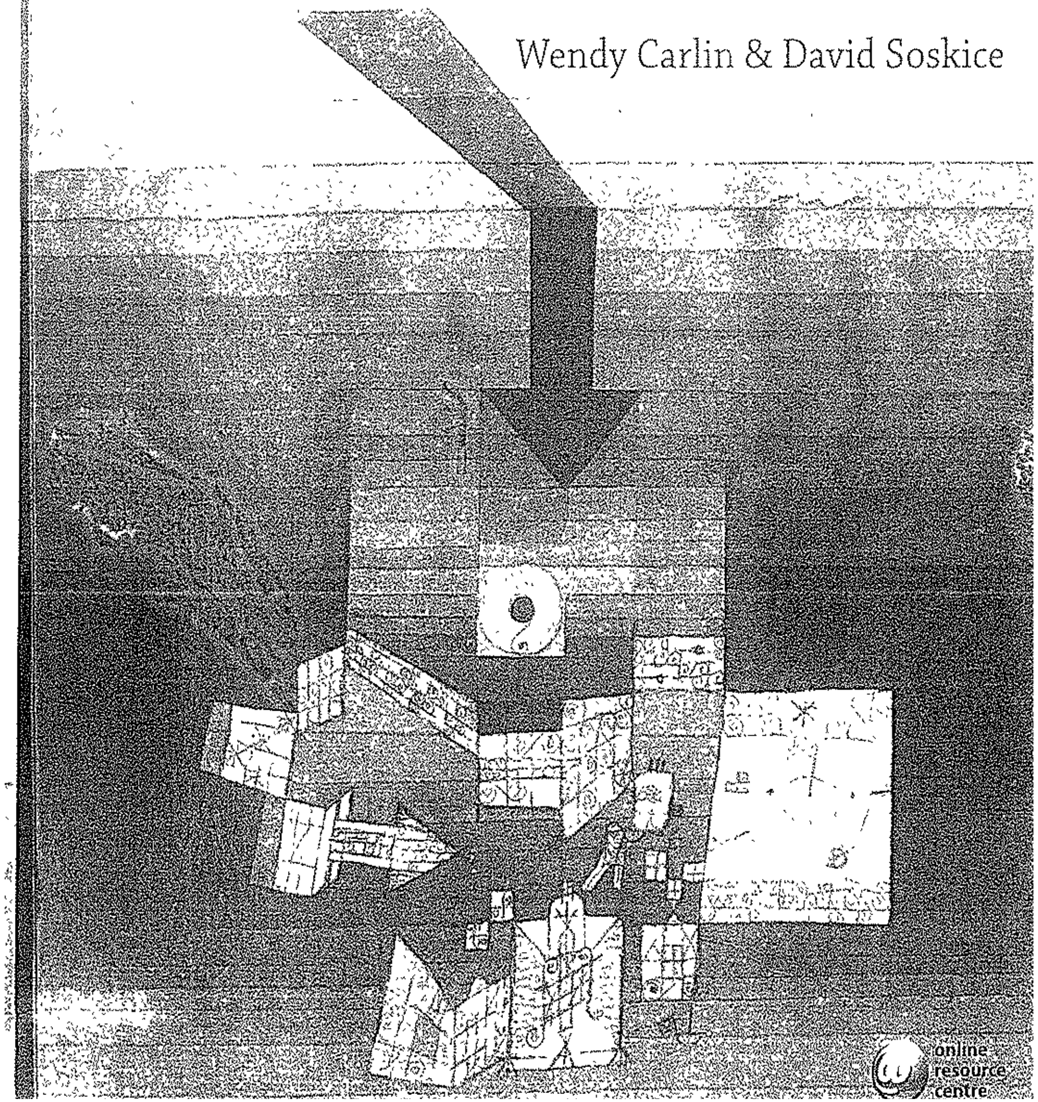
ƷRO. SOSKICE, David y CARLIN, Wendy - Macroeconomics

MACROECONOMICS

Institutions, Instability, and the Financial System

Wendy Carlin & David Soskice



Reviews

'Carlin and Soskice have produced a gem of a book The teaching of macroeconomics after the crisis has changed surprisingly little, limiting itself to incorporating 'frictions' into otherwise standard models that failed during the crisis. The authors embark on a much more ambitious venture. They show how the financial cycle and macroeconomics are inextricably linked, with the risk-taking channel as the linchpin. Their exposition is refreshingly original and yet lucid and accessible. This book will appeal to serious students of economics and to all inquiring minds who have wondered about the role of the financial cycle in macroeconomics.'

- Hyun Song Shin, Economic Adviser and Head of Research, Bank for International Settlements and Hughes-Rogers Professor of Economics, Princeton University

'This is, I believe, the first macroeconomic textbook effectively to incorporate the lessons of the Great Financial Crisis and to describe how financial frictions can impact the macro-economy. The authors weave together the old mainstream, three-equation, model with the newer account of potential financial disturbances in a lucid and efficient manner. As such, it has a major advantage over almost *all other extant textbooks, and will be a boon not only for* undergraduates, but also for graduates and those wishing to understand the current working of our macroeconomic system, beset as it has been with financial strains.'

- Professor Charles Goodhart, Director of the Financial Regulation Research Programme, The London School of Economics and Political Science

'This illuminating book introduces the reader to macroeconomics in a revolutionary fashion. Namely, by means of very elegant and accessible models based on sound micro foundations and developed against a narrative of the performance and policy regimes of the advanced economies over the post-war period Unlike most other macro textbooks, this book builds on the most recent research and debates to teach macroeconomics the way it should now be taught: by emphasizing the interplay between macro and finance; by linking growth to innovation, market structure and firm dynamics; and more generally by taking institutions seriously into account when looking at growth, business cycles, and unemployment and the interplay between them. This book is an absolute must-read for students and policy makers, even those with little initial background, who need to be fully acquainted with modern macroeconomics.'

- Philippe Aghion, Robert C. Waggoner Professor of Economics, Harvard

'This is an exciting new textbook. It offers a clear and cogent framework for understanding not only the traditional macroeconomic issues of business cycles, inflation and growth, but also the financial crisis and ensuing Great Recession that have recently shaken the world economy. The paradigm it offers is highly accessible to undergraduates. Yet at the same time it is consistent with what goes on at the frontiers of the field. Overall, the book confirms my belief that macroeconomics is alive and well!'

- Mark Gertler, Henry and Lucy Moses Professor of Economics, New York University

'To be relevant, economics need to help society understand those phenomena which do it greatest harm—unemployment, inflation and deflation, financial instability, fiscal and banking crisis. Pre-crisis, mainstream economic models failed that societal test and therefore failed society. Wendy Carlin and David Soskice's important new book is the first step towards redemption, providing students and scholars with a rigorous but accessible framework for understanding what troubles society most.'

- Andrew G Haldane, Chief Economist, Bank of England

'The Carlin and Soskice book does a wonderful job of covering the economics behind macroeconomics and the financial system, alongside presenting the latest research on this and the drivers of the great recession It also has an impressive array of data and examples woven in with theory explained in a beautifully intuitive way For any student interested in a refreshingly modern take on the financial crisis and the economics that underlie this, this book is invaluable.'

- Nicholas Bloom, Professor of Economics, Stanford University

'One of the first macro textbooks to integrate the lessons of the crisis. An elegant bridge between introductory undergraduate and graduate macro texts.'

- Olivier Blanchard, Chief Economist, IMF, and Professor of Economics, MIT

'In the light of the events of the past decade, it is important that a new macroeconomics text attempts to satisfy the demands of those learning and using macroeconomics to be able to access relatively simple models which reflect the ways in which the financial sector interacts with the real economy This is by no means an easy task. The new Carlin and Soskice book represents a significant step forward in this regard. Consequently undergraduates, post-graduates and their teachers should be grateful that they can now access teaching materials which have something useful to say about the *financial crisis*.'

- Professor Stephen Nickell, CBE, FBA. Honorary Fellow of Nuffield College, Oxford

Macroeconomics

Institutions, Instability, and the Financial System

Wendy Carlin
David Soskice

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University press is a department of the University of Oxford
It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

Impression 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number 2014953103

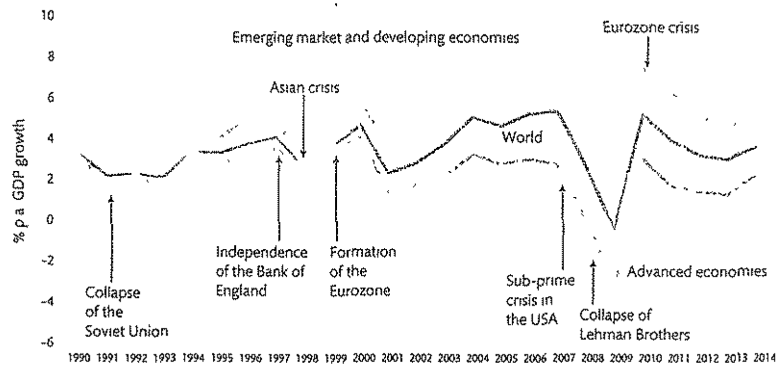
ISBN 978-0-19-965579-3

Printed in Great Britain by Ashford
Colour Press Ltd, Gosport, Hampshire

For Niki, and in memory of Andrew

The chart shows the dramatic events of the past decade and a half in the global economy. Following the Asian crisis in 1998, the emerging market and developing economies grew strongly and became large enough to drive global growth in the 2000s. The financial crisis in the advanced economies plunged the world economy into recession The Eurozone crisis has depressed global growth from 2010

Growth and crisis in the global economy



Source. International Monetary Fund, World Economic Outlook Database, April 2014
Series Gross domestic product, constant prices, percent per annum growth rate

Cover image Paul Klee, *betroffener Ort* (translated as *Affected Place*), 1922

Preface

This is two books in one. It is first a textbook and second the result of our on-going research into building small scale but realistic and comprehensible models of the macro-economy. On both counts it is a response to the challenges facing students, teachers and many professional economists since the calm of the Great Moderation (the period of low inflation and unemployment from the early 1990s to the start of the crisis in 2007) was rocked by the credit crunch of 2007 and the full-blown financial crisis of 2008–9. Further challenges have followed: the emergence of sovereign debt problems in the Eurozone, the persistence of global imbalances in the aftermath of the crisis and longer term problems regarding sustainable growth in emerging and advanced economies.

Stephen Nickell wrote about our previous book published in 2006 that it 'is based on the mainstream monetary macro model which is now widely used by both academics and policy makers In a straightforward manner, it shows how this model can be used to address an enormous variety of practical questions without heavy use of mathematical technique This is modern macroeconomics for undergraduates, post-graduates and business economists alike'. Today there is less of a consensus around a satisfactory 'mainstream model'. During the Great Moderation period of relatively stable growth and low inflation, shocks never pushed economies too far from an equilibrium defined as a zero output gap with inflation at the central bank's target. As a result, what later became serious problems did not arise in the Great Moderation for instance, sharp rises in the share of debt-constrained households in the economy, 'abnormal' cuts in lending by commercial banks, the freezing up of the inter-bank market or the inability of central banks to set their desired interest rates. As long as the financial system provided continuity of core banking services, there had been an unspoken agreement that there was no need for it to be part of the macro model, and it was not

Our goal here is to integrate the financial system with the macroeconomic model. In doing this we take account of the gaps in the mainstream model exposed by the financial crisis and the Eurozone crisis We hope to supply readers with a simple but realistic model which they can use both to analyse the state of the economy and to think systematically about responding to its problems. We integrate the modelling with the presentation of descriptive data, evidence about the empirical testing of the important relationships among variables, and institutional arrangements.

What this book offers to undergraduates studying economics

It is likely you were attracted to study economics because you wanted to know more about why the world was plunged into a financial crisis and how the policies adopted by central banks and finance ministries in response might work. We aim to equip you with a framework for understanding macroeconomics and the financial system that you will feel comfortable using to explain to your friends what is going on in the global economy.

Just as when studying labour economics, you expect to be able to participate in debates about the minimum wage and the effects of migration, and in industrial economics, about the effects of competition policy and of incentives to boost R&D, so too in macroeconomics you should be able to discuss important real world problems. Some of the questions for which we will provide the tools to address are the following: the role of the global banking system in the upswing to the financial crisis, how a period of economic stability and confidence in macroeconomic management could sow the seeds of a crisis of a scale not seen since the Great Depression; the merits and failings of austerity policies in the Eurozone and the prospects for recovery, why the oil price hike in the mid 2000s did not result in stagflation or high unemployment as had the oil shocks of the 1970s; the consequences for the real exchange rate of the discovery of new natural gas deposits; whether fiscal policy should be delegated to a 'fiscal policy council'; the options for macroeconomic policy in an independent Scotland or Catalonia; how housing bubbles can be explained and why they occur in some economies and not others.

This book gives you a systematic way of thinking through problems like these. You will learn ways of connecting the behaviour of households, firms, governments and central banks to aggregate economic outcomes. We look at how households try to even out fluctuations in their income so that their consumption is fairly stable, but also why they may fail to do this, especially following financial crises when they are debt-constrained. Whereas household behaviour tends to smooth out fluctuations in the economy, the investment decisions of firms are lumpy. We look at how firms set wages and prices, how banks set the lending rate and how the need to maintain the continuity of banking services affects bank risk-taking. Institutions matter to macroeconomic performance and we shall see, for example, that union behaviour or whether households can re-mortgage their house when house prices rise help explain differences across countries in economic outcomes.

In many jobs you may get, whether working in finance or in think tanks, or in management or consultancies, or for governments or central banks, you need to be able to interpret national and international economic trends and policy debates. The confidence to do so comes from learning about how the different actors in the economy behave and how they interact. What are they trying to achieve and what limits their ability to put their intentions into practice? The macroeconomic environment is always changing and we need to be prepared for surprises. Knowing about previous periods of growth, stability and crisis is very helpful in preparing for shocks that might come It has frequently been said that Federal Reserve President Ben Bernanke's bold response to the crisis in 2008 owed as much to his research on the Great Depression as to his knowledge of economic models. As you work through the chapters, you will build up a picture of how the world economy has evolved since the Great Depression and how economics has affected and been affected by those developments. Chapter 8 provides a longer-run perspective and shows you how to link the analysis of macroeconomic behaviour in the short- and medium-run to long-run growth.

What this book offers graduate students, professional economists and other interested readers

Our approach is to reduce the complex, mathematically dense models used in frontier research and in central bank and fiscal forecasting exercises to a relatively simple intuitive and unified model—one that can be understood using diagrams and a small number of equations. The model incorporates realistic institutional settings in labour and product markets as well

as the financial system, and analyses closed, open and global economies It is a unified model in that the same tools are used to analyse both stability and crisis.

Long before the crisis, the economist Hyman Minsky had argued that over periods of prolonged prosperity and optimism when both data about macroeconomic volatility and the behaviour of the rating agencies signal declining risk, financial institutions invest in riskier assets. This makes the economic system more vulnerable to a crisis Minsky's insight suggests that the very tranquillity of the economy in the Great Moderation created the seeds of the crisis by causing banks to take excessive risks. As vividly depicted in the film *The Inside Job*, many macro-economists participated in the general sense of unproblematic well-being

Nevertheless the mainstream modern monetary framework, captured in the 3-equation model of the real economy (explained below), remains an important foundation of the unified model in this book. But now the link between the financial system and that mainstream real economy model is central to our approach. We also believe that the crisis has shown more clearly how consumer behaviour is affected—and therefore modelled—when the economy is operating with high unemployment and where households have high levels of debt In addition, we pay more attention to real and financial interactions among global economies than was customary in pre-crisis models, reflecting the importance of global and Eurozone imbalances. It is probably fair to say that in none of these developments is there clear consensus of the kind that had previously characterized the modern monetary framework.

The modern monetary framework: the 3-equation model

The workhorse model for macroeconomic modelling of economies with an inflation targeting central bank is the 3-equation model. It is often referred to as New Keynesian, the 'Keynesian' drawing attention to the fact that it allows for 'prolonged departures of economic activity from its optimal level as a consequence of instability in aggregate spending' (Woodford, 1999).

The three equations are the IS, which models aggregate demand, the PC, which reflects wage and price-setting behaviour in the economy, and the MR, which represents the best-response behaviour of the inflation-targeting central bank.[1] In the 3-equation model business cycles are driven by shifts in aggregate demand and supply (in contrast to the Real Business Cycle model set out in Chapter 16 where supply side shocks are the drivers of cycles) Output and employment are affected by fluctuations in aggregate demand because of structural features of the economy, often referred to as nominal rigidities, which prevent wages and prices from adjusting rapidly. The central bank is modelled as adjusting interest rates in response to shocks to the economy so as to achieve its inflation target. Aggregate demand fluctuations shift the economy away from the equilibrium rate of unemployment at which inflation is constant. The equilibrium rate of unemployment is the outcome of wage- and price-setting behaviour. Even in the absence of imperfect competition in the labour market, there is involuntary unemployment at equilibrium because employment contracts are incomplete. At the wage chosen by the firm, there are workers who would be willing to

[1] For readers who would like to see the 3-equation model presented in contrast to the old workhorse models of IS/LM and Mundell-Fleming, we refer you to our book published in 2006.

take a job but who are unemployed. Demand shocks shift output and employment away from the equilibrium, and inflation rises when the labour market tightens and falls when it slackens. Supply shocks shift the equilibrium rate of unemployment while institutional and policy differences across countries imply different national rates of unemployment.

The zero lower bound, as well as a deflation trap, can be modelled within the common framework. The best-response function of the policy maker establishes the output gap that will get the economy onto a path back to its constant inflation equilibrium. When interest rate based monetary policy is inoperative, because of the zero lower bound, the best response function is reinterpreted in terms of the stance of discretionary fiscal policy and the extent to which unconventional monetary policy such as quantitative easing can have a direct effect on the lending rate, inflation expectations and asset prices

Macroeconomics and the financial system

The widespread adoption of the mainstream monetary macro model of inflation-targeting reflected the tendency in economics teaching and research to separate macroeconomics from finance. The first step to redress this is by integrating a model of the banking system with the 3-equation model, showing how the margin of the lending rate over the policy rate is set in the commercial banking sector, how money is created in a modern banking system and how the central bank can take account of the working of the banking system in order to achieve its desired policy outcome. The basic set-up is one in which money is credit; the money supply is endogenous and the real lending rate is controlled by the central bank through its interest rate policy. Shocks to the banking system can cause macroeconomic shocks unless the central bank makes the appropriate adjustments to the policy rate.

The second step is to highlight characteristics of the financial system that can create vulnerability to a financial crisis, with implications for fiscal balance. This is ignored in the 3-equation macro model. The economy depends on the continuity of core banking services and governments cannot afford to let them fail. This means that systemically important banks do not bear the full cost of their lending decisions. They will have an incentive to take on excessive risk. Schemes of deposit insurance guard against panic-based bank runs but reduce the incentive for small depositors to monitor the risk-taking strategies of banks, even though these strategies can threaten their solvency In principle banking regulation can impose higher private capital cushions on banks to create a better alignment of private and social costs but the crisis revealed the weakness of pre-existing arrangements.

In the third step, we set out a simple model of the behaviour of highly-leveraged financial institutions as the basis for the development of a leverage or financial cycle in the economy. There is no widely agreed model of this. We use recent research by Shin (2009) and Geanakoplos (2010) as the basis of a model suitable for undergraduate level and for other non-specialist readers The model explains how risky loans made by retail banks can be transformed into marketable securities by investment banks. By connecting this to Minsky's insight that a period of prolonged macro-economic calm can incubate a crisis, we have the tools to analyse how a leverage cycle can take hold in the economy. The modelling of a financial crisis and the consequences for policy makers of a subsequent balance sheet recession is also set out within the same framework (reflecting the approach of Eggertsson and Krugman, 2012)

Open economy

The extended 3-equation model

In this new book, we extend the 3-equation model to deal with policy making in the open economy. We show how arbitrage in the forward-looking foreign exchange market interacts with the way that the central bank forecasts the effects of shocks. The interaction between these two forward-looking agents in the economy also determines the mix of exchange rate and interest rate adjustment to shocks [2]

Global imbalances

Although global imbalances were steadily building up during the 2000s, they were not seen as a cause for concern. Inflation targeting was the focus of policy makers' attention and its success in the presence of growing international imbalances in current accounts and real exchange rates prompted the development of models accounting for imbalances as equilibrium phenomena. But once the crisis started it became very apparent that we needed a clearer analysis of the interrelation of open economies in which growing imbalances could signal unsustainable combinations of growth strategies across the world's major economies We set out a simple 2-bloc version of the 3-equation model to show how the pursuit of different growth strategies in each bloc is nevertheless consistent with the achievement of the inflation targets of the central banks in each bloc.

Eurozone

Imbalances among Eurozone economies had also been largely ignored as a source of concern for policy makers. This reflected the presumption in open economy New Keynesian macroeconomic models that real exchange rates would respond in a stabilizing way to the creation of a common currency area. Little attention had been paid either to the current account and associated real exchange rate imbalances in the Eurozone or to the fragile governance structure of the Eurozone itself. All of this came into sharp focus in the Eurozone crisis of 2010.

The case of a common currency area is handled within the core model. Monetary policy making by the ECB at the level of the Eurozone can be understood using the open economy 3-equation model to explain how the Eurozone economy was able to keep close to its inflation target in the first decade of operation The core model is extended to explore the behaviour of member countries of the Eurozone. It is used to highlight how an attitude of benign neglect by national policy makers towards stabilization policy in their own country contributed to the origins of the sovereign debt crisis of 2010.

Growth and innovation

The analysis of long run growth is placed in historical perspective by presenting data from the year 1000 for GDP per capita in several European countries as well as Japan, China and India. This draws attention to the relatively recent experience of steadily rising living standards. We highlight the interconnection of great changes in technology with the emergence of

[2] For an initial presentation of this, see Carlin and Soskice, 2010

capitalism as the dominant economic system in Europe in the 19th century, which marks the beginning of the transformation of the way of life of humans across the globe.

The modelling focuses on capital accumulation and innovation as the proximate sources of growth. The Solow model provides a workhorse to explain the concept of a balanced growth path and to analyse the role of capital accumulation in long run growth. To understand the central role of technological progress in economic performance over the last 250 years, we focus on the dynamism of the capitalist economy and the role of Schumpeterian (or innovation) rents The Aghion and Howitt model of Schumpeterian growth is used to explain the role of competition in growth and to look at the relationship between fluctuations and growth.

Understanding microeconomics

Consumption behaviour: credit-constrained households and balance sheet effects

In a stable world, the permanent income hypothesis and associated Euler equation appear to be reasonable ways of modelling consumption Faced with the major disruption of the crisis and recession that followed, however, the importance of credit constraints and of balance sheet effects on the behaviour of indebted households came to the fore. The need for a richer model of consumption following the crisis also made clear the neglect by many economists of financial accelerator effects, which had operated in the household sector in the years of upswing of the financial cycle. In economies where it is possible for households to withdraw equity from their house, house price booms led to a relaxation of credit constraints, prompting households to spend more on housing and on the purchase of other goods and services.

Imperfect competition and output determination

In the book we assume that firms operate in imperfectly competitive product markets This explains why their incentive to change prices immediately in response to demand shocks is limited The combination of sluggish price adjustment with a sizeable responsiveness of consumption to changes in current income produces demand-driven fluctuations. The multiplier is therefore a useful tool in understanding business cycles

Inertial inflation expectations

In the modelling of inflation in the New Keynesian DSGE (Dynamic Stochastic General Equilibrium) models, which are routinely used in macroeconomic research, wage and price setters use rational expectations to calculate, in a forward-looking way, the expected inflation rate. In the presence of price stickiness, the result is that inflation, like the exchange rate is a jump variable that responds to forecasts of future output gaps. Estimated NK DSGE models used in policy making then include a variety of ad hoc elements to produce sufficient persistence in inflation and output to match the persistence found in the data of actual economies. In the core modelling of this book, we include the persistence of inflation explicitly in the Phillips curve by including lagged inflation. We draw the attention of students to the institutional sources of the differences in the way wages, prices and exchange rates (along with the prices of financial assets) behave. We assume, realistically, that wage setting takes place periodically. In the baseline model, wage setters update their inflation expectations based on past inflation and firms update their prices following the wage round We show how expectations that are more firmly anchored to the central bank's inflation target reduce the costs of adjusting to shocks

We explain the importance of the Lucas critique, which highlights problems raised by using a backward-looking Phillips curve when the economic or policy environment changes. These problems are less acute in an environment in which the motivation of policy makers is modelled explicitly as is the case in this book.

Non-clearing labour markets: incomplete contracts, bargaining and institutions

When an employer advertises vacancies at a particular wage or salary and hires a worker, she cannot be sure of what she will get in terms of work produced. The wage contract is by its nature incomplete and can be contrasted with a contract to buy a mobile phone, where the purchasers know exactly what they are getting Because contracts in the labour market are incomplete, there is involuntary unemployment: in order to get the worker to work hard, the firm deliberately sets the wage above the worker's reservation wage, which means there is a cost of job loss Involuntary unemployment is characteristic of the economy even when wages are flexible and there is no imperfect competition.

In many countries unions play an important role in wage-setting. Even when they don't, there is often bargaining between employer and prospective employee over the wage. This will result in wages being set above the level that the employer would choose. An exception is when unions behave in a coordinated way. This means they take into account the effect of the wage bargain on economy-wide inflation and hence on the response of the policy maker. The precise institutional arrangements for wage setting are important in understanding their implications for equilibrium unemployment.

Asymmetric information in credit markets

In credit markets, asymmetric information between lenders and borrowers means that banks will ration credit. Although households are forward-looking, some face constraints in borrowing to smooth their consumption. Other households are impatient and don't save when they could. As a result of credit constraints and impatience, demand shocks are amplified by a multiplier process.

The information asymmetry between banks and the government in turn helps explain why banks take on too much risk—they can lend 'too much'. The problem is exacerbated because of the obligation of governments to maintain the functioning of core banking services in the economy In the event that excessive risk taking by banks results in bank failures, the government normally bails them out. Banking crises become fiscal crises. The interaction between the institutions in the housing and financial markets creates the basis for a financial accelerator that can drive the upswing of a financial cycle and create vulnerability to a financial crisis.

Using the book in a variety of courses

The book is designed to be used in core macroeconomics courses at intermediate and advanced undergraduate level. To increase the flexibility of the book and to make it accessible to a wider audience, non-technical overviews are provided for the main modelling chapters. The overviews encourage students to see the structure of the argument as whole before they get drawn into the details. There are appendices to many chapters in which more technical material is set out; web appendices are used where the technical presentation is lengthier.

There is a macroeconomic simulator on the book's website developed by Javier Lozano, which allows readers to conduct a range of exercises using the closed and open economy versions of the model.

Below are illustrations of the use of different combinations of chapters (chapter numbers are in parenthesis) for a variety of courses.

Intermediate macroeconomics

Demand side (1)
Supply side (2)
The 3-equation model and macroeconomic policy (3)
Expectations (4)
Money, banking and the macro economy (5)
The financial sector and crises (selected sections of 6)
The 3-equation model in the open economy (9)
The open economy demand and supply sides (10)
Growth, fluctuations and innovation (8)
Selected applications from the chapters on Monetary policy (13)
Fiscal policy (14)
The global financial crisis: applying the models (7)

Advanced macroeconomics

Microfoundations of the demand and supply sides (1, 2, 4, 15; web appendices)
Money and the financial sector in the macro model (5, 6, 7, web appendices)
Open economy macro (9-12)
Macroeconomic policy (13-15)
Real business cycle and New Keynesian models (16)

Macroeconomic policy

Overview sections of Chapters 1-3, 9-11
Expectations (4)
Monetary policy (13)
Fiscal policy (14)
Supply-side policy, institutions and unemployment (15)
The Eurozone economy (12)

Macroeconomics of money and banking

The 3-equation model and the macro economy (3; the Overview sections of Chapters 1 and 2 as preparatory reading)
Expectations (4)
Monetary policy (13)
Money, banking and the macro-economy (5)
The financial sector and crises (6)

The global financial crisis: applying the models (7)
Banking and macro-prudential regulation (13, Section 13.6)

Labour economics: macro

Overview section of Chapter 1
The supply side (2)
The 3-equation model and the macro economy (3)
Supply-side policy, institutions and unemployment (15)

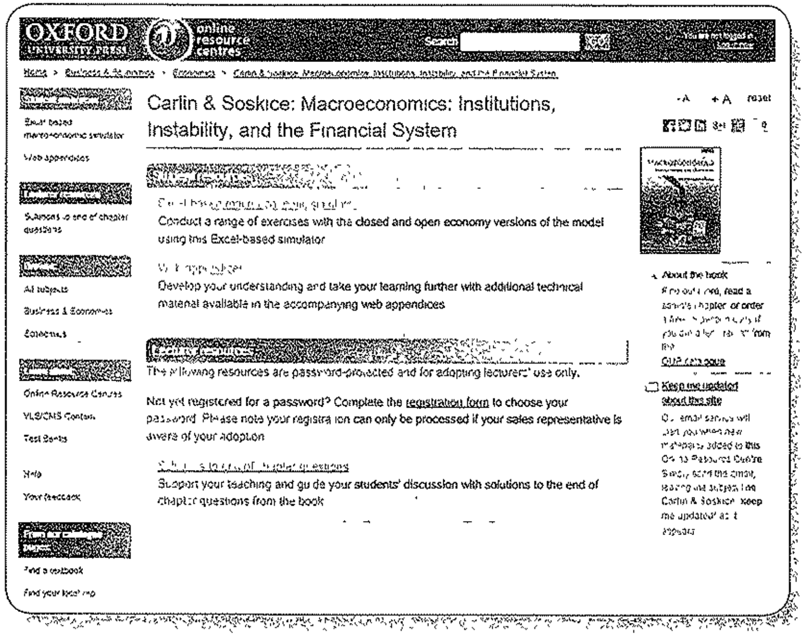
International monetary economics

Overview sections of Chapters 1 and 2
The 3-equation model and the macro economy (3)
The 3-equation model in the open economy (9)
The open economy: demand and supply sides (10)
Extending the open economy model. oil shocks & imbalances (11)
The Eurozone (12)

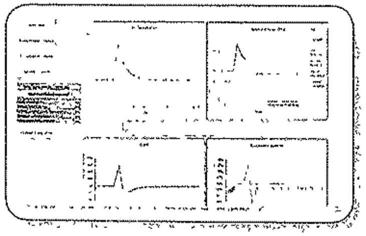
How to use the Online Resource Centre

This textbook is accompanied by a number of online resources available for students and registered lecturers.

Visit the Online Resource Centre at www.oxfordtextbooks.co.uk/orc/carlin.soskice/ to access all of the supporting content.

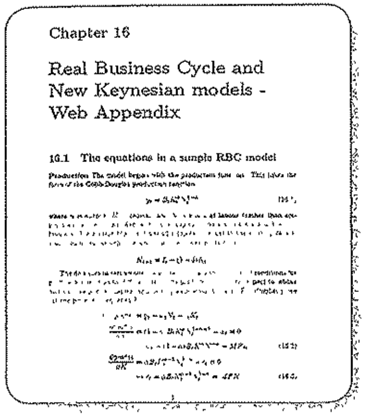


For students



Excel-based macroeconomic simulator

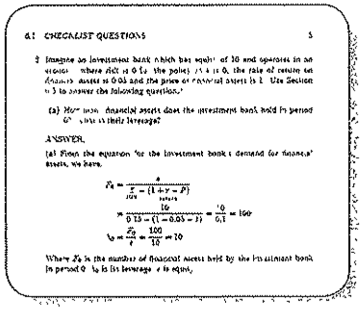
Conduct a range of exercises with the closed and open economy versions of the model using this Excel-based simulator



Web appendices

Develop your understanding and take your learning further with additional technical material available in the accompanying web appendices

For registered lecturers



Solutions to end-of-chapter questions

Support your teaching and guide your students' discussion with solutions to the end of chapter questions from the book.

Acknowledgements

This book would not have been possible without the contribution of David Hope, who has worked as a researcher - and much more - on the project over a number of years. Marvellously organized, clever, focused and calm, he has subjected our arguments to critical scrutiny and kept the interests of our readers uppermost in our minds.

The development of the simulator for students to use—they can get to know the models by trying out different scenarios—was the brainchild of Javier Lozano. We are deeply grateful to Javier for his imagination and skill in developing the simulator.

A number of people have read and commented on various drafts of the book. We want to single out in particular: Philippe Aghion, Andrea Boltho, Sam Bowles, James Cloyne, Liam Graham, Javier Lozano, Colin Mayer, John Muellbauer, Bob Rowthorn, Hyun Shin, Bob Sutcliffe, David Vines and Stephen Wright. Their reactions and suggestions have immeasurably improved the final product and we are indebted to them for their consistent intellectual support.

Many others have given us comments and ideas for improvement: Chris Allsopp, Hannes Ansorg, Erik Berglof, V. Bhaskar, Richard Blundell, Ciaran Driver, Christian Dustmann, Fabian Eckert, Jonathan Glyn, Ian Goldin, David Goll, Georg von Graevenitz, Peter Hall, Bob Hancke, Soren Harck, Matthew Harding, Frank Harman, David Hendry, Michal Horvath, David Howell, Torben Iversen, Arjun Jayadev, Cloda Jenkins, Simon Johnson, Tom Josephs, Vijay Joshi, Michael Landesmann, Kieran Larkin, Deborah Mabbett, Steve Machin, Andy Martin, Bill Martin, Felix Martin, Massimo di Matteo, Costas Meghir, Jacques Melitz, Thomas Michl, David Miles, Hoang Minh Vu, Jamie Murray, Steve Nickell, Andrew Oswald, Susana Parrago Rodriguez, Tom Pybus, Ming Qiu, Nick Rau, Andy Ross, Mark Schaffer, Waltraud Schelkle, Paul Seabright, Rajiv Sethi, John Sharma, Luigi Spaventa, Andrew Sykes, David Tuckett, John Van Reenen, Charles Vincent and Simon Wren-Lewis. The book has benefited greatly from Wendy's engagement with the Institute for New Economic Thinking at the Oxford Martin School and in New York, and especially from joint work with the members of the CORE project. The reviewers commissioned by OUP provided very helpful guidance.

We had much to learn about the financial system when we began this project. Our 'teachers' have been extraordinarily generous with their time and expertise. We would especially like to thank Claudio Borio, Charles Goodhart, Sujit Kapadia, J.R. (Dick) Sargent, Hyun Shin, and John Vickers.

We have benefited from using teaching notes of Chris Bowdler, David Laibson, Campbell Leith, Sebastien Walker, Karl Whelan, and Simon Wren Lewis.

For help with data and specialist knowledge, we are particularly indebted to Nick Bloom, Nick Crafts, Bishnupriya Gupta, Mathias Drehmann, Kevin O'Rourke and Luigi Pistaferri.

Many thanks to Davide Melcangi who has worked out the answers to the questions and to Kieran Larkin who has done the technical proof-reading, making many improvements.

Wendy's colleagues in the Department of Economics at UCL have very generously supported this project, which has been judged by successive heads of department (Costas Meghir, Steffen Huck and Morten Ravn) a worthwhile use of research time. Fatima Cetin,

John McGlynn and Brian Wallace have provided excellent technical support and Tina Fowler and Nirusha Vigi have helped keep morale high.

Students at UCL in several courses over a number of years have provided feedback on various parts of the book. Their questions and suggestions have pushed us to improve the presentation and fine-tune the coverage. Their enthusiasm has kept us going. Doctoral students in political science at Duke have been a great and critical audience for various parts of the book; and encourage us to believe that it can be used to teach macroeconomics to well-trained social scientists.

We would like to thank the team at OUP led by Kirsten Shankland and Jo Hardern for their cheerful and effective marshalling of the book through the processes of review, revision and production.

Our families and friends have been patient and encouraging. Their probing questions about the crisis and their scepticism about the potential for economists to contribute to dealing with it have provoked us to make the arguments clearer and more accessible to non-specialists. Special thanks from Wendy to Jonathan, Tessa, Lucy and Miles Glyn, to Joan, John, David, and Michael Carlin, and to Mayerlene Engineer, Lizzie Fricker, Liz Harman, Carinna Hockham, Rosemary Joseph, Marion Kozak, Belinda Probert, and Jonathan Rée.

And special thanks for support and close friendship from David to William Soskice, Juliet Soskice-Rosenfeld and Andrew Rosenfeld, to Catherine Soskice and Vikram Gandhi, Oliver, Janet and Izzy Soskice, Ciaran Driver and Anne Phillips, Rebecca Murray and Kevin Prosser, and to Torben Iversen and Charla Rudesill - and in particular, thanks for the probing questions from William and Ciaran. David takes this opportunity to acknowledge the key intellectual relationships he has had during this period with Torben Iversen and with Niki Lacey, which have added greatly to the book from a broader perspective.

We dedicate this book to our two spouses, with deepest love:

Niki Lacey, David's wife and very close friend of Wendy, well-known legal and feminist theorist, has been central in every way to the writing of this book, providing sustenance of all sorts (including on and round the kitchen table), raising occasional eyebrows at the excess of rational choice, and being an inspiration affectively and intellectually for both of us.

Andrew Glyn, Wendy's husband and one of David's closest and oldest friends, died quite unexpectedly just as we had started planning this book. At the peak of his highly influential intellectual career (he already has an entry in the Dictionary of National Biography), he had just finished his prescient book Capitalism Unleashed (OUP 2007). Our other two macroeconomics books (1990 and 2006) had been written with his continuous and unstinting help; his presence and luminosity (and modesty) have remained with us as we wrote this book; but on many occasions we have been keenly aware of his absence.

Outline contents

Contents

List of figures

List of tables

Abbreviations

ABS	Asset backed securities
AD	Aggregate demand
APK	Average product of capital
BIS	Bank for International Settlements
BoE	Bank of England
BT	Balance of trade
CA	Current account
CAGR	Compound annual growth rate
CBO	Congressional Budget Office
CCA	Common currency area
CDO	Collateralized debt obligations
CDS	Credit default swap
CPB	Central Planning Bureau
CPI	Consumer price index
CRA	Credit rating agency
CRS	Constant returns to scale
DSGE	Dynamic stochastic general equilibrium
ECB	European central bank
ERU	Equilibrium rate of unemployment
FOMC	Federal Open Market Committee
FPC	Fiscal policy council
GDP	Gross domestic product
IB	Investment bank
ICA	Intertemporal model of the current account
ICT	Information and communications technology
IMF	International monetary fund
LIBOR	London inter bank offered rate
LOLR	Lender of last resort
LOP	Law of One Price
LTV	Loan-to-value
MBS	Mortgage backed securities
MM	Modigliani-Miller
MMF	Money market funds
MPC	Monetary Policy Committee
MPK	Marginal product of capital
MRE	Medium-run equilibrium
MRS	Marginal rate of substitution
NAIRU	The non accelerating inflation rate of unemployment
NBER	National Bureau of Economic Research
NFA	Net foreign asset
NK	New Keynesian
NKPC	New Keynesian Phillips curve
OBR	Office for Budget Responsibility

OECD	The Organization for Economic Co-operation and Development
PC	Phillips Curve
PDE	Price dynamic equation
PFPR	Prudent fiscal policy rule
PIH	Permanent income hypothesis
PPP	Purchasing Power Parity
PR	Policy Rule
PS	Price setting
QE	Quantitative easing
RBC	Real Business Cycle
REH	Rational expectations hypothesis
SGP	Stability and Growth Pact
SIV	Structured investment vehicle
SOMA	System Open Market Account
SRM	Single Resolution Mechanism
SME	Small and medium-sized enterprises
TARP	Troubled Asset Relief Program
TFP	Total factor productivity
UIP	Uncovered interest parity
VaR	Value at risk
VAT	Value added tax
WS	Wage setting
ZLB	Zero lower bound

1 The demand side

1.1 Overview

In the late 2000s, households and firms across the world cut back their spending and the global economy went into recession. There is little disagreement among economists that the dramatic fall in global GDP in 2008–09, which is now known as the global financial crisis, was a negative shock to aggregate demand.

This chapter begins the task of building a macroeconomic model with the demand side of the economy, which concerns the spending decisions of households, firms and government. As we shall see in Chapter 2, the supply side refers to the production activities in the economy. Dramatic changes in aggregate demand—that is, spending in the economy as a whole—as occurred in 2008-09 are unusual. In more normal times, shifts in aggregate spending decisions as well as shifts on the supply side of the economy are sources of the irregular fluctuations from recession to boom that form the business cycle.

If a period of depressed spending is forecast, there is much discussion in the financial press about whether the central bank will intervene and offset the likely recession by cutting the interest rate. The central bank would cut the interest rate because it expects this to encourage spending and help return the economy to stability. Similarly, if a boom in spending were forecast, the central bank would try to dampen it down by raising the interest rate. As well as the central bank, the government also needs to know how spending patterns are likely to evolve and affect output. A recession will depress tax revenue and increase spending on unemployment benefits. Forecasting patterns of spending is therefore a priority, not only for businesses, but also for the monetary and fiscal policy makers.

Aggregate demand and spending decisions

In this chapter, we focus on what lies behind the spending decisions in the economy and how they influence the level of economic activity. By the level of activity, we mean output or income. When output changes, employment also changes. For example, when output rises, more workers or longer hours are needed to produce the higher level of output. With more hours worked, the wage bill is larger and with higher sales, total profits are higher. This is why changes in economic activity are thought of as both changes in output and in income (i.e. wages and profits). When we think of a real-world economy, output is normally growing and recessions and booms produce fluctuations around a trend growth rate. However, we will often simplify by working with levels of output rather than its rate of growth.

Spending decisions are complex. For consumers they involve both a static component (what shall I buy today given my current income and the prices of the goods and services that are available?) and an intertemporal one (how do I allocate my spending over time given my expectations about how my income will evolve in the future?). Decision making for firms and the government also involves an intertemporal aspect. Firms make decisions to purchase machinery, equipment and premises based on a business plan that includes forecasts about how input costs and demand for their products will evolve over time. The government must also forecast demographic trends when making plans for building schools and hospitals.

Decisions by the entities that make up the economy—firms, families and government bodies—lie behind the demand side of the economy. However, macroeconomics is concerned with the aggregate sum of spending decisions of these groups, and the consequences of those decisions for economy-wide outcomes such as the rate of unemployment or inflation. Household spending decisions add up to aggregate consumption, C; firms' investment decisions add up to aggregate investment, I (note that I refers to spending on machinery, equipment and new houses and other buildings); and government spending on different goods and services adds up to a single number, G.

Including the purchase of new housing as part of investment highlights the difference between on the one hand, aggregate demand, which refers to spending on goods and services and on the other hand, the purchase of assets such as company shares or second-hand property. The purchase of a second-hand house does not contribute to aggregate demand. It is the transfer of the ownership of an asset from one household to another. In contrast, the building and selling of a new house uses resources and produces income, hence it influences aggregate demand.

Although the first chapters in the book relate to the closed economy (a single nation without links with others) as noted above, in the fuller model (set out in Chapters 9 and 10) the demand side includes foreign spending on home goods and services, exports (X), and domestic spending on foreign goods and services, imports (M). Aggregate demand is real expenditure on goods and services produced in the home economy. This can be summarized by an equation relating real expenditure, which is called y^D, to its individual components:

$$y^D = C + I + G + (X - M), \qquad \text{(aggregate demand)}$$

where we add to the trio of C, I and G, the expenditure of foreigners on home output, X and subtract the spending of home agents on output produced abroad, M.

Aggregate demand and government policy

Two of the major tools of macroeconomic policy, monetary and fiscal policy, work by influencing different elements on the demand side. Policy makers worry about fluctuations in aggregate demand because they affect unemployment and inflation.

Monetary policy seeks to stabilize aggregate demand by changing interest rates, which affect the investment decisions of firms and the purchase of durable goods like new cars and houses by households. A rise in the interest rate increases the cost of financing investment projects, and projects that would have gone ahead with lower interest rates are postponed or cancelled. Monetary policy also has indirect effects because the interest rate affects incentives to save and therefore shifts spending decisions over time. For example, by making new

borrowing more expensive and increasing the return on saving, a higher interest rate will encourage households to postpone consumption.

Through changes in government spending on goods and services (G), fiscal policy affects aggregate demand directly. Fiscal policy can also be used to affect demand indirectly through its influence on household incomes and through that channel, on household spending. This is how changes in taxation and in the transfers made by the government to households in the form of pensions, disability and unemployment benefits feed through to affect aggregate demand. We will denote taxes minus transfers by T.

An important reason to study the demand side is to construct a model of the transmission mechanism by which monetary and fiscal policy, via the spending decisions of households, firms and the government, affect the economy. Before explaining the first building block of the model, which is the *IS* curve, we provide some facts about the demand side.

1.1.1 Facts about the demand side and business cycles

Shares of GDP

Table 1.1 sets out the average composition of gross domestic product (GDP) in five major economies between 2000 and 2005. GDP is the national accounts measure of national output. We discuss the calculation of GDP next subsection. Table 1.1 shows that consumption makes up the largest proportion of GDP in all the five economies. However, the importance of consumption in GDP ranges from 43% in China to nearly 70% in the US. A large part of this cross-country variation can be explained by differences in the contribution of investment. For example, an exceptionally high proportion of Chinese GDP arises from investment spending, which crowds out consumption spending. The table also highlights the variation across countries in the contribution to GDP from net exports. On average, between 2000 and 2005, the UK and the US ran current account deficits with exports in excess of imports, whereas the other three countries were in surplus (see Chapter 10 for further discussion and definitions).

The upper panel of Fig. 1.1 shows how the shares of consumption, investment and government spending in GDP changed over time for the UK from 1948 to 2010. The share of consumption follows a shallow U shape over the period; investment is mildly hump shaped and the share of government spending drifts upward throughout the period. Government spending is dominated by spending on services (e.g. health and education) where

Table 1.1 Shares of GDP (in %), current prices, average 2000–2005.

	Consumption	Investment	Government spending	Net exports
China	43.0	37.5	15.1	2.9
Germany	58.6	18.7	19.0	3.5
Japan	57.1	23.7	17.7	1.4
United Kingdom	65.3	16.7	20.0	−2.5
United States	69.9	19.0	15.3	−4.5

Note: GDP in current prices is a measure of nominal GDP; shares might not add up to 100% due to stockbuilding and statistical discrepancies.
Source: OECD National Accounts (data accessed November 2011).

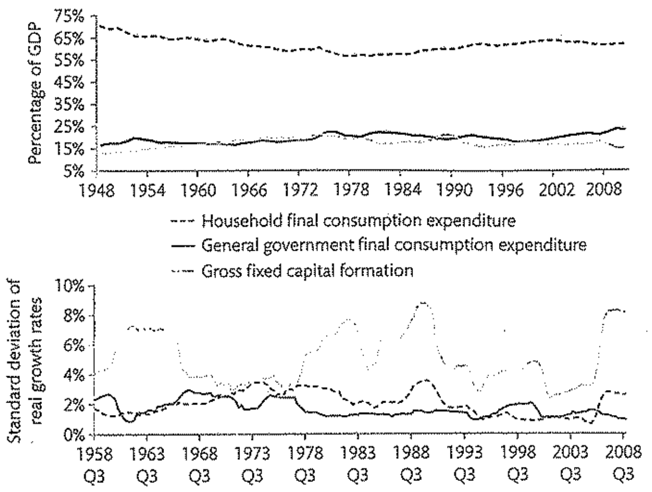


Figure 1.1 Components of GDP in the UK: as a percentage of GDP between 1948 and 2010 (upper panel) and volatility of growth rates between 1958 Q3 and 2008 Q4 (lower panel).

Source: UK Office for National Statistics (data accessed October 2011).

Note: Volatility has been calculated as the standard deviation of the GDP growth rate over a rolling 21-quarter period. The upper graph uses data from the series 'Gross Domestic Product by category of expenditure, current prices' and the lower graph uses data from the series 'Gross Domestic Product by category of expenditure, chained volume measures'.

productivity growth is slower than in industry. In sectors with comparatively slow productivity growth, prices rise relative to the rest of the economy. This is an important structural reason for the tendency for government expenditure, when measured in current prices, to rise as a proportion of GDP.

National accounts

The national accounts are used to measure the output of an economy. The most commonly used measure for calculating national output is gross domestic product, or GDP. GDP can be measured in three different ways. All three equations for calculating GDP are identities, which means that because of the way the variables are defined, the left hand side must always equal the right hand side. This special feature of an identity is signalled by the use of the equals sign with three bars.

Firstly, the *expenditure* method, which is one we use in the model of the demand side. This method measures GDP as the total expenditure on the economy's output of goods and services:

$$y \equiv C + I + G + (X - M), \qquad \text{(GDP identity, expenditure method)}$$

where $y \equiv$ output (or GDP), $C \equiv$ consumption, $I \equiv$ investment (including changes in stocks of raw material and finished goods), $G \equiv$ government spending and $(X - M) \equiv$ net exports. This is an identity, as it simply breaks down GDP into its constituent components. A fall in one of the components on the right hand side, for example consumption, will therefore always result in an equivalent reduction in measured GDP on the left hand side.

When a *new* house is bought, just like a new piece of machinery, this is an investment decision—the house provides a flow of services to the household over many years—and is included in the national accounts under I. The house and the piece of machinery form part of the economy's capital stock. National accountants have to make many tricky classification decisions: although households buy a variety of durable goods that provide services over many years, by convention it is only housing that is considered as investment in the national accounts. Cars and furniture, for example, are treated in the national accounts as consumption.

It is important to note that only government spending on goods and services is part of aggregate demand. G does not include government expenditure on transfers (e.g. pensions or social security payments). When the recipients of transfers spend their income from benefits or pensions, this is then recorded as consumption.

The expenditure method only includes *final* goods. For example, during the production process, firms buy raw materials and *intermediate* goods to make their final goods. The only purchases that are counted in GDP are when the firm sells the finished goods to the consumer. This avoids any double counting. In this way, GDP captures only the *value added* created in the economy. This leads onto the second approach for calculating GDP; the *value added* method. This method measures GDP as the value added created in all sectors of the economy, such that:

$$y \equiv \text{value of output sold} - \text{costs of raw materials and intermediate goods.}$$

(GDP identity, value added method)

The third and final approach for calculating GDP is the *income* method. This method measures GDP as the total income of all agents in the economy, such that:

$$y \equiv \text{salaries of workers} + \text{profits of the owners of capital.}$$

(GDP identity, income method)

The three methods of calculating GDP are all identities and each holds at each point in time. It makes intuitive sense that the total income in the economy is equal to the total expenditure, because for every transaction there is both a buyer and a seller. What is expenditure for the buyer is income to the seller. In practice, however, GDP calculated using the three methods may differ. This is a matter of *measurement error*, which can arise for several reasons. Measurement error may be more serious for one rather than another of the three methods. For example, tax evasion or the existence of a substantial black market are likely to produce a greater problem of underestimation of GDP when measured by incomes than for the other methods. If the underlying components could be measured completely accurately, then each of the three approaches would yield exactly the same estimate of GDP.

Relative volatility

The lower panel of Fig. 1.1 shows the relative volatility of growth rates for the three components of aggregate demand: a higher standard deviation means higher volatility. We can see that investment is much more volatile than consumption and government spending:

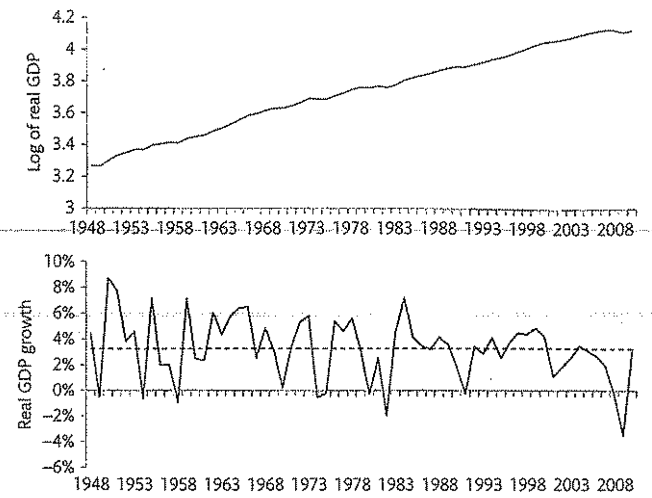


Figure 1.2 Business cycles in the United States between 1948 and 2010–Log of real GDP (upper panel) and real GDP growth (lower panel).

Source: US Bureau of Economic Analysis (data accessed October 2011).

Note: The shaded grey areas represent recessions as defined by NBER. Both graphs use data from the series 'Real GDP, chained dollars, billions of chained (2005) dollars'.

the line showing the standard deviation for investment lies above that for consumption and government spending. Investment depends on expected post-tax profits and is very dependent on how optimistic firms are, so it tends to flourish in boom periods and collapse in recessions, making it more volatile than the other components of GDP, or GDP itself. In addition to this, investment can also be postponed in recessions, whereas government spending and consumption cannot be as easily delayed. For example, a household still has to spend money on food and drink in a recession, whereas a firm may choose to wait until the economy has recovered before it undertakes investment. We shall also see—in Section 1.2.6—that fixed investment decisions are often bunched, further accounting for the greater volatility of investment.

Growth and cycles

The idea that economies fluctuate between phases of boom and recession is confirmed by looking at the data for the USA. Figure 1.2 illustrates the long-run growth and business cycles in the US economy. In the top panel, the log of GDP in constant prices is plotted and shows the rather steady growth rate of GDP over the long run. Using the log scale, a straight line would represent a constant growth rate from 1948 to 2010.[1] In fact, annual growth rates in each decade from 1948 were: 3.3%, 4.8%, 3.2%, 3.0%, 3.1% and 2.5%.

[1] See Section 3 of Chapter 8 for a detailed discussion of growth concepts and the use and interpretation of natural logs.

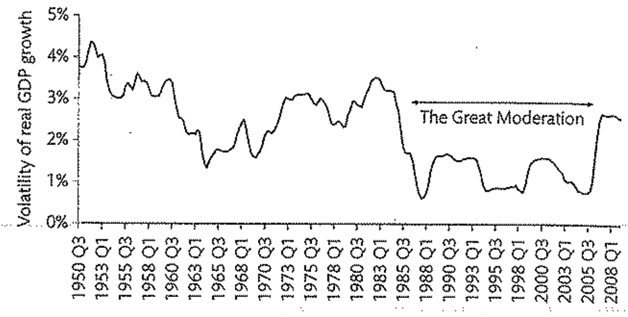


Figure 1.3 Volatility of real GDP growth in the United States: 1950 Q3–2009 Q1.

Source: US Bureau of Economic Analysis (data accessed October 2011).

Note: Volatility has been calculated as the standard deviation of the GDP growth rate over a rolling 21-quarter period. Graph uses data from the series 'Real GDP, chained dollars, billions of chained (2005) dollars'.

By plotting the *annual* growth rate of GDP in the lower panel, the fluctuations of the economy around the long-run trend are highlighted. The dotted line represents the average growth rate between 1948 and 2010 of 3.3%. Peaks and troughs of growth occur about twice a decade. In the US, an independent body called the National Bureau of Economic Research (NBER) establishes the dates of US business cycles and states there have been eleven separate contractions (i.e. recessions) over this period, which are represented by the shaded grey areas on the graphs. Recessions start at the peak of a business cycle and end at the trough: during recessions, economic activity is contracting. In the period we are analysing, the recessions range from 6 to 18 months in duration. This is why the *annual* GDP growth rate shown in Fig. 1.2 is not negative in all of the recessions (e.g. the 2001 recession). The longest of these recessions was the 2007–09 global financial crisis (GFC), which spanned a year and a half.

Before the global financial crisis, many macroeconomists claimed that better policy making had produced greater macroeconomic stability. This led to the use of the term 'the Great Moderation' to describe the calmer macroeconomic conditions from the mid-1980s (fewer and shorter recessions as illustrated in Fig. 1.2).

A more precise way to document the Great Moderation is to look at the volatility of GDP growth. Figure 1.3 shows the volatility of US GDP growth since the late 1950s. By plotting the standard deviation of the growth rate over a 21-quarter rolling period, changes in volatility can be seen easily. There is a noticeable fall in the volatility of GDP growth from the mid 1980s. However, the spike in volatility experienced during the recent recession brought the era of the Great Moderation to an end.

1.1.2 Introducing the *IS* curve

In this chapter, we begin to build up a way of capturing essential elements of the complex real-world macro-economy in a model. The model is called the 3-equation model because there are three core elements, representing the demand side, the supply side and the policy maker. The *IS* is the curve that represents the demand side. It is called the *IS* curve because it refers to planned investment and savings decisions; hence the *I* and the *S*. It was originally formulated as a simple shorthand version of John Maynard Keynes's description of the demand side in his General Theory.[2]

Models are important for providing a framework in which to think about complex problems more easily and more carefully. By modelling the goods market we can answer interesting questions about the demand side that we might otherwise struggle with, such as:

1. If you give an individual a $100 bonus, how much of it will they spend?
2. How much will output increase following a rise in government spending or a boost in the confidence of households and firms about their future prospects?
3. To what extent will a rise in the interest rate curtail investment in new housing, machinery and equipment?

The *IS* curve is a way of summarizing the way that aggregate output in the economy is affected by changes in the spending decisions of families, firms and government bodies. For example, when firms increase their spending on new equipment, this triggers increased production in the capital goods producing sector of the economy. More people are employed to produce the extra capital goods and as they spend their wages, demand for consumer goods goes up and employment and output expand in those sectors as well. As a result of this process, the economy will move to a higher level of output and employment. We shall see that in response to the initial expansion of spending, output will expand until the extra *S*aving households want to make just balances the extra spending on *I*nvestment. At that point, the process by which the impact of the initial increase in aggregate demand is multiplied through the economy in subsequent rounds of higher employment, spending by newly employed workers, higher demand for consumer goods, etc. comes to an end.

Aggregate demand in the private sector

Focusing attention on decisions by the private sector, the level of aggregate demand will be affected by current income and by the following factors.

1. *Expectations about the future*: the plans of firms to invest in new equipment and premises depend on their expectations of *future post-tax profits*. If firms anticipate high levels of capacity utilization and strong order books, they will increase investment in new capacity. Households prefer to have smooth rather than fluctuating consumption, which means they need to save and borrow in order to spread their consumption more evenly over time. To make their saving and borrowing decisions, they must form a view about the future growth of their income. The *life-cycle motive* for saving refers to the planning of saving taking into account the projected pattern of income during a person's working life and retirement. Households will revise upward their estimate of how much they can spend each period if they have new information that leads them to expect their income to grow more strongly.

[2] See Keynes (1936). The *IS* curve also underlies the famous *IS/LM* model, which was introduced by John Hicks in 1937 (Hicks, 1937). For a concise and interesting discussion of its origins and impact, see Durlauf and Hester's article entitled *IS-LM* in the New Palgrave Dictionary of Economics: 2nd edition (2008).

Firms and households form their expectations in the face of uncertainty. For example, an increase in the unemployment rate in the economy could be a signal to households that *uncertainty* about their future income has risen, which would trigger an increase in saving for *precautionary* reasons.

2. The extent of *credit constraints*: these arise because of problems faced by banks in assessing the creditworthiness of households and firms. It is impossible for banks to have full information about borrowers' projects and actions. Borrowing by households and by small and medium-sized firms is therefore often restricted by banks. Households and business that cannot borrow as much money as they would like are said to be *credit constrained*. The information problems inherent in bank lending mean that access to credit is often highly dependent on the amount of *collateral* the borrower has with which to secure the loan. The most common form of collateral for households is the value of their house. This means that changes in the value of collateral, which occur for example when house prices change, affect consumption and investment because they either tighten or relax credit constraints.

3. The *interest rate*: there are a number of channels through which the interest rate will affect aggregate demand. When interest rates go up, households find it more expensive to get a mortgage. This reduces the demand for new houses and reduces the demand for furnishings and other consumer durables that go along with moving house. A higher interest rate will lead firms to rein in their spending plans on new capital equipment and buildings. Households will tend to postpone consumption spending because of the improved returns from saving. However, creditor and debtor households would be expected to react differently. For example, a creditor household will find their income has gone up when the interest rate rises and this will boost spending through the income effect. For debtor households, the effect will be the opposite. The first and third effects are normally stronger than the second one and a higher interest rate is associated with lower aggregate demand.

The IS curve

The *IS* curve is a way of summarizing in a diagram the demand side in the macroeconomic model. It shows the combinations of the interest rate and output at which aggregate spending in the economy is equal to output.

Figure 1.4 shows the *IS* curve. This is a downward-sloping relationship. To see this, think of the combination of a high interest rate and low output. When the interest rate is high, spending on housing, consumer durables, machinery and equipment will be low. This means aggregate demand is low and a low level of output will satisfy the low demand. Hence, we have the first point on the *IS* curve. Now, take a combination of a low interest rate and high output. Here the situation is the opposite: buoyant spending on new houses, consumer durables and investment goods generates a high level of output and high incomes for households. This is the second point on the downward-sloping *IS* curve.

To show how changes in profit or income growth expectations, uncertainty and the value of collateral can be captured in the diagram, we hold the interest rate constant and look at shifts in the *IS* curve. As an example, in a situation of depressed profit expectations, we would expect firms to postpone new investment. The result is lower investment spending at any interest rate. The *IS* curve shifts to the left (shown by the curve labelled *IS* (pessimistic expectations)).

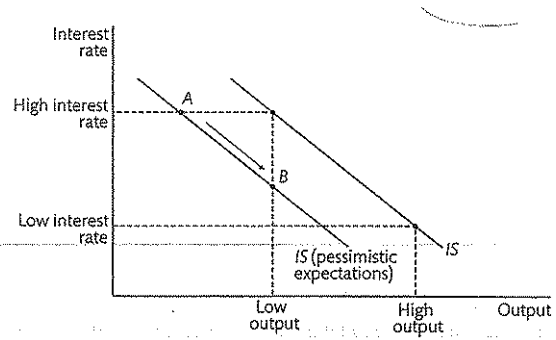


Figure 1.4 The *IS* curve—the effects of changes in optimism and economic policy.

Using the *IS* diagram, we can show how the central bank or the government can affect the demand side. Following up the situation where the business environment becomes more pessimistic, the central bank could lower the interest rate to stimulate investment. This would be shown as a move *along* the *IS* curve from point *A* to point *B*. As an example, there was much discussion about the decision by the US central bank, the Federal Reserve, to reduce the policy interest rate and keep it low as a way of stimulating investment in the aftermath of the collapse of the Dotcom bubble in 2001. The long period of low interest rates stimulated investment in new house construction, and we shall see in Chapter 7 the role that this played in the background to the global financial crisis of 2008.

Another response to a leftward shift of the *IS* curve due to pessimistic profit expectations could be action by the government rather than central bank. If the government decides to launch a major expenditure programme, such as to improve healthcare provision or to install an information super-highway, this will *shift* the *IS* curve to the right: at a given interest rate, the government purchases a larger amount of goods and services. Under our assumption that suppliers will respond to the higher demand, the economy moves to a higher level of output and employment. In the diagram, this would shift the *IS* curve to the right back towards its initial position.

A much debated issue in macroeconomics is by how much a government expenditure programme would increase output in the economy. Would extra government spending increase output one-for-one, by more or by less? This debate is frequently couched in terms of the size of the 'multiplier': i.e. by how much would expenditure of an extra dollar raise GDP? As we shall see in Section 1.2, in the *IS* diagram, the larger the multiplier,

1. the larger is the rightward shift of the *IS* curve associated with any given amount of additional government spending and
2. the flatter is the *IS* curve, since with a larger multiplier, a given cut in the interest rate has a larger effect on output.

The demand side is only part of the macroeconomic model. To understand how the spending decisions of households, firms and the government fit into the bigger picture, we need to include the supply side and the motivations and behaviour of the policy makers.

Chapter 2 addresses the supply side and Chapter 3 brings in the policy maker. Before moving on, we set out the modelling of the demand side in more detail.

1.2 Modelling

The Modelling sections of the chapters in the book provide more details of the models and their components introduced in the Overview sections. Detailed derivations of some of the results are available for download from the book's website.

In order to better understand the demand side of the macro-economy, we build a simple model, which is captured by the *IS* curve. This section focuses on the closed economy *IS* curve and we will come back to the open economy *IS* curve in Chapters 9 and 10. The closed economy *IS* curve splits demand into three components; consumption, investment and government spending.

As a first step in this chapter, we introduce the concept of goods market equilibrium and the definition and mechanics of the multiplier process. Throughout this section, we assume that firms are willing to meet the higher demand for their goods and services, and workers are willing to take the extra jobs or work the extra hours that are offered. In simple terms, we assume that the supply of output adjusts to meet the demand for goods, services and labour. The next step is to bring in additional determinants of aggregate spending listed under points (1) to (3) in the previous section and to modify the *IS* equation to take these into account. Along the way, we refer to empirical evidence to support the richer model.

1.2.1 **Goods market equilibrium**

We begin our discussion and modelling of the goods market by considering each of the components of aggregate demand in the closed economy. The aggregate—or economy-wide—demand for goods and services consists of:

Consumption demand: expenditure by individuals on goods and services. Spending is on both *durable* products such as a car, laptop or sofa and on *non-durable* products such as a theatre show or groceries.

Investment demand: expenditure on capital goods (machinery, equipment, and buildings). Spending is by households on new houses, by firms on new capital goods, including structures, and building up inventories of materials or finished goods, and by government (on public infrastructure projects, such as building a new high speed rail line).

Demand stemming from government purchases: government expenditure on salaries, goods and services. Spending includes public sector wages (e.g. civil servants, doctors), purchases of goods (e.g. educational supplies, ammunition for the army) and purchases of services (e.g. waste disposal and contract cleaning).

The model of goods market equilibrium

We now set out a model of the goods market. In the closed economy, aggregate demand, y^D, is given by

$$y^D = C + I + G, \tag{1.1}$$

and we ignore imports and exports. All variables are in real terms, which is also referred to as constant price terms. Equilibrium in the goods market requires that planned real expenditure on goods and services (i.e. aggregate demand) is equal to real output

$$y^D = y, \qquad \text{(goods market equilibrium)}$$

where y is output. y is also income: spending on the output of the economy in turn becomes the income of those producing it (wages and profits). This circular flow of income to expenditure and output, and back to the incomes of producers means we can use the terms output, income and expenditure interchangeably. We shall come back to the goods market equilibrium condition after providing some more detail about consumption and investment behaviour.

To start, we assume that aggregate consumption is a simple linear function of after-tax or disposable aggregate income

$$C = c_0 + c_1(y - T), \tag{1.2}$$

where T is total taxes net of transfers.

To get to the consumption function we use as the core of the model of the demand side, we make the additional assumption that taxes are a fixed proportion of income—i.e. $T = ty$, where $0 < t < 1$. The consumption function then becomes

$$C = c_0 + c_1(1 - t)y. \qquad \text{(Keynesian consumption function)}$$

This is referred to as a Keynesian consumption function. It consists of a constant term, c_0, which is often referred to as autonomous consumption, and $c_1(1 - t)y$, which shows households spending a fixed proportion of their disposable income. c_1 is referred to as the marginal propensity to consume and lies between zero and one ($0 < c_1 < 1$).

Note that the marginal propensity to consume (MPC) shows the change in consumption as the result of a change in post-tax or disposable income:

$$MPC \equiv \frac{\Delta C}{\Delta y^{disp}} = c_1, \qquad \text{(marginal propensity to consume)}$$

where $y^{disp} = (1 - t)y$.

In this model of consumption, saving is $S = y - C$ and adds to one's assets. Income is the amount that can be consumed without reducing one's assets. The corresponding marginal propensity to save is s_1 and $c_1 + s_1 = 1$ (i.e. disposable income can be either saved or spent).

If disposable income is zero, the consumption function predicts consumption equal to c_0. For this to be the case, households must have some savings and or they must be able to borrow. A more satisfactory model of consumption must therefore include forward-looking behaviour in order to account for savings, and we build this in to the model of consumption later in the chapter.

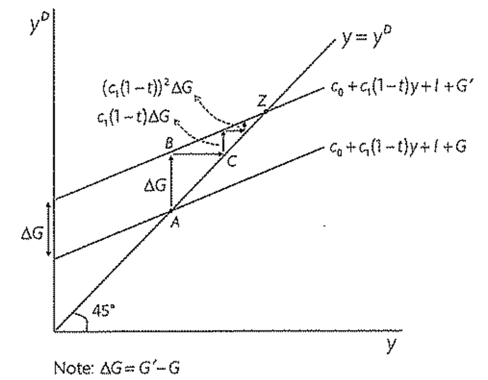


Figure 1.5 Keynesian cross—Increase in government spending.

1.2.2 The multiplier

We begin with the Keynesian consumption function in order to introduce in the simplest way the concept of how equilibrium in the goods market is determined and how the multiplier process works. If we substitute the consumption function into the equation for aggregate demand, then we get a relation between aggregate demand,

$$y^D = c_0 + c_1(1 - t)y + I + G, \qquad \text{(aggregate demand)}$$

and output, y, which we can draw on a diagram with aggregate demand on the y-axis and output on the x-axis (Fig. 1.5). The intercept of the curve will be $(I + G + c_0)$ and the slope $c_1(1 - t)$ which, given the assumptions that both c_1 and t are between 0 and 1, will itself be between zero and one. We can also draw the goods market equilibrium condition on the same graph, which will be a 45° line, because in equilibrium

$$y = y^D.$$

The point at which the two curves intersect shows the level of output where planned real expenditure on goods (and services) by firms, households and government is exactly equal to the level of goods being supplied in the economy. To the left of the 45° line, the demand for goods is greater than the supply.

To understand how the model works, we disturb the initial equilibrium at point A. Suppose there is an increase in government spending. From the aggregate demand equation, we can see that this shifts the aggregate demand curve upwards by the change in government spending, ΔG. Aggregate demand now exceeds output (Point B). As the government increases its purchases of goods (e.g. of office equipment) the stocks of these goods in warehouses decline. The inventory management software records the fall in stocks and triggers an increase in production: output rises. This is the move from B to C, where once again $y = y^D$. The higher output in turn raises incomes (in the form of wages of the additional workers employed and the profits of the owners of the firms making higher sales) and according to the consumption function, some proportion of the higher incomes are spent on goods and services in the economy, raising aggregate demand further. The process continues until the

We model the proposal to encourage savings by a fall in autonomous consumption from c_0 to c_0'. Aggregate demand falls initially by $(c_0 - c_0')$ and the multiplier process works in the downwards direction. Using the same logic as illustrated in Fig. 1.5 in reverse, this process continues until we get to the new equilibrium at lower output. At the new lower output level, y', planned savings on the left hand side of Equation 1.9 is equal to planned investment plus government spending on the right hand side, which have not changed:

$$(1 - c_1(1 - t))y' - c_0' = \bar{I} + \bar{G}. \tag{1.9}$$

An important insight emerges from this example. The initial equilibrium was disturbed by a fall in autonomous consumption as households sought to increase their savings. However, the intention to save more did not lead to higher aggregate savings because *income fell* ($\Delta y = k(c_0 - c_0')$). This must be the case for Equation 1.9 to hold.[4] This is called the *paradox of thrift*, because if greater thriftiness is not matched by higher investment in fixed capital, income will fall and there will be no overall increase in savings in the economy.

To summarize, the answer to the question of whether saving should be encouraged in a recession depends on the model of the economy the economist is using–and of course, on how well that model matches the real economy under study. Using the model we have developed so far in this chapter, the answer is clear: encouragement to save more will not help the economy to exit recession. The reason is that there is no mechanism in *this* model through which higher saving is translated into higher investment. Investment remains at $\bar{I}$ throughout. Hence the result of a higher propensity to save is that aggregate demand falls, output falls and the recession is deepened.

By contrast, if the model includes a central bank, then the recession could be averted by the central bank cutting the interest rate and boosting investment to offset the fall in savings. In the next section, we remove the simplifying assumption of exogenous investment to show the impact on the demand side of investment being responsive to changes in the interest rate.

1.2.3 The *IS* curve

Deriving the IS equation

Throughout this book, we use the *IS* curve to represent the demand side of the economy. In this subsection, we derive the *IS* curve and use it as a starting point to think about how monetary and fiscal policy work. As we discussed in Section 1.1.2, the *IS* curve shows the combinations of output and the real interest rate at which the goods market is in equilibrium.

We first set out the relationship between the real interest rate, r, and the nominal interest rate, i. This relationship is shown by the Fisher equation:

$$r = i - \pi^E. \quad \text{(Fisher equation)}$$

[4] We can confirm that the change in aggregate saving is equal to zero. Using Equation 1.7:

$$\Delta y = \frac{1}{1-c_1(1-t)}\Delta c_0,$$

$$\Delta S = \left[(1 - c_1(1-t))\frac{1}{1-c_1(1-t)}\Delta c_0\right] - \Delta c_0 = 0.$$

This equation says that the real interest rate is simply the nominal interest adjusted for expected inflation. It is the real interest rate that is most important for investment and saving decisions, as it represents the true cost of borrowing (and the true return on saving). This interest rate is therefore the one used in the investment equation and in the *IS* curve. When the central bank sets the nominal interest rate, it does this with the intention of achieving a particular real interest rate since it aims to affect interest-sensitive spending. In the Appendix, we show how the Fisher equation is derived.

At this stage, we assume there is just one interest rate in the economy that applies to all borrowing and saving. We make the assumption to keep the maths as simple as possible. In reality, there is a spectrum of interest rates. For example, the interest rate on bank lending will typically be higher than the interest rate set by the central bank. We discuss the banking mark-up in Chapter 5 and the difference between short-term and long-term interest rates in Chapter 7.

Although we shall amend this later in the chapter, in the model to this point, consumption is independent of interest rates. This means that the transmission of monetary policy will work through its effect on investment spending, including on new houses. Up to this point in the chapter, we have assumed that investment is determined by expected future profits, which we have assumed to be exogenous. We now incorporate the interest rate into the investment function as follows:

$$I = a_0 - a_1 r, \quad \text{(investment function)}$$

where r is the real interest rate, a_0 and a_1 are constants and $a_1 > 0$. The main determinant of investment is expected future post-tax profits, which is captured by the term, a_0.

We derive the *IS* curve as follows. Substituting the investment function into Equation 1.4 gives us a relationship between the real interest rate and output, which is the *IS* curve. We can then use k to denote the multiplier and simplify to achieve a simple equation for the *IS* curve:

$$y = \underbrace{\frac{1}{1 - c_1(1-t)}}_{\text{multiplier}}[c_0 + (a_0 - a_1 r) + G] \tag{1.10}$$

$$= k[c_0 + (a_0 - a_1 r) + G] \tag{1.11}$$

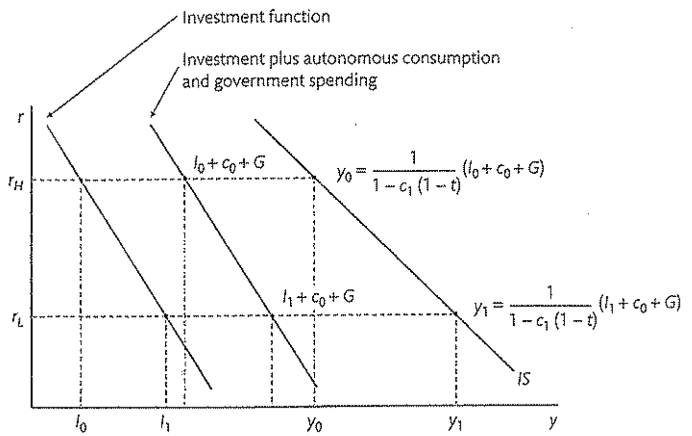
$$= k(c_0 + a_0 + G) - k a_1 r \tag{1.12}$$

$$= A - ar, \quad \text{(IS curve)}$$

where $A \equiv k(c_0 + a_0 + G)$ and $a \equiv k a_1$.

The derivation of the equation for the *IS* curve highlights the fact that given r, equilibrium output, y, is found by multiplying autonomous consumption and investment, and government spending by the multiplier $k = \frac{1}{1-c_1(1-t)}$. This fixes an r-y combination on the *IS* curve. It is clear from Equation 1.12 that a higher multiplier, k, or higher interest-sensitivity of investment to the interest rate, a_1, increases the effect of a change in the interest rate on output, making the *IS* curve flatter.

The *IS* curve is derived graphically in Fig. 1.6 using the *IS* curve equation shown above. Three curves are shown on the figure: the investment function, the investment function plus autonomous consumption and government spending and the *IS* curve. Two real interest



Note that: $I_0 = a_0 - a_1 r_H$ and $I_1 = a_0 - a_1 r_L$

Figure 1.6 Deriving the IS curve.

rates are shown in the figure: a high real interest rate, r_H, and a low real interest rate, r_L. At r_H, investment is more costly, so the level of planned investment is low, whereas the opposite is true at r_L. This leads to a downward-sloping investment function. Autonomous consumption and government spending are unaffected by the interest rate, so when added on, the curve shifts out parallel. The last step to derive the *IS* curve is to multiply $I_0 + c_0 + G$ by the multiplier. This gives output equal to planned expenditure, y, on the *IS* curve. The *IS* curve shows all the combinations of output and real interest rate where the goods market is in equilibrium. It is flatter than the other two curves because the multiplier effect of investment is higher when interest rates are low.

Using the *IS* curve equation and the diagram, we can summarize its properties.

1. The *IS* curve is downward sloping because a low interest rate generates high investment, which will be associated with high output. By contrast, when the interest rate is high, investment and consequently equilibrium output is low.
2. The slope of the *IS*:
   (a) Any change in the size of the *multiplier* will change the slope of the *IS* curve. For example, a rise in the marginal propensity to consume, c_1, or a fall in the tax rate, t, will increase the multiplier, making the *IS* flatter: it rotates counter clockwise from the intercept on the vertical axis.
   (b) Any change in the *interest sensitivity of investment* (a_1) will lead to a consequential change in the slope of the *IS* curve: a more interest-sensitive investment function (i.e. $\uparrow a_1$) will be reflected in a flatter *IS* curve.
3. Shifts in the *IS* curve: any change in *autonomous consumption, autonomous investment* or *government expenditure* (c_0, a_0, G) will cause the *IS* curve to shift by the change in autonomous spending times the multiplier.

1.2.4 Forward-looking behaviour

The spending decisions of agents in the present are influenced by their expectations of the future. This means there is an intertemporal component to both consumption and investment.

1. Households adjust their current spending based on their expected income in the future. For example, a final year economics undergraduate who secures a lucrative job contract (which is set to commence after university), has low current income, but high expected future income. Having got the job, this individual may borrow in the present, consume more and pay back the money when they start their job. This behaviour is referred to as *consumption smoothing*.

2. Firms make decisions (e.g. purchasing machinery, equipment and premises) based on a business plan that includes forecasts about future demand for their products and input costs. For example, a firm selling cars to China might choose to build a new factory (i.e. undertake investment) based on forecasts that Chinese incomes (and therefore their demand for cars) will continue to rise rapidly over the next 20 years. Investment is intrinsically forward looking, as it incurs a cost today, but the stream of benefits occurs in the future.

Present value calculation

We can calculate the present value of a flow of income or profits received in future periods in the following way. We take the example of a firm making an investment decision.

We assume that firms aim to maximize profits, so they undertake investment projects if these offer a return that is higher than costs. The outlay on investment typically precedes the returns, which may be lumpy and spread over a number of years. The way to deal with this is to calculate the 'Present Value' (V) of the expected flow of profits Π, where Π_t, Π_{t+1}, ... is the flow of profits in period t, $t+1$ etc. Suppose the interest rate is 10%. Then if I save €100 today I will have €110 in a year's time (100 (1 + 10%)). Expressed the other way around, we could say that €110 in a year's time has the same value as €100 today—its present value is €100.

More generally, if interest rates are constant at r, the value of X in n years' time is the same as that of $X/(1+r)^n$ today. We can calculate the present value of our stream of expected profits, Π^E, from an investment project with profits over T years by:

$$V_t^E = \Pi_t^E + \frac{\Pi_{t+1}^E}{(1+r)} + \frac{\Pi_{t+2}^E}{(1+r)^2} + \cdots = \sum_{i=0}^{T} \frac{1}{(1+r)^i} \Pi_{t+i}^E.$$

(expected present value calculation)

We use the superscript E to denote the expected value of a variable.

If the cost of the machine is greater than the present value of the flow of profits from the machine, then it would be more profitable not to buy the machine but instead put the money in the bank or in bonds (which earn interest of r). Similarly, if the money to purchase the machine is being borrowed, then if the cost of the machine is greater than the present value, buying the machine will be unprofitable. On the other hand, if the present value is greater than the cost, then this investment is profitable, and a profit-maximizing firm will go ahead with it.

We can also apply the same logic to modelling consumption decisions. We can account for the fact that the future will affect consumption decisions by calculating the present value of the stream of expected income over a person's lifetime. If we assume that individuals live forever, we can use the formula for calculating present value to calculate the expected present value of lifetime wealth (Ψ^E, pronounced 'sigh'), which is defined at time t as follows,

$$\Psi_t^E = (1+r)A_{t-1} + \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} y_{t+i}^E,$$

(expected present value of lifetime wealth)

where $\sum_{i=0}^{\infty} \frac{1}{(1+r)^i} y_{t+i}^E$ is the present value of expected post-tax lifetime labour income and $(1+r)A_{t-1}$ are the resources available in period t from the assets the individual held at the end of the previous period.[5]

1.2.5 Consumption

People's income fluctuates over their life cycle; it also fluctuates when they lose their job, move to a different job or get a promotion. Given that people prefer to smooth out the fluctuations in their income in their spending behaviour, they must take account of the future and they must be able to save and borrow. The modelling of consumption should include how households form expectations about the future and how they borrow and save.

The desire to smooth consumption in the face of fluctuations in income is captured by the assumption of diminishing marginal utility of consumption. A simple way of visualizing this is to consider the choice of consumption in a world of just two periods. If we know income will be higher next period, how does that affect consumption now? Consider the choice between either low consumption equal to income in the first period and high consumption equal to income in the second period, or having the average of the two consumption levels in each period. If there is diminishing marginal utility of consumption, more consumption always increases utility, but successive increases in consumption deliver smaller and smaller benefits. Therefore, households will make the second choice, because consuming the average in both periods offers higher utility than the first choice.

The permanent income hypothesis (PIH)

The permanent income hypothesis (PIH) states that individuals optimally choose how much to consume by allocating their resources across their lifetimes.[6] Their resources include their

[5] The assumption that individuals 'live forever' can be thought of as assuming that households have children or heirs and that they incorporate the utility of their children or heirs into their consumption function. In other words, households behave 'as if' they last forever. The assumption is clearly unrealistic, but it greatly simplifies the mathematical derivation of the permanent income hypothesis (PIH). From this point onward, however, we will refer to lifetime income and earnings in the text, as it is more intuitive to think about forward-looking behaviour in terms of a household maximising utility over their lifetime.

[6] This view of the consumption function was first developed by Modigliani and Brumberg (1954). See also Friedman (1957). For a review of Friedman's theory, its influence on modern economics and the relevant empirical literature, see Meghir (2004).

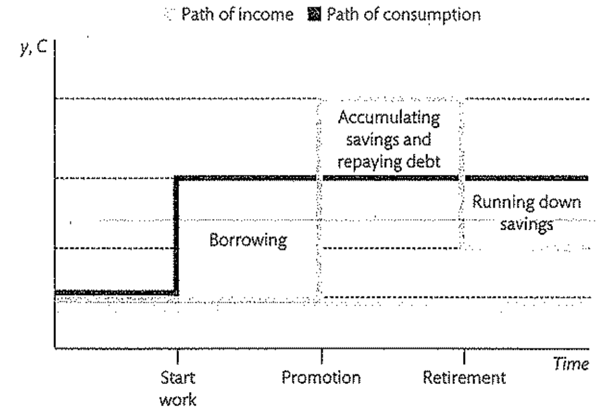


Figure 1.7 The permanent income hypothesis over the life cycle.

assets and their current and future income. This is a forward looking decision and will depend on interest rates, asset values, expectations of future income and expectations of future taxes. The PIH predicts that optimal consumption is smooth as compared to income. For example, when individuals start work, their income is low and they will borrow to consume more; when income increases, they keep consumption constant and use the excess income to pay off debts and save for retirement; then at retirement, their income falls and they draw down their savings. Figure 1.7 shows how consumption and income change over the life cycle in this simplified example. The important point is that the PIH predicts that households will borrow and save in order to smooth consumption over their lifetimes. Likewise, over the business cycle, if an individual becomes unemployed, the model predicts that they will borrow in order to sustain consumption during the spell of unemployment. As we shall see in Chapter 14 on fiscal policy, the government plays an important role in smoothing consumption through the provision of unemployment benefits.

The PIH model of consumption provides a stark contrast with the Keynesian consumption function where there is no explicit consideration of the future. Aggregate consumption of households is modelled there as the consumption of a fixed amount, c_0 and a fixed proportion of the current period's disposable income.

The PIH model is derived in more detail in Section 1.4.2 of the Appendix, but it is useful to set out the intuition and predictions of the model here, as it provides a framework for thinking about forward-looking consumption decisions. Given that income will fluctuate over a person's lifetime, the starting point of the PIH is their desire to smooth out fluctuations in consumption and their ability to save and borrow in order to achieve this.

The next question that arises is whether an individual prefers their smooth consumption path to be one of constant consumption in each period or of rising or falling consumption. This will depend on the relationship between the interest rate on saving and borrowing and the rate at which the individual trades off consumption in the future for consumption in the present. The latter is the subjective discount rate ρ, called rho which is a measure of an individual's impatience.

The household chooses a path of consumption to maximize its lifetime utility subject to its lifetime budget constraint, which is

$$\Psi_t^E = (1+r)A_{t-1} + \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} y_{t+i}^E. \qquad \text{(lifetime budget constraint)}$$

Note that when considering the PIH, income, y is defined as post-tax income. The key to solving this problem can be thought about over just two periods, since the same considerations will apply to every subsequent pair of periods.[7] The answer is a simple relationship between consumption this period and next period, which is called an Euler equation (the derivation is in Section 1.4.2 of the Appendix):

$$C_t = \frac{1+\rho}{1+r} C_{t+1}^E. \qquad \text{(Euler equation)}$$

The intuition is clearest if we take the case where the interest rate, r, and the subjective discount rate, ρ, are the same. This means the household gets the same (objective) return from saving, r, as their (subjective) willingness to trade off consumption in the present for consumption in the future, ρ. In this case, the agent prefers a constant level of consumption each period, $C_t = C_{t+1}^E$.

We can see that if the discount rate is above the interest rate, $C_t > C_{t+1}^E$ then consumption is falling over time, reflecting the 'impatience' of the household. The opposite is the case if the discount rate is below the interest rate; a patient household will have a path in which consumption rises over time, i.e. $C_t < C_{t+1}^E$. The crucial point to note is that whichever of these consumption patterns is chosen, it is independent of the period by period changes in income.

From now on we focus on the PIH where the subjective discount rate is equal to the real rate of interest, i.e. ($\rho = r$). To implement a life-time consumption plan like this, given each period's income the household must do whatever saving and borrowing is called for in order to deliver the constant level of consumption each period. If current income is above permanent income, individuals save and if it is below, they borrow.

The next step is to use the Euler equation to find out how much individuals consume each period (the derivation is in Section 1.4.2 of the Appendix). When $\rho = r$, consumption is

$$C_t = \frac{r}{1+r} \Psi_t^E. \qquad \text{(PIH consumption function)}$$

The intuition is easy to understand: an individual with this consumption function will borrow and save to deliver a perfectly smooth consumption path (in expectation). The amount they consume each period is equal to the annuity value of expected lifetime wealth and is called 'permanent income'. The individual consumes their permanent income and the formula ensures that in expectation, they will be able to do this forever. Note that consumption remains constant unless Ψ_t^E changes.

7 See Hall (1978) for a full derivation of the PIH and an empirical test of the main implications of the theory.

Predictions and empirical evidence on the PIH

The predictions of the PIH can be tested. How does consumption react when a change in income occurs? The answer depends on the nature of the change in income.

1. *Anticipated* or foreseen changes in income should have no effect on consumption when they occur. The reason is that anticipated changes will already have been incorporated into consumption through the recalculation of permanent income. Hence, when the change in current income is recorded, the marginal propensity to consume is predicted to be zero (the multiplier is equal to one). A finding of '*excess sensitivity*' to anticipated income changes would contradict the full smoothing behaviour predicted by the PIH.
2. *News or unanticipated* changes in income should affect consumption because they require the recalculation of future lifetime wealth, Ψ_t^E.
   (a) News of a temporary increase in income. If current income y_t increases unexpectedly by one unit, consumption increases by the extent to which this raises permanent income. Since the increase in one unit will be spread over the entire future, the PIH consumption function tells us that permanent income and hence consumption rise very little, just by $\frac{r}{1+r}$ times the increase in lifetime wealth. The marginal propensity to consume out of temporary income is $\frac{r}{1+r}$. (For example, if the real interest rate is 4%, the MPC is 3.8%, which implies a multiplier barely above one.)
   (b) News of a permanent increase in income. If there is news that income y_t is higher from now and for every future period by one unit, then permanent income and hence consumption rise by the full one unit. This means the marginal propensity to consume out of post-tax permanent income is one. A finding of '*excess smoothness*' of consumption in response to news of permanent income changes would contradict the simple PIH.

Excess sensitivity to anticipated *changes in income*

The first testable hypothesis suggests that there should be no change in consumption at the time income changes, if the change in income was known in advance. This is because consumption should already have adjusted as soon as the news arrived of the change in future income. An influential study by Campbell and Mankiw (1989) tested this hypothesis econometrically using aggregate data on consumption and income from the G7 countries. The study rejected a model in which all consumers were following the PIH but could not reject a model in which half of all consumers were simply following the 'rule of thumb' of spending their current income. In other words, they found that current consumption was overly sensitive to expected changes in income across the G7.

A recent study used the 2001 federal income tax rebates in the US as a testing ground: according to the PIH, the one-off (temporary) tax rebate should have very little effect on spending and if there is any effect, it should occur when the announcement was made and not when the cheques arrived. Johnson, Parker and Souleles (2006) were able to identify the causal effect of the tax rebate by using household data and the fact that the timing of the sending of cheques was based on the taxpayer's social security number, which is random. They found that the average household spent 20–40% of the (predictable) tax rebate in the three month period when the cheque was received rather than when the programme was announced. They also found that it was households with low income and wealth that responded most, underlining the likely role of limits on the ability of households to borrow, i.e. credit constraints. This is referred to as the *excess sensitivity* of consumption and is evidence against the strong predictions of the PIH.[8]

Is consumption excessively sensitive to the predictable fall in income at retirement? What is observed is that when income falls in a predictable way on retirement, consumption falls. Although this is a violation of the simple PIH, explanations consistent with consumption smoothing in a more detailed model have been proposed. These include the possibility that consumption falls because spending was related to being at work (i.e. complementary to working, such as the purchase of business suits) and that in the light of the increased leisure of the household, home production substitutes for consumption spending (e.g. cooking replaces the purchase of more expensive ready-made meals). Jappelli and Pistaferri (2010) conclude that evidence on the so-called consumption–retirement puzzle is not clearly in contradiction with the spirit of the PIH. Putting this together with the above discussion, the fact that there is excess sensitivity to anticipated rises in income but not so much to falls in income underlines the important role of credit constraints, which are ignored in the PIH.

Excess sensitivity to news about temporary income and excess smoothness to news about permanent income

Turning to the hypotheses that relate to news about income, studies have found that consumption over-responds to temporary income shocks. In the case of positive income shocks, this strongly violates the PIH. An early example of this was the large consumption response of US veterans after the Second World War to an unexpected windfall payout of the National Service Life Insurance. This kind of evidence questions the assumption in the PIH that households make optimizing decisions over a time horizon stretching a long way into the future and use a low discount rate (equal to the interest rate) to compare consumption at near and far horizons. The fact that people respond to a windfall by raising spending suggests that discount rates are higher than assumed in the simple PIH: people appear to be more *impatient* than the hypothesis assumes.

It is also likely that uncertainty about whether observed income changes are temporary or permanent prevents households from acting exactly as PIH would predict. For example, if a household mistakenly thought a temporary change in their current income was permanent, then they would consume more of the income change than would be consistent with PIH behaviour.

Jappelli and Pistaferri (2010) review the recent evidence and suggest that it shows that consistent with the PIH, consumption responds more to permanent than transitory income shocks, but that households do not revise their consumption fully into line with permanent shocks.

[8] For a detailed summary of the early literature on excess sensitivity see Deaton (1992). A recent excellent overview of the empirical testing of the excess sensitivity and excess smoothness hypotheses is Jappelli and Pistaferri (2010).

Credit constraints, impatience and uncertainty

The evidence suggests there are three reasons for the failure of the simple PIH to provide an adequate model of aggregate consumption:

1. The presence of credit constraints, which prevent borrowing by households who lack wealth or collateral.
2. Impatience, which prevents some households from saving as would be indicated by a permanent income view.
3. Uncertainty about future income, which explains precautionary saving over and above the level predicted by the PIH.

Credit constraints

If people prefer to smooth their consumption, but are prevented from doing so because they cannot borrow to bring forward consumption when their current income is below their expected permanent income, they are said to face credit constraints. The prevalence of credit constraints is an important factor that stops the world from operating as the simple PIH theory would imply. Because of information problems facing banks in assessing creditworthiness, banks are not always willing to lend to households without the wealth or collateral to secure a loan in order that they can smooth their consumption (see Chapter 5). Figure 1.8 shows the consumption response for PIH and credit-constrained households to an anticipated rise in income. The PIH households borrow to raise consumption as soon as the news of the future income increase is received, whereas the credit-constrained households do not have

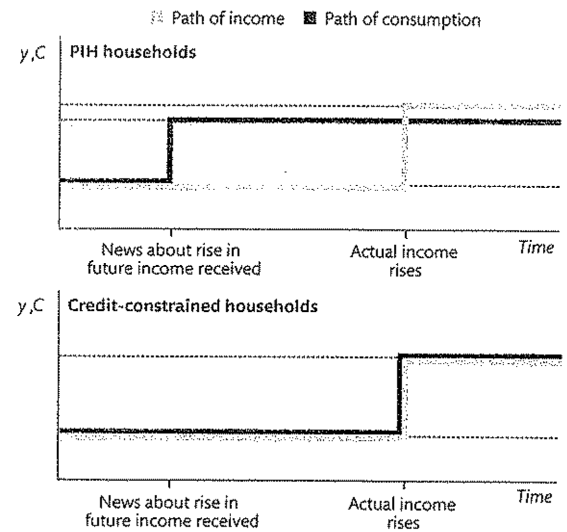


Figure 1.8 The consumption response of PIH and credit-constrained households to an anticipated rise in income.

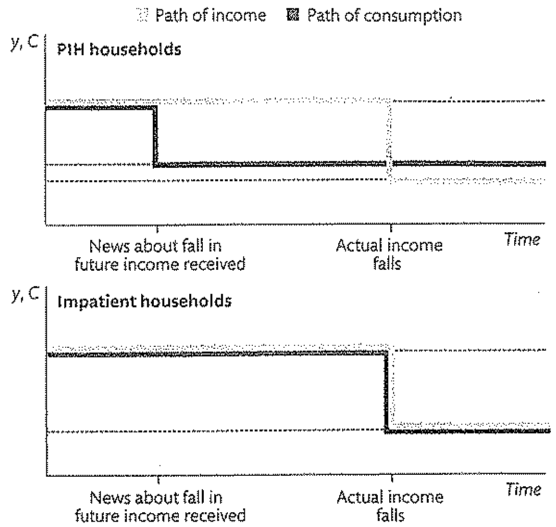


Figure 1.9 The consumption response of PIH and impatient households to an anticipated fall in income.

this option, so consumption can only rises when income does. Estimates from studies using household data have found that around 20% of the population in the United States is credit constrained.[9] A much higher estimate of closer to two-thirds comes from a study of the way people respond to automatic increases in the limits on their credit cards: if spending responds to such automatic increases, it suggests the card-owner was credit constrained.[10]

Impatience

The observation that consumption of a substantial fraction of households changes one-for-one with their current income (i.e. no borrowing or saving) reflects two features of the real world. The most important is the presence of credit constraints, which can prevent households from borrowing to smooth their consumption. This seems to especially affect young and low income households. Faced with a spell of unemployment, for example, constraints on borrowing will prevent a household from maintaining its level of consumption.

Whereas credit constraints can explain the inability of households to borrow to smooth consumption, it cannot explain the failure of some households to save, which would allow for some consumption smoothing without relying on bank loans. It seems that there are households who are not only credit constrained, but also impatient. Figure 1.9 shows the consumption response for PIH and impatient households to an anticipated fall in income. The PIH households start saving as soon as the news of the fall in future income is received, which alllows them to smooth consumption over the two periods. In contrast, the impatient households fail to reduce current consumption upon receipt of the news, meaning consumption falls dramatically when income actually falls.

[9] See Japelli (1990), Mariger (1986) and Hall and Mishkin (1982).
[10] See Gross and Souleles (2002).

Experimental evidence on impatience

The impatience of households is represented in a model of consumption by a higher discount rate for the short run than the longer run. A simple illustration of what this means is to consider the following experiment: if you were choosing today for eating *next week*, would you choose fruit or chocolate? About three-quarters of experimental subjects choose fruit. When asked whether they would choose fruit or chocolate to eat *today*, 70% choose chocolate. From this, we can see that preferences are not consistent over time, because if we run time forward to next week, the three quarters who said they would prefer fruit would have shrunk to 30%. When next week comes, they would in fact, choose chocolate.[11]

Impatience is also documented in experiments related to consumption and saving behaviour. For example, subjects have a budget and can choose between two accounts paying the same interest rate: a 'freedom account' with no constraints on withdrawal or a 'commitment account' where there are restrictions on withdrawal before a date chosen by the participant. From a purely economic point of view, the freedom account is preferable because it imposes no constraints on withdrawal. In spite of this, more than half of investment was put in the commitment account (no withdrawal before the goal date), which reveals the desire of the experimental subjects to commit themselves to saving.[12]

Implications of impatience for consumption and saving

Very large differences between a high short-run and a low long-run discount rate help explain why people simultaneously save in illiquid assets such as housing (where the interest rate is low) and borrow on credit cards, paying a very high interest rate. Households with behaviour of this kind will have a marginal propensity to consume of one. Although a person can agree that their expected well-being is higher if they save more now, they are unable to resist consuming and blowing a hole in their saving plan when they get an unexpected increase in their income. One way of committing to saving is to take out a mortgage. Another is for the government to automatically enrol people in schemes for saving for retirement (from which they can opt out) such as pensions.

Uncertainty and precautionary saving

If there is uncertainty about future job opportunities or about health, then the household may save as insurance for future contingencies. For example, having a 'buffer stock of savings' could be important to sustain consumption in the event of job loss for a credit-constrained household.[13]

In the face of uncertainty, households would tend to save more early in life than the PIH predicts. Instead of the average propensity to consume falling as income rises over the working life as in the simple PIH, the desire for precautionary savings leads to saving early on with result that consumption rises with income (not by more than income) early in the

[11] See Read and van Leeuwen (1998). [12] See Ashraf et al. (2006).
[13] See Carroll (1997) for an exposition of the buffer-stock savings model.

life cycle and the average propensity to consume rises later on. If saving for a rainy day is important, then the utility provided by having assets to tide the household over if need be outweighs the utility that would have come from higher consumption in the early years.

An empirical consumption function

To conclude the discussion of consumption, we consider a consumption function estimated on macroeconomic data for the UK, US and Japan.[14] Although the empirical model does not impose the straitjacket of the simple PIH, it nevertheless allows for permanent income considerations to affect consumption; households will have a notion of sustainable consumption constrained by their budget, which includes their income forecasts. They will therefore know that they cannot spend their existing assets now without impairing future consumption and that future income affects their sustainable consumption.

In addition to permanent income considerations based on forecasts, the consumption function that is estimated includes the following factors.

1. Consumption is predicted to respond more strongly to current income than the PIH predicts due to the presence of credit constraints.

2. The effects of income uncertainty on the desire of households to have precautionary savings. Income uncertainty is measured by the change in the unemployment rate and an increase is predicted to reduce consumption.

3. The housing collateral effect through higher house prices is expected to interact with increased credit availability to increase consumption (in the presence of credit constraints). The presence of this effect will depend on the institutional arrangements for mortgage borrowing in the country in question. For example, the availability or otherwise of home equity loans is a crucial distinction. A home equity loan is a loan secured against the equity in an individual's home. Home equity is the difference between the current market value of the house and the remaining mortgage, hence it increases as house prices rise. The collateral channel works through credit constraints being relaxed, but if it is not possible to get a home equity loan, as in France, Germany, Japan and China, this channel does not operate. In fact, if home equity loans are unavailable and high down-payments are required to get a mortgage (as in Japan), then higher house prices can actually dampen consumer spending, as young people have to save more to accumulate the required down-payment.

4. The current real interest rate. There are many interest rate channels, some of which point to a negative and some to a positive effect of a higher real interest rate on consumption.

5. Increases in households' wealth to income ratio are expected to increase consumption.

The consumption function includes the housing and financial variables that are at the centre of the discussion in Chapters 5–7, where the banking system, financial cycles and financial crises are brought into the macro model.

[14] See Aron et al. (2012).

For the UK and US, the consumption functions estimated using macroeconomic data suggest that roughly 50% of household consumption is accounted for by permanent income considerations. Current income has an important positive effect on consumption and the real interest rate has a negative effect. The wealth to income ratio has a positive effect but different types of wealth have different effects. For example, cash in a bank or savings account is more spendable than pension wealth. Also, the evidence is that for a given degree of access to new credit, higher debt of \$100 reduces consumer spending more than higher pension or stock market wealth of \$100 increases spending.

For these two economies, where home equity loans are important, increased credit availability and house prices interact to raise consumption through the collateral effect. A house price boom in conditions where access to credit is easier boosts consumption. We return to look at the implications of the richer analysis of consumption for the *IS* curve in Section 1.2.7.

1.2.6 Investment

The decision a firm makes over whether to invest or not will depend on their expectations about future after-tax profits. The role of expectations helps explain the excess volatility of investment over the other components of GDP visible in Fig. 1.1. In developing the *IS* equation, we used a simple linear relation to model investment in Section 1.2, where the amount of investment depended positively on expected future profits (captured by the constant term, a_0) and negatively on the real interest rate (a measure of firms' borrowing costs). In this subsection we introduce a more sophisticated forward-looking modelling approach—the q theory of investment.

Tobin's q theory of investment

This theory was developed by Nobel Prize winner James Tobin.[15] Like the PIH, Tobin's q theory of investment is forward looking. Firms choose the amount of investment to undertake with a view to maximising the expected discounted profits over the lifetime of the project. The q theory amounts to comparing the benefits from investment in an increase in the capital stock with the costs of doing so: if the expected benefits exceed the costs, investment should take place.

Marginal and average q models

The theory is derived explicitly in the Web Appendix to this chapter, but it is useful to set out the equation for *marginal* q here, as it provides a framework for thinking about forward-looking investment decisions:

$$q = \frac{\text{MB of investment}}{\text{MC of investment}} = \frac{Pf_K}{\delta + r}. \qquad \text{(marginal } q \text{ model of investment)}$$

The model is intrinsically forward looking since the firm is considering the future benefits from investment spending now. Before looking at the expression on the right, we note that the optimal amount of investment occurs when $q = 1$. This is where the expected marginal

[15] See Tobin (1969) and Tobin and Brainard (1977).

benefits of investment are equal to the expected marginal cost. Thus, if $q > 1$, the marginal benefit of investment exceeds the marginal cost, so firms should invest to increase the capital stock until $q = 1$. If $q < 1$ the firms should reduce their capital stock.

It is important to note that the q theory of investment is a model in which firms are assumed to make their investment decisions (a flow) by pinning down the optimal capital stock (a stock). Small changes in the desired capital stock can therefore translate into large changes in the flow of investment. This is one reason for the excess sensitivity of investment over GDP (see Fig. 1.1).

The marginal benefits from investment are higher when the price of output is higher and when the investment is more productive. The equation shows that firms should undertake more investment if there is

1. an increase in the price of output, P,
2. an increase in the marginal productivity of capital, f_K, which indicates the increase in output the new capital equipment will produce,
3. a reduction in the rate of interest, r, or
4. a reduction in the rate of depreciation, δ. For example, a rise in the expected rate of depreciation (e.g. as a result of uncertainty about future legislation banning fuel-inefficient cars) would lead to a reduction in the level of current investment because it would reduce the expected benefits to the firm from additional investment in the auto industry.

Is q-theory a good way to represent real-world investment behaviour? Marginal q implies firms adjust their level of investment in each period to equate the marginal benefits and marginal costs of investment. This does not fit with the real-world data on investment. When economists looked at very detailed microeconomic data for very large numbers of plants, they found that, contrary to the idea of a smooth investment flow, investment was in fact very bunched.[16] For example, a growing manufacturing firm might build a new factory once every 10 years.

The share price as indicator of expected profits

Testing q theory directly is difficult because q itself is difficult to measure. This is because it depends on the marginal product of capital and a measure of technology, which cannot be observed. In the real world, firms do not know their production functions and do not think explicitly in terms of marginal products. To operationalize the theory, the market value of the firm as reflected in its stock market valuation is compared with the replacement cost of the capital stock. If the market value is higher then this signals that the firm should increase investment. On the other hand, if the market value is below the replacement cost, then the firm would not want to build a *new* factory because it could buy an existing one more cheaply.

We can define average Q as follows:

$$Q = \frac{\text{Market value of firm}}{\text{Replacement cost of capital}}. \qquad \text{(average } Q \text{ model of investment)}$$

16 See the survey chapter on 'Investment' by Caballero (1999).

This is often called *average* Q to distinguish it from marginal q. Whereas q is the ratio of the marginal benefit of a unit of investment to its marginal cost, Q depends on the total expected return from the firm's capital divided by the total cost. For publicly listed companies, the stock market provides a forward-looking measure of the market value of the firm, which can be used in the numerator of the Q equation: when the firm's market value rises relative to the replacement cost of the firm, as reflected in a rise in the price of its shares, the model suggests that investment should go up. A higher interest rate and depreciation rate will raise the replacement cost of capital.

The idea is that the market value incorporates information about how well the firm is expected to be able to implement the investment, whether new competitors will enter the market, whether there are new technological innovations likely to affect the firm's value, the state of the macro-economy and the labour market and the future path of the interest rate. Since investing in a firm is a wager on this uncertain future, investors continuously evaluate these factors and, under certain conditions, the share price and hence the market value of the firm will reflect all the information available.

The stock market is notoriously volatile. For example, in 2008, the UK's leading share index—the FTSE 100—fell 31% and GDP contracted by just 1.1%. The following year, the FTSE 100 rose 22%, whilst GDP continued to contract, but this time by 4.4%.[17] In times of extreme uncertainty, the share price may not reflect the fundamental value of a firm. Large movements in share prices heavily affect market capitalization, which is the measure used to proxy for the market value of the firm in the calculation of average Q. As the stock market is volatile and subject to bubbles, fads and herd behaviour, it may not be a good indicator of the firm's prospects.

However, a study using US micro data (over 1,000 firms from 1982 to 1999) shows that a Q model of investment can be successfully estimated when analysts' forecasts of a firm's profits are used instead of the stock market price to measure the numerator in the Q equation.[18]

Empirical evidence on investment

The empirical evidence on the q theory is that Q helps to explain investment but it is not the only influence. In particular, just as was the case for consumption, credit constraints play an important role in explaining investment spending.

Credit constraints—the role of cash flow

A testable prediction of q theory is that current cash flow should have no impact on investment. The reason is that forward-looking firms should take into account any credit constraints that they face: these should already be incorporated into the stock market valuation, Q. However, the role of cash flow in empirical studies of investment suggests that *capital market imperfections* are important. The importance of cash flow variables in estimated investment equations is reminiscent of the importance of current income in estimated

17 GDP data from IMF World Economic Outlook database, April 2012. Stock market data from Bloomberg News articles.
18 See Bond and Cummins (2001).

consumption functions. As we have seen, the presence of credit-constrained households causes excess sensitivity of consumption to predictable changes in income. Similarly, the presence of cash flow terms in estimated investment functions strongly suggests that firms are credit constrained. Although some firms are able to borrow as much or sell as much equity as they would like in order to finance their investment plans, for others, their investment will be limited by their internal funds.

This is sometimes referred to as the *excess sensitivity of investment to internal funds*. For firms like this, cash flow would be expected to be an important determinant of investment.[19] The reason why firms may be rationed in their access to bank loans is explored in Chapter 5.

A study using UK company data for the period 1975–86 found that company investment was significantly influenced by Q and by credit constraints. However, the impact of Q was found to be very small: a 10% rise in the stock market value of a company was associated with an immediate rise in the investment rate of only 2.5%. Cash flow, on the other hand, was very important. The sample period was divided into two, and it was found that the impact of Q was lower and the impact of cash flow higher in the first part of the sample, during which the UK economy was in a deep recession. This is consistent with the idea of credit constraints biting especially hard in a recession.

Uncertainty—the option value of waiting

The decision of firms about whether to invest or not is also influenced by uncertainty about the future. Under certain circumstances, there can be a value in waiting to undertake an investment project, since with the passing of time more information arrives, the costs of delay (lost profits from the project) may be outweighed by more secure information on the balance between the costs (including those sunk in the project) and the benefits of undertaking it. The upshot of such considerations is that an expected rate of return considerably higher than the cost of capital will be required to trigger investment. Dixit (1992) gives examples to show that including the so-called 'option value of waiting' can double the hurdle rate (i.e. the return required to trigger investment) for an investment project to proceed.

1.2.7 Consumption, investment and the *IS* curve

In the previous subsection, we introduced models of consumption and investment that took account of forward-looking behaviour. Expectations of future income and of future profits play a key role in the spending decisions of households and firms. We also reported empirical evidence on aggregate consumption and investment, which highlighted the role of credit constraints for households and firms. What are the implications for the *IS* curve?

The basic form of the *IS* curve was presented in Section 1.2.3 and we shall see that we can bring together the additional insights about consumption and investment by focusing on the factors that affect the slope of the *IS* curve and which shift it.

[19] See, for example, Chirinko (1993).

[20] See Blundell et al. (1992).

The *IS* equation is:

$$y = k(c_0 + a_0 + G) - ka_1 r \quad (1.13)$$
$$= A - ar, \quad \text{(IS curve)}$$

where $A \equiv k(c_0 + a_0 + G)$ and $a \equiv ka_1$; k is the multiplier. A larger multiplier makes the *IS* curve flatter, which increases the impact on output of a change in the interest rate; it also increases the impact on output of a shift in c_0, a_0 or G for a given interest rate. This makes it clear that the size of the multiplier is important for understanding the effectiveness of both fiscal and monetary policy.

As we shall see in Chapter 14 on Fiscal Policy, there has been a great deal of debate about the empirical size of the multiplier in the context of the use of fiscal stimulus to respond to the Great Recession following the global financial crisis. We shall see that that there is no simple answer to the question: what is the size of the multiplier? It depends on characteristics of the economy and on the context in which, for example, a fiscal stimulus is being applied.

We can summarize the implications of the analyses of consumption and investment from Sections 1.2.5 and 1.2.6 by thinking about how they affect the multiplier and the *IS* curve.

Factors affecting the multiplier

The PIH predicts that the size of the multiplier is dependent on the permanency of the income shock. The hypothesis suggests the marginal propensity to consume from temporary unanticipated changes in income is close to the size of the interest rate, i.e. only $\left(\frac{r}{1+r}\right)$. The multiplier is therefore predicted to be very close to one for temporary income shocks. On the other hand, permanent income shocks will result in a multiplier larger than one as households update their consumption to reflect their new, higher permanent income.

The presence of credit-constrained and impatient households also impacts the size of the multiplier. For these households, the multiplier effect will be larger than one even for temporary income shocks, as their marginal propensity to consume out of changes in income is equal to one. The larger is the proportion of these 'hand to mouth' or 'rule of thumb' households in the economy, the larger will be the multiplier.

Uncertainty over the permanency of income shocks is another factor that will influence the multiplier. If a substantial proportion of observed income fluctuations are perceived as permanent, then this will result in a multiplier larger than one. In this case, households could be following the PIH, but uncertainty leads their behaviour to be consistent with a Keynesian consumption function.

A change in the multiplier will shift the *IS* curve and change its slope (a higher multiplier makes the *IS* flatter).

Other factors affecting the slope of the IS curve

In addition to the effect of the size of the multiplier, the slope of the *IS* curve is affected by the interest sensitivity of investment and of consumption. The theoretical prediction of the impact of the interest rate on consumption is ambiguous; a reduction in interest rates boosts consumption through some channels and dampens it through others. The empirical evidence on the consumption function suggests that national institutional structures, particularly in relation to the financial system, are important for determining the strength and direction of the relationship. Aron et al. (2012) find a negative relationship between consumption and the real interest rate in the financially liberalized UK and US economies, but they find a positive relationship in Japan, where households' huge liquid deposits far outweigh household debt.

Other factors shifting the IS *curve*

In addition to the effect of the size of the multiplier, the *IS* curve is shifted by a number of other factors. Looking first at consumption, the PIH predicts that anything that changes expected lifetime wealth, Ψ_t^E, such as changes in asset prices or news about a future promotion, shifts the *IS*. The empirical findings about consumption highlight three other factors that can shift the *IS* curve:

1. The role of uncertainty: a rise in the rate of unemployment will raise savings for precautionary purposes, shifting the *IS* to the left.
2. A house price boom will boost consumption in a country with home equity loans by loosening credit constraints, shifting the *IS* to the right. However, a house price boom could also shift the *IS* to the left in countries where home equity loans are unobtainable and large down-payments are needed to get a mortgage.
3. A shift in credit market architecture that increases household access to credit, such as financial innovation or deregulation, will shift the *IS* to the right (at least until the accumulation of debt eventualy cancels out some of the shift).

Turning our attention to investment, Tobin's marginal q predicts that the following factors will shift the *IS* curve to the right:

1. an increase in the price of output, P,
2. an increase in the marginal productivity of capital, f_K; and
3. a reduction in the rate of depreciation, δ.

Lastly, the average Q equation highlights the role of expectations of future profits as a shift factor for the *IS* curve: a rise in the stock market tends to boost fixed investment as it signals a rise in the value of companies relative to their replacement cost.

1.3 Conclusions

This chapter has provided the first building block in the model of the macro-economy. We have set out the *IS* curve, which is used to model the demand side of the economy and is one of the three equations that underpin the 3-equation model that will form the core model of this textbook.

The *IS* curve allows us to think systematically about how changes in the spending behaviour of firms, households and governments can influence output and drive business cycles. The *IS* curve shows the combinations of the real interest rate and output where the goods market is in equilibrium. It slopes downwards to account for the fact that households' consumption and new housing decisions respond negatively to the interest rate; firms will also undertake fewer investment projects as the cost of borrowing increases.

We return to the questions at the start of the modelling section (1.2) and see how the model we have developed helps us work towards an answer. This acts as an exercise to summarize what we have learnt in this chapter.

1. If you give an individual a \$100 bonus, how much of it will they spend? This depends on the model of household spending we are using. If consumer behaviour is well modelled by a Keynesian consumption function (i.e. $C_t = c_0 + c_1(1-t)y_t$) then households will spend a fixed proportion of any extra income they receive. In the case of a \$100 bonus, they will spend everything left over once they have paid tax and put some money into savings. However, if consumer behaviour is better modelled by the permanent income hypothesis (PIH), then individuals' current consumption is a function of their expected lifetime wealth (i.e. $C_t = \left(\frac{r}{1+r}\right)\Psi_t^E$). As the bonus is a one-off, it will only increase expected lifetime wealth by a small amount, resulting in a small increase in consumption in this and all future periods. This leads the individual to save the majority of the bonus in the current period. In Section 1.2.7, we showed how the proportions of rule of thumb and forward-looking households in the economy affect the multiplier and the *IS* curve. We also discussed how the extent to which observed changes in income are perceived as permanent can help to reconcile the Keynesian and PIH consumption functions.

2. How much will output increase following a rise in autonomous demand, such as higher government spending or a boost in the confidence of households and firms? The multiplier will determine the extent to which an €x increase in autonomous demand will increase output. In the simple model presented in Section 1.2.1, the multiplier is always greater than one, as both c_1 and t are between zero and one. This means that in the short run, an injection of autonomous demand will always boost output by more than one for one. The extent of the boost of output will depend on the size of the multiplier. The multiplier will be larger if the tax rate (t) is low and households have a high marginal propensity to consume (c_1). In Section 1.2.7, we saw that if all households are described by the PIH, the permanency of the income change is important for determining the size of the multiplier. When income changes are permanent, the multiplier is greater than one as households update their consumption to reflect the increase in their permanent income. When the income change is temporary, the marginal propensity to consume out of current income is essentially zero as it has a negligible effect on permanent income, and the multiplier is therefore equal to one. Temporary and unanticipated income shocks can, however, result in a multiplier greater than one if some proportion of households are credit constrained or consume according to a rule of thumb. We discuss the multiplier further, with a particular focus on government spending, in Chapters 4, 7 and 14.

3. To what extent will a rise in the interest rate curtail investment? In the simple model in Section 1.2.3, where $I = a_0 - a_1 r$, a rise in the interest rate of Δr will reduce investment by $a_1 \Delta r$. In the marginal q theory of investment (with $q = \frac{Pf_K}{\delta + r}$), an increase in the interest rate will increase the marginal cost of investment. If we assume that q is initially equal to one, (i.e. the marginal cost and marginal benefit of carrying out the investment project are equal) then an increase in r will lead firms to reduce investment. The extent to which investment

will fall depends on how quickly the marginal productivity of capital (f_K) rises as investment falls to bring q back toward a value of 1.

This chapter has provided a model of the demand side which can shed light on macroeconomic questions. We have discussed how the models of consumption and investment relate to characteristics of the real world. For example, the greater volatility of investment as compared with consumption observed in the data can be partly explained by the factors that influence investment spending decisions and by consumers borrowing and saving to smooth their consumption over the economic cycle. Households vary—some are impatient and find it difficult to save, whereas others are prudent and save for precautionary reasons. The government contributes to consumption smoothing through provision of unemployment benefits. Credit constraints facing both households and firms help better align the models of consumption and investment with the empirical data.

Although the demand side is a major influence on economic activity, it is only part of the story of how the macro-economy works. To develop our understanding of economic fluctuations and longer-run trends, we need to introduce the supply side. This will provide a framework for thinking about how wages and prices are set and what determines the unemployment rate. The supply side is the subject of the next chapter.

1.4 Appendix

1.4.1 Real and nominal interest rates and the Fisher equation

To clarify why it is the real rather than the nominal interest rate that affects real expenditure decisions in the economy, think about a firm considering an investment project. A higher money or nominal rate of interest will not impose a greater real burden on the firm if it is balanced by correspondingly higher inflation, because the expected profits from the investment project will be higher in money terms and the balance between the real cost and the real return on the project will not have changed.

The real interest rate is defined in terms of *goods* and the nominal interest rate, in terms of *money*. Thinking of a consumer good, the real rate of interest, r, is how much extra in terms of units of this good - namely (1+r) units - would have to be paid in the future in order to borrow one unit of the goods today. The nominal rate of interest is how much extra in euros would have to be paid in the future in order to borrow one euro today. If goods prices remain constant then it is clear that the real and nominal interest rates are the same: if you lent one euro today, you would be able to buy $(1+r)$ goods in the future. In general,

$$1 + r = (1 + i) \cdot \frac{P}{P^E_{t+1}},$$

where it is the expected price level in the future (P^E_{t+1}) that comes into play since at time t, we do not know what the price level will be at $t+1$. If we use the following definition of expected inflation:

$$\pi^E = \frac{P^E_{t+1} - P}{P},$$

then

$$\frac{P}{P^E_{t+1}} = \frac{1}{1 + \pi^E}.$$

By rearranging the above expression, it follows that

$$(1 + r) = \frac{(1 + i)}{(1 + \pi^E)},$$

and therefore that

$$r = \frac{i - \pi^E}{1 + \pi^E}.$$

When expected inflation is low, the denominator of this expression is close to one and we have the standard approximation for the relationship between the real and the nominal rate of interest:

$$i \approx r + \pi^E. \qquad \text{(Fisher equation)}$$

Inflation expectations will drive the divergence between the real and nominal interest rates. It should be noted that only one of these three terms is observable: the nominal interest rate, i. The real interest rate can be estimated from historical data on the nominal interest rate and the rate of inflation: this gives a measure of the so-called *ex post* real rate of interest. Alternatively, an *ex ante* measure can be derived from a model that is able to predict inflation. Finally, if bonds have been issued in the economy that are protected against inflation because the face value is indexed by the rate of inflation, then the yield on such a bond is a real rate of interest and can provide a third measure. But there are only a few countries that have issued index-linked or inflation-proof bonds (UK in 1981, the USA in 1997, France in 1998).

1.4.2 Deriving the Euler equation and the PIH consumption function

Deriving the Euler equation

The first step in deriving the Euler equation is to set out an expression for the present value of a consumer's utility:

$$V^E_t = U(C_t) + \frac{U(C^E_{t+1})}{1 + \rho} + \frac{U(C^E_{t+2})}{(1 + \rho)^2} + \ldots \qquad \text{(expected present value of consumption)}$$

The present value equation uses the expectations operator (i.e. C^E_{t+i} for all $i > 0$) to show that future consumption is uncertain in the present period. In addition, the utility of consumption is discounted to the present period by the consumer's rate of time preference. Future consumption will be worth less in the current period to more impatient consumers.

To derive the Euler equation, we need to assume a specific functional form for the utility function. We choose $U(C_t) = \log C_t$ because it exhibits diminishing marginal returns and is easy to work with. We can take the first derivative of the utility function to find the marginal utility of consumption, which is $U'(C_t) = \frac{1}{C_t}$.

The Euler equation shows the optimal C_t in relation to C^E_{t+1}. If we just think about these two periods, then it is easy to derive the Euler equation. In period t the consumer must weigh up the gain from consuming more in this period, against the discounted loss of consuming less in the next period. The gain from consuming one unit more in this period is simply the marginal utility of consumption $\frac{1}{C_t}$. The subjective present value loss of utility in the next period is equal to $\frac{1}{C^E_{t+1}}\frac{1}{1+\rho}(1+r)$. This is because the consumer's income will have fallen by (1+r) next period, multiplied by the marginal utility of consumption next period multiplied by the impatience factor.

The consumer cannot gain any more utility when the gain from consuming one more unit in the current period is exactly the same as the subjective present value loss of consuming one unit less in the next period. This gives us the Euler equation:

$$\frac{1}{C_t} = \frac{1}{C^E_{t+1}}\frac{1}{1+\rho}(1+r)$$

$$C_t = \frac{1+\rho}{1+r}C^E_{t+1}. \qquad \text{(Euler equation)}$$

Deriving the permanent income hypothesis consumption function

The simplest case of the Euler equation is when the rate of interest is equal to the consumer's subjective discount rate $r = \rho$. This is the case we have focused on throughout this chapter. We can see from the Euler equation that this implies $C_t = C^E_{t+1}$. In other words, it implies that consumption is constant in expectation in all future periods.

In order to derive the permanent income hypothesis, we make the additional assumption that future consumption is known with certainty, so that $C_t = C_{t+1} = \cdots = C_{t+i} = C$. To calculate C, we need to set out the expression for the present value of consumption, $\overline{C_t}$:

$$\overline{C_t} = C_t + \frac{C_{t+1}}{1+r} + \ldots + \frac{C_{t+i}}{(1+r)^i} + \ldots = C\left(1 + \frac{1}{1+r} + \ldots + \frac{1}{(1+r)^i} + \ldots\right)$$

$$= C\left(\frac{1}{1-\frac{1}{1+r}}\right) = C\left(\frac{1+r}{r}\right). \qquad \text{(present value of consumption)}$$

The permanent income hypothesis implies that the present value of consumption has to be equal to the present value of income, Ψ^E_t, so that:

$$\overline{C_t} = C\left(\frac{1+r}{r}\right) = \Psi^E_t$$

$$C_t = C = \left(\frac{r}{1+r}\right)\Psi^E_t, \qquad \text{(PIH consumption function)}$$

where C_t is the amount the consumer consumes in each period. It is their permanent income and is equal to the annuity value of their expected lifetime wealth.

1.5 Questions

1.5.1 Checklist questions

1. What is the *IS* curve? Why does it slope downwards?
2. Use the equation for the *IS* curve shown in Section 1.2.3 to answer the following questions:
   (a) If we assume that $0 < c_1, t < 1$, then what can we say about the size of multiplier?
   (b) If there is a decrease in government spending of ΔG, by how much does this decrease output?
   (c) Describe the feedback process that means a decrease in government spending can decrease output by more than 1 : 1.
3. Use the Keynesian cross to show the effect of a decrease in autonomous investment on the economy. Discuss the path of the economy as it adjusts to the new medium-run equilibrium. Why does the economy not continue to contract?
4. Use the Keynesian cross to illustrate the paradox of thrift. Model the change in savings behaviour as an increase in the marginal propensity to save, s_1 (remember that $c_1 + s_1 = 1$). Show how a rise in investment can counteract the reduction in output associated with the rise in savings.
5. Use the equation for the *IS* curve shown in Section 1.2.3 and Fig. 1.6 to discuss what happens to the *IS* curve in response to the following shocks. In each case provide a real world example of what might cause the shock.
   (a) An increase in autonomous consumption (i.e. $\uparrow c_0$).
   (b) A reduction in the interest sensitivity of investment (i.e. $\downarrow a_1$).
   (c) An increase in the marginal rate of taxation (i.e. $\uparrow t$).
6. According to the permanent income hypothesis, how will the paths of borrowing and consumption change in response to:
   (a) A temporary decrease in income when it occurs.
   (b) A permanent decrease in income when it occurs.
   (c) Are the answers different if the changes in income are unanticipated, i.e. if they are 'news'? Comment on the size of the marginal propensity to consume and the size of the multiplier.
7. Assuming the real interest rate is 4%, calculate how, according to the PIH, consumption and borrowing would change in each of the following cases
   (a) A stock market crash permanently reduces the value of an individual's assets by 1,000.
   (b) Households are told that in a year's time, they will receive a one-off bonus of 1,000. Then in one year's time, it is not paid.
   (c) Comment briefly on your results.
8. Explain the concepts of excess sensitivity and excess smoothness that arise from the empirical literature on the permanent income hypothesis. What could explain these findings?
9. What does Tobin's q tell us firms' investment decisions depend upon? According to Tobin's q, when should a firm invest?

10. What is the key problem with measuring marginal q? Is there an alternative measurement that can be used instead? Is this alternative measurement likely to be an accurate proxy for marginal q?
11. Use Section 1.2.7 to discuss what is expected to happen to the *IS* curve in response to the following shocks:
    (a) A crash in the stock market.
    (b) An increase in the retirement age.
    (c) A decrease in the rate of depreciation.
    (d) An increase in the cost of oil.
    (e) An increase in the rate of technological progress.

1.5.2 Problems and questions for discussion

1. This question involves collecting data from national statistics agencies (e.g. the UK Office for National Statistics) and/or international organizations (e.g. the OECD or the IMF). First, select an emerging and a developed economy.
    (a) Collect annual data on real GDP as far back as it is available for both countries. Convert the data to the log scale and plot on a graph. Comment on your findings. Calculate the growth rates of GDP for the two countries over the period and plot them on a graph. How do the business cycles of the two countries compare? To what extent do they appear synchronized?
    (b) Collect current price data on GDP and its components for the two countries. Calculate the percentage of GDP according to each of the four main types of expenditure—i.e. household consumption, government consumption, gross fixed capital formation and net exports (i.e. exports minus imports) and plot the series. How does the composition of GDP and changes over time in the two countries compare? Discuss possible reasons for the differences.
2. We start this question with a simple version of the aggregate demand function, where, to keep the maths simple, we omit taxation and government spending:

$$y^D = c_0 + c_1 y + \bar{I}. \tag{1.14}$$

Now, assume that $c_0 = 200$ and $\bar{I} = 200$. In addition, assume that $c_1 = 0.8$, such that:

$$y^D = 200 + 0.8y + 200. \tag{1.15}$$

(a) What is the level of output in goods market equilibrium?
(b) Assume there is a fall in the marginal propensity to consume, c_1, to 0.6—i.e. there is a rise in the savings rate, as $s_1 + c_1 = 1$. If we assume that the rise in planned savings leads to the accumulation of unplanned inventories of goods and does not increase planned investment ($\bar{I}$), then what is the level of output that satisfies the new goods market equilibrium?
(c) Compare the level of savings in the new and old goods market equilibria. Note that in goods market equilibrium; $S = s_1 y - c_0$ and $I = \bar{I}$.
(d) Comment on your findings in (c) in response to the question of whether it is advisable for the policy maker to encourage households to increase their savings to help escape a recession. Describe in words a mechanism, which is not included in this model, that could provide a connection from a policy encouraging households to save more and exit from a recession.

3. Use Section 1.2.4 and appropriate readings from beyond this book to decide whether the following statements (S1 and S2) are both true or whether one of them is false. Justify your answer.

    S1: According to Tobin's q theory, the path of investment is independent of current cash flow (and profits).

    S2: The empirical evidence shows that current cash flow is an important determinant of investment.

4. Aggregate consumption varies less than GDP and aggregate investment varies more. Can you reconcile these observations with the assumption that consumption and investment decisions are taken by rational, forward-looking agents?

2 The supply side

2.1 Overview

2.1.1 Unemployment

Unemployment is a characteristic feature of a market economy. There are unemployed people looking for work who would be prepared to take a job at the going wage but cannot get a job offer. This is involuntary unemployment and it reflects the fact that the labour market does not clear.

Figure 2.1 illustrates how unemployment varies over time and across countries. We can see that there is a large amount of heterogeneity in unemployment across the developed economies in the figure; following the global financial crisis, average unemployment between 2009 and 2012 ranged from 4.5% in the Netherlands to 21.2% in Spain. In addition, the figure highlights that an individual country can experience large changes in unemployment over time; Irish unemployment fell from 16% in the late 1980s to just above 4% in the early 2000s, before rising again to 14% in the wake of the economic crises of 2009–12.

Unemployment is costly to the economy since it represents a waste of resources, and it is associated with unhappiness and psychological distress to those affected. It is a major source of concern for policy makers.

In this chapter, we develop the supply side of the macroeconomic framework to provide a model of unemployment. Whereas the demand side refers to spending decisions, the supply side refers to production activities in the economy. Although both capital and labour are inputs to the production process, we concentrate on the input of labour. The supply side refers to both the supply of labour by workers and the demand for labour by firms. We return to discuss capital accumulation when we look at long-run growth in Chapter 8. On the supply side, we ask how firms decide how much output to produce, how many workers to employ and what prices to set. We want to know about decisions to work and how wages are set.

2.1.2 Why the labour market does not clear

When a market clears, the price is such that there is neither excess supply of, nor demand for, the good. If the market is for 'labour', then the wage would rise in response to excess demand and fall in response to excess supply, with the result that there would be no 'unsold' labour. As we shall see, this is not a good way of thinking about the labour market. The data in Fig. 2.1 covers many countries over more than five decades. It reveals unemployment as

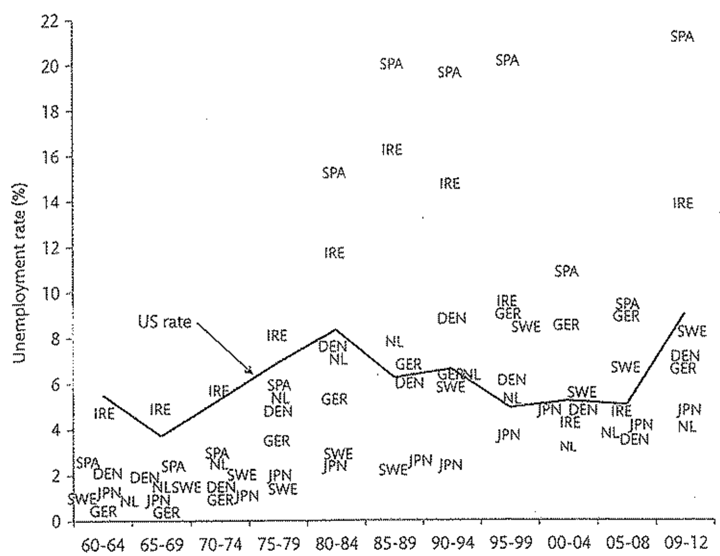


Figure 2.1 Trends and heterogeneity in unemployment for selected OECD economies, 1960–2012.

Source: Howell et al. (2007), Fig. 1.1, p. 10. Updated to 2012 using OECD harmonized unemployment rates.

characteristic of capitalist economies. The general case is one in which there are workers who are willing to work at the going wage but cannot find a job. Consistent with the phenomenon of involuntary unemployment is the large amount of evidence that workers are unhappy when they are unemployed. For example, in an early study, Clark and Oswald (1994) use data for a sample of six thousand British workers in 1991 and find that the unemployed have twice the level of mental distress of the employed.[1]

To understand why the labour market does not clear, we ask the obvious question of why an employer would not offer a job to a worker who is willing to work at a lower wage. One answer is that to be useful to an employer, a worker has to exert effort on the job and when the employer hires the worker at an hourly wage or a monthly or annual salary, she cannot pin down the supply of labour effort. Moreover, she cannot easily observe the effort the worker is making. The way the employer creates an incentive for the worker to work conscientiously is to create a cost to the worker of losing his job. When the employer chooses to set the wage above the level at which a worker would be prepared to take a job, the employee knows that if he loses his job he runs the risk of being unemployed. This makes him work effectively enough to keep his job. The term *efficiency wage setting* is used to describe this aspect of the supply side of the economy.

Thinking of the labour market in this way highlights two key factors that will influence the wage the employer sets.

1 For a review of the literature on unemployment and happiness, see Oswald (1997).

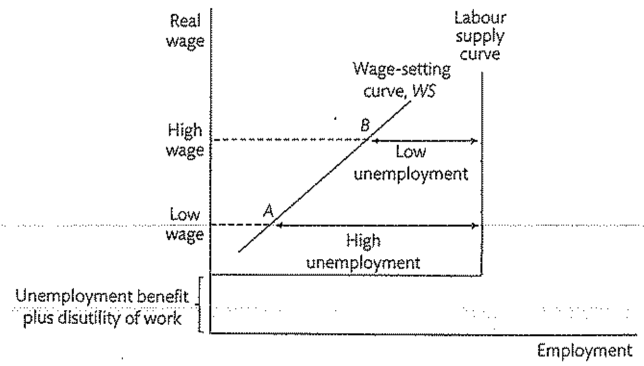


Figure 2.2 The wage-setting curve, *WS*.

1. First, is the income available to the worker if he is unemployed. If the alternative to working is claiming unemployment benefit and assuming there is a disutility associated with working, the utility the worker gets from working and receiving the wage must be higher than the utility from the alternative of not working and receiving unemployment benefit.
2. Second, the employer will have to offer a higher wage when the probability of the worker getting another job is higher: the wage she sets will rise as the economy-wide unemployment rate falls.

Figure 2.2 is a diagram with employment on the horizontal axis and the real wage on the vertical axis. The wage-setting curve is upward sloping. To get the worker to work hard, the wage is set above what the worker gets if he is not working (the unemployment benefit) plus the disutility of turning up to work, and it rises as unemployment falls. At point *A*, unemployment is high and a low wage is set: the cost of job loss is relatively high. At point *B*, unemployment is low. Given the relatively low cost of job loss, a higher wage is set.

A rise in the unemployment benefit will shift the wage-setting curve upwards and vice versa.

The labour supply curve is inverse-L shaped. The wage has to be above the unemployment benefit plus the disutility of work for workers to be willing to work effectively. The vertical portion of the labour supply curve is a simplification. Two factors lie behind a steep or 'inelastic' labour supply curve at the level of the economy as a whole. For workers who are already working, a rise in the wage has both an income and a substitution effect. The substitution effect suggests that labour supply rises with the wage because working becomes more worthwhile. The income effect suggests the opposite: since it is assumed that there is a disutility of working, if the same income can be earned from fewer hours of work, labour supply will fall. It is even possible for the labour supply curve to be backward bending if the income effect outweighs the substitution effect. The second factor concerns the decision to participate in the labour market. As the wage increases, participation goes up. There will be a steep labour supply curve if the tendency of workers to enter the labour force when the wage goes up is offset by a fall in hours of work by those who are working.

A point on the wage-setting curve is an answer to the question: what wage has to be paid to secure adequate worker effort (effective labour input) at a given level of unemployment (and associated level of employment)?

A point on the labour supply curve is an answer to the question: if a given wage were to be offered, how much labour would workers be willing to supply?

The horizontal gap between the wage curve and the labour supply curve is therefore involuntary unemployment, because it shows the extra supply of labour that would be offered at that wage but is not employed when wages are set on the wage-setting curve. The vertical gap between the wage-setting curve and the horizontal part of the labour supply curve is inversely related to the cost of job loss at the relevant level of employment: a higher efficiency wage has to be paid when the labour market is tight and it is relatively easy to get another job.

The efficiency wage explanation of why the labour market does not clear is based on the nature of the relationship between employer and worker.[2] Using the language of microeconomics, the special features of labour make it impossible to write a complete contract for the supply of labour in the same way that a contract can be written for the supply of energy or other inputs to the production process. The employer cannot specify in a contract every aspect of the work she requires of the worker and she is unable to perfectly monitor his output. Since worker effort is essential to the employer, it is the wage-setting curve and not the labour supply curve that is relevant for how wages in the economy are set, and, as we shall see, for pinning down the unemployment rate consistent with constant inflation. The outcome is that involuntary unemployment is characteristic of the economy.

It might be suggested that the employer use piece rates rather than an hourly wage to pay workers. By paying for the worker's output rather than the worker's time, the problem of not being able to observe effort is solved. But taking a look at any economy quickly reveals that very little economic activity is organized around piece rates for labour. This is particularly true of modern service-based economies, where output is less easily quantifiable. In addition, much of the production activity in the economy is based on team production, where the contribution of an individual worker is hard to determine.

2.1.3 Supply side effects on unemployment

The supply side effects on unemployment relate to both the labour supply and labour demand parts of the economy. The wage-setting curve relates to the *supply of labour* and shows that the real wage that has to be paid to workers rises as employment rises and lies above the labour supply curve. This curve shows the *wage-setting real wage*. This is positively sloped for the efficiency wage reasons explained in the previous subsection. In the modelling section, we discuss a number of other explanations for the wage-setting curve. One of these relates to the case where workers have market power. If the wage is set by a union, then the union can set wages above the efficiency wage. In this case, the wage-setting curve would be above the one that applies in the absence of unions.

[2] The efficiency wage model was first formalized in Shapiro and Stiglitz' landmark 1984 paper *Equilibrium Unemployment as a Worker Discipline Device.*

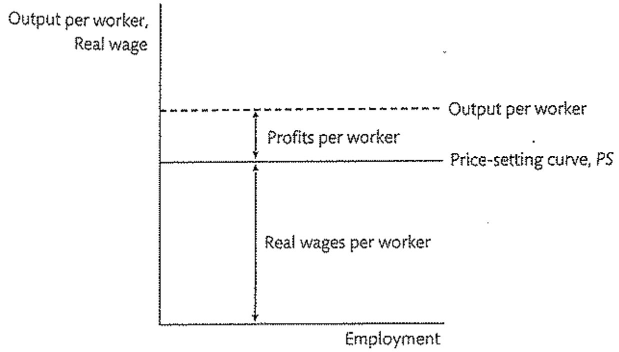


Figure 2.3 The price-setting curve, *PS*.

The other side of the labour market reflects the *demand for labour* by imperfectly competitive firms. This defines a second real wage: this is the real wage that makes production (and hence, the employment of workers) profitable for the price-setting firm. This curve shows the *price-setting real wage*.

As our baseline case, we use a simple model of labour demand where there is a constant real wage at which price-setting firms find it profitable to employ workers. In the modelling section, we derive the price-setting real wage from the firm's pricing decision. Here we focus on the outcome. We assume that the firm's production function is very simple: labour is the only input, and productivity, which is output per worker, is constant. In order to produce, the firm requires a fixed profit margin and this implies the real wage at which it finds it profitable to employ workers. Given these assumptions, the price setting real wage curve is flat (see Fig. 2.3). As shown in the diagram, output per worker is divided into two chunks: profit per worker and the real wage.

Figure 2.4 shows the wage-setting and price-setting curves. The intersection of the price-setting (*PS*) curve and the upward-sloping wage-setting (*WS*) curve determines the real wage, employment and unemployment when the supply side is in equilibrium. The intersection of

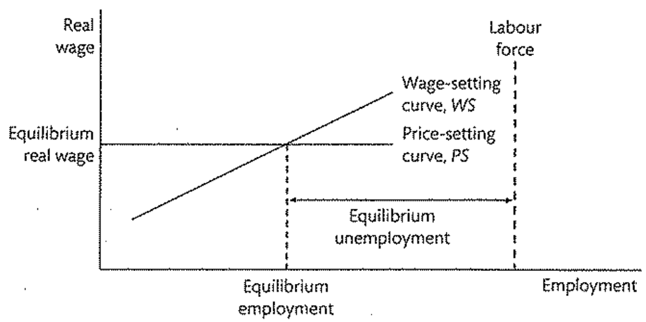


Figure 2.4 Supply-side equilibrium: the intersection of the wage-setting and price-setting curves.

the wage- and price-setting curves fixes the equilibrium real wage and employment level. At the equilibrium, the real wage is consistent with what is needed to secure sufficient labour (on the *WS* curve) and for production to be profitable (on the *PS* curve).

We can use Fig. 2.4 to show how the unemployment rate is defined. The labour force is shown by a vertical line. This is the total number of people reported by the statistical agency who are working or looking for work. The unemployment rate is defined as unemployment divided by the sum of employment and unemployment, that is, the numbers unemployed divided by the labour force:

$$\text{Unemployment rate} \equiv \frac{\text{unemployed}}{\text{employed} + \text{unemployed}} \equiv \frac{\text{unemployed}}{\text{labour force}}.$$

Equilibrium unemployment is the difference between employment at the *WS* – *PS* intersection and the labour force. From Fig. 2.5, it is clear that equilibrium unemployment will rise if the wage-setting curve shifts up or if the price-setting curve shifts down. Taking the case of a higher unemployment benefit first, we have already seen that a higher unemployment benefit will shift the *WS* curve up. In Fig. 2.5a, this shifts the *WS* curve up to *WS*′. The new *WS* – *PS* intersection is at a lower level of employment. The economics behind this can be explained as follows. For equilibrium, the real wage has to be on the *PS* curve; otherwise firms will not be prepared to produce output. An increase in unemployment benefits reduces the cost of job loss, which shifts the *WS* curve upwards. In order that the real wage on the wage-setting curve is brought down into line with the *PS* curve, it is necessary for unemployment to be higher. Given the higher out-of-work benefits, a higher risk of job loss (which is associated with higher unemployment) is necessary to make an unchanged real wage consistent with wage setters' behaviour.

Equilibrium unemployment will also be higher if the *PS* curve shifts down (from *PS* to *PS*′ in Fig. 2.5b). If there is a reduction in the extent of competition in the goods market, which allows the mark-up of price over unit labour costs to rise, the price-setting real wage will be reduced: under the new conditions, firms will only produce if the real wage is lower. In terms of Fig. 2.3, a higher mark-up increases the share of output per worker that goes to profits

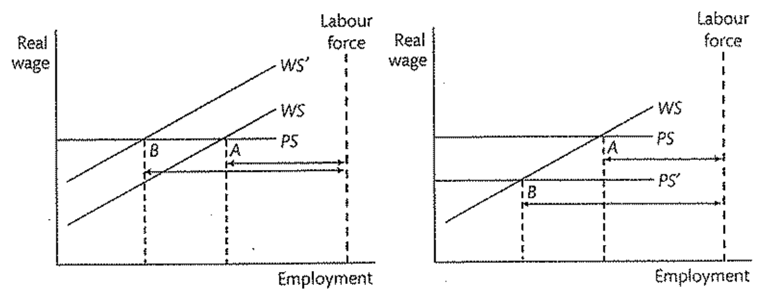


Figure 2.5 The impact on equilibrium unemployment of shifts in the *WS* and *PS* curves.

Note: the double-headed arrows show equilibrium unemployment for each case.

and reduces the share to wages, and hence, the real wage. The new equilibrium is at a lower real wage and higher unemployment. For the wage-setting real wage to be reduced to the new equilibrium level, there must be higher unemployment in the economy. With a higher cost of job loss, the efficiency real wage is lower.

The concept of equilibrium unemployment is a medium to long-run concept. It tells us the unemployment rate at which both wage and price setters are content with the prevailing real wage. The kinds of policies that can be implemented to affect equilibrium unemployment are those that shift the *WS* or the *PS* curve. These are *supply-side policies*.

In the modelling section and in Chapter 15 on supply-side policies, we enrich the model of the supply side by introducing additional institutions and policies that shift the wage- and price-setting curves. A more detailed examination of both institutions and policies is presented in Chapter 15. These include the role of unions, taxes, workplace policies that affect the disutility of exerting effort, labour and product market regulations, and competition policy. These factors help explain both the cross-country variation and the changes over time in the 5-year average unemployment rates shown in Fig. 2.1.

2.1.4 Nominal rigidities and demand-side policy

When discussing equilibrium unemployment, the focus is on the medium run when the economy is at a *WS* – *PS* intersection, and hence on the effect of supply-side policies and institutions on unemployment. In this subsection, we introduce *nominal rigidities* and the role of *demand-side* policies. The term nominal rigidity refers to the fact that nominal wages (in dollars or euros) and prices do not adjust immediately to fluctuations in aggregate demand to keep the economy at equilibrium unemployment.

In Chapter 1, we assumed an extreme form of this rigidity, namely that prices and wages don't change at all in response to shifts in demand. This meant that when aggregate demand shifted, as a result, for example of an investment boom, output and employment in the economy responded fully to bring supply into line with the higher demand. We used the multiplier process to model the response of output and employment based on the assumption that wages and prices remained unchanged. An alternative extreme assumption is that the economy is always at the medium-run equilibrium, i.e. at a *WS/PS* intersection, and therefore that fluctuations in aggregate demand do not affect output and employment. The terms 'fixed price' and 'flex price' are sometimes used to label these two extreme cases.

In building a model suitable for analysing real-world economies, we need a model that is neither 'fixed price' nor 'flex price'. The model we set out includes nominal rigidities but wages and prices are not completely fixed. Wages are changed at intervals, for example, at an annual wage review. Nominal rigidities characterize both wage and price setting. More detailed arguments and evidence are provided in the modelling section. Here we summarize the main points. Before doing so, it is important to distinguish between real and nominal wages.

Nominal wages, real wages and inflation

To this point, all of the discussion about the supply side has been in terms of the *real* wage. This is because it is the real wage that workers care about (i.e. how much their dollar or euro wage buys in goods and services) and therefore, that the employer has to be concerned with when setting the wage to secure the worker's effort. From the labour demand side, it is the cost of labour in terms of the price they get for their output that firms care about. However, workers are paid a wage in money terms: this is the *nominal* wage and firms set prices.

We can see how real and nominal wages are related in the model as follows. Given the unemployment rate, employers (or unions if they are setting wages) set the nominal wage so as to achieve their desired real wage on the wage setting curve. This means they must make an assumption about the *consumer price level* when they set the nominal wage. Similarly, firms set their *product price* making an assumption about the nominal wage to achieve the price-setting real wage on the *PS* curve. The consumer price level is the outcome of price-setting by firms across the economy.

Note that when wages are set by the employer, it would be the firm's human resources division that sets the nominal wage taking account of the consumer price level to ensure workers supply adequate effort for the firm's operations. Given the nominal wage, the marketing and strategy divisions set the price for the firm's product.

Modern economies are typically characterized by a positive rate of inflation. Central banks often target an inflation rate of 2%. If nominal wages rise by 2% per annum and prices rise by 2% per annum, the real wage is unchanged. At a *WS* – *PS* intersection, the real wage is constant, which means nominal wages and prices are rising at the same rate. This is a constant inflation equilibrium.

Rigidities in wage setting

Nominal wages are set periodically, usually at an annual wage round or review. They are not continuously adjusted by the wage setter. Barattieri et al. (2010) analyse US survey data and find that wages are indeed very 'sticky'; they are most likely to be adjusted just once every year. The authors also find that there is little variation in the frequency of wage changes across different industries and occupations.

An important reason for the stickiness of wages relates to the interactions between employers and workers, where considerations of fairness and morale play an important role. Campbell and Kamlani (1997) surveyed 184 firms and find strong support for two explanations of wage rigidity (in this case, why wages are not cut as much as possible in a recession). First, firms were concerned that wage cuts would lead their most productive workers to quit, whereas layoffs could target the least productive workers. Second, respondents linked wage cuts to reduced worker effort, particularly when wage cuts were viewed as being 'unfair'. Bewley (2007) reviews the empirical literature on fairness and wage rigidity and finds that fairness is one of the primary determinants of company morale. He also finds that wage cuts are only viewed as being fair when they are seen as saving a large number of jobs.

Rigidities in price setting

The frequency of price changes varies a great deal across industries. However, for large parts of the economy, firms are cautious about frequent price adjustments in response to fluctuations in demand because of their concern about their competitors' and customers' reactions.

A detailed survey of how firms set prices using a structured questionnaire of a random sample of 330 US firms found that almost half of all prices were changed no more than

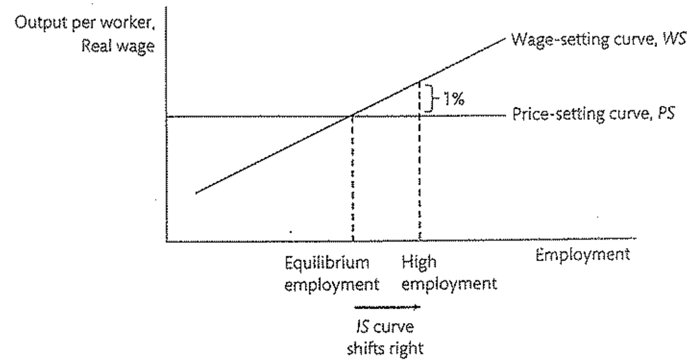


Figure 2.6 The response of wages and prices to a cyclical upswing.

annually. The survey identified the following major reasons for price stickiness: firms are deterred from raising prices because of concerns their competitors will not follow suit; cost increases are generally industry-wide and can serve as convenient signals that other firms are probably under pressure to raise prices; implicit contracts with customers deter price hikes when demand rises but permit them when costs rise; and firms hold prices steady until the next regularly scheduled price review.[3]

Aggregate demand shocks and the role of nominal rigidities

In the macroeconomic model, we make simplifying assumptions about how wage and price rigidities operate. The simplifying assumptions reflect the arguments presented about why wages and prices are adjusted infrequently. We assume that wage contracts are reviewed annually at the 'wage round'. This means that if an aggregate demand shock occurs in the interval between wage rounds, the wage is not adjusted. We assume that prices are not adjusted in response to shifts in aggregate demand; however, prices are adjusted immediately following the wage round. Since the adjustment in wages will reflect changes in aggregate demand, prices respond to changes in aggregate demand with a lag.

The timeline for wage and price setting, shocks and changes in output and employment can be summarized as follows:

Aggregate demand shock → output and employment change

Next wage round → nominal wages change

Immediately after the wage round → prices change

Before looking at an aggregate demand shock, we look at what happens when the economy is at supply-side equilibrium in an environment of constant inflation. The equilibrium is where the *WS* and *PS* curves intersect in Fig. 2.6. Inflation in the previous year was 2%.

At the next annual wage round, wage setters observe that prices have risen 2% over the previous year and they will set a 2% nominal wage increase to keep the real wage constant.

[3] See Blinder et al. (1998).

After the wage round, firms will observe that the 2% increase in wages has increased their costs by 2%. If they raise their prices by 2%, this will keep their profit margin constant (the real wage will remain on the *PS* curve). As long as the economy remains at the equilibrium level of employment, wage and price inflation will remain unchanged at 2% per year and the real wage will remain constant.

We now look at what happens when there is an aggregate demand shock of the kind discussed in Chapter 1: for example, an investment boom raises output. Since the investment boom does not affect the *WS* or the *PS* curve, it has no effect on equilibrium unemployment. The investment boom raises output and employment in the short run: we assume output per worker is constant. As Fig. 2.6 shows, the economy is no longer in medium-run equilibrium.

What happens the next time there is an opportunity for wages to be adjusted? At the next wage round, employment is higher and wage setters will respond by setting a nominal wage increase to take the real wage up to the point on the *WS* curve at the higher level of employment. They will need a 2% nominal wage increase as usual, plus an additional increase. In the example shown in Fig. 2.6, this is an extra 1%. Nominal wages go up by 3%. How do price setters respond to the higher wages? Their costs have risen by 3% (rather than the usual 2%), so they mark up this cost increase in their prices and price inflation goes up from 2% to 3%. Firms are likely to feel comfortable about raising their prices at this point, because they can observe that wages have gone up during the annual wage round for their competitors as well. We have the result that an expansion of aggregate demand that pushes output and employment above the equilibrium level is followed by a rise in wage and price inflation.

To summarize, when the economy is at equilibrium unemployment, inflation is constant. A demand shock will shift the economy away from equilibrium because wages and prices do not adjust instantly to keep the economy at equilibrium. Instead, output and employment respond to the change in aggregate demand. In the example of an investment boom, this leads to a fall in unemployment and a rise in inflation. More generally, a positive demand shock leads to a movement away from equilibrium with higher employment and a rise in inflation; a negative demand shock leads to lower employment and lower inflation.

This behaviour of the economy, which is called an upswing or downswing of the *business cycle*, creates a role for a policy maker who aims to improve welfare by keeping the economy close to the medium-run equilibrium unemployment rate. The involuntary unemployment characteristic of the economy when it is at the constant inflation equilibrium (i.e. *WS* – *PS* intersection) is raised in a business cycle downswing and lowered in an upswing. A policy maker who focuses on stabilization and uses demand side policy is introduced in Chapter 3.

2.1.5 Facts about the supply side

Efficiency wage setting

There is a large body of empirical evidence documenting the existence of a 'wage curve', which is represented in the model by the *WS* curve.[4] Microeconomic data on local labour

[4] For an introduction to the empirical literature on using microeconomic evidence to estimate wage curves, see Blanchflower and Oswald (1995).

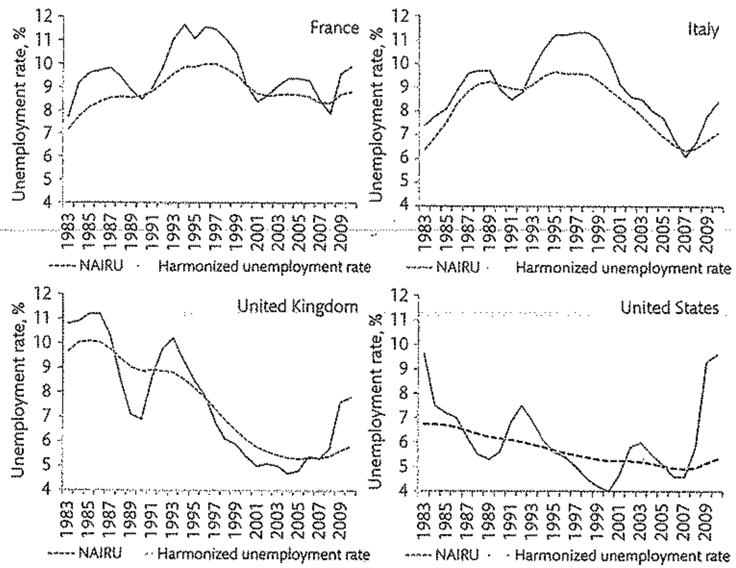


Figure 2.8 Non-Accelerating Inflation Rate of Unemployment (NAIRU) and harmonized unemployment rates in France, Italy, the United Kingdom and the United States: 1983–2010.

Source: OECD Economic Outlook (accessed December 2011).

Patterns of unemployment differed between the continental European economies (France and Italy) and the UK and US over this period. Unemployment rates in France and Italy were consistently above equilibrium, whereas there were periods when unemployment was both above and below equilibrium in the UK and the US.

Figure 2.8 also shows that actual and equilibrium unemployment were on average higher in these continental European economies than in the Anglo-Saxon ones over the period, with all countries excluding France experiencing a sustained drop in the equilibrium rate of unemployment from the mid nineties to the onset of the global financial crisis in 2008. In all four countries, actual unemployment rose dramatically during the global economic downturn and was slow to fall thereafter.

In Fig. 2.8, we saw that unemployment differs from equilibrium unemployment over the economic cycle. In the *WS* – *PS* model, when output is above equilibrium (i.e. when unemployment is below equilibrium), this creates upward inflationary pressure and vice versa.

Figure 2.9 displays the inflation rates in selected OECD countries from the 1960s to the present: whilst inflation in the past decade has been low at 2% to 3% in many of the OECD countries, it was much higher, and often at rates between 10% and 20% in the 1970s. By comparing Fig. 2.8 and Fig. 2.9, we can see that the decline in inflation in France and Italy from high rates at the end of the 1970s to very low rates in the mid 2000s is consistent with the persistence of unemployment above the estimated constant inflation rate during this period.

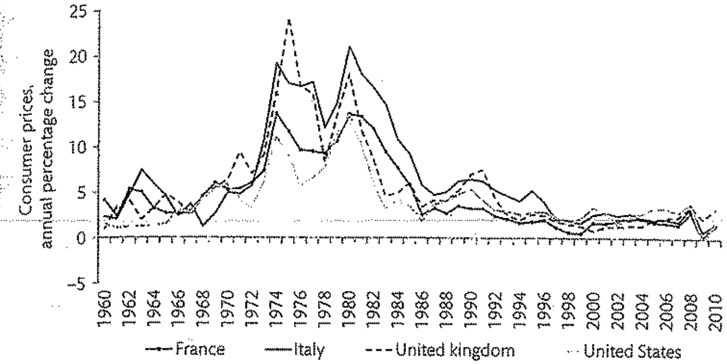


Figure 2.9 Consumer price inflation rates in France, Italy, the United Kingdom and the United States: 1960–2010.

Source: OECD Monthly Economic Indicators (accessed January 2012).

In Chapter 3, we look at the costs of getting inflation down from its high levels in the 1970s in terms of prolonged high unemployment. The years of so-called 'stagflation'—a combination of high unemployment and high inflation—imposed heavy costs on economies. As a consequence, policy makers sought to find better ways of managing the economy. One outcome was a move to inflation targeting and central bank independence.

2.2 Modelling

Following the distinction between the role played by real and nominal rigidities introduced in Section 2.1, we break down the modelling of the supply side in this section into two parts, each of which is focused on an important question.

1. Supply-side effects on unemployment. In Section 2.2.1, we shall look in more detail at the structural features of the economy that determine the medium-run equilibrium level of output and unemployment around which the economy fluctuates over the course of the business cycle.
2. Nominal rigidities, inflation and the business cycle. What happens when the economy is away from equilibrium—e.g. as a consequence of fluctuations in aggregate demand? In particular, what is the response of inflation? Answering this question in Section 2.2.2 involves extending the discussion of inflation from Section 2.1 and introducing Phillips curves.

In the applications section (2.3) at the end of the chapter, we show how we can combine the demand and supply sides to model the business cycle. Shocks to aggregate demand and to the supply side move the economy away from equilibrium. This leads us into Chapter 3, where we introduce the policy maker and show how policy intervention can stabilize the economy following a shock.

2.2.1 Supply side effects on unemployment

In Section 2.1, we begin with the fact that the employer cannot write and enforce a complete contract specifying work effort. Faced with this problem, the employer will *choose* to pay a wage that is higher than the wage at which the worker would be prepared to work. The implication is that there is involuntary unemployment in equilibrium. In the Appendix, we set out the textbook model of a perfectly competitive labour market where the labour market clears in equilibrium (we call this the hiring hall case). Some readers may find it is a useful reference point because it shows how the textbook model differs from the model of the supply side that we use in this book.

Incomplete employment contracts and price-setting firms

Since involuntary unemployment is a ubiquitous characteristic of market economies, this is a defining feature of the supply-side equilibrium in the macro model we are building. As we discuss in Section 2.1, wages will be set above the level at which a worker would supply labour. This defines the wage-setting curve. Facing downward-sloping demand curves for their products, firms set prices as a mark-up over their costs. This defines the price-setting curve. The real wage and equilibrium employment are the outcome of these wage- and price-setting decisions across the economy.

The labour supply side and wage setting

Efficiency wage setting means there must be a cost of job loss. At the prevailing real wage, say, w_0, additional workers would be prepared to work. However, they cannot bid down the wage because the employer knows that given the state of the labour market, unless w_0 is paid, the cost of job loss will not be high enough for workers to stay and work diligently.

The economics and management literatures stress a broad range of incentive problems faced by employers because of information problems in the labour market. In addition to the role of motivation that we have discussed, employers face problems of retention and recruitment.

Worker turnover is costly to employers. Workers move from job to job, from employment to unemployment (and vice versa) and also leave and enter the labour force. In the simple explanation of efficiency wages set out above, the worker either worked hard enough to be valuable to the employer or not. A richer model would include the employer's need to pay a wage that takes account of the costs she incurs due to turnover. The employer's optimal response to this is to raise the wage—it pays the employer to pay a higher wage in order to reduce turnover. Moreover, when unemployment is lower, it is easier for a worker to find another job and turnover will go up. This provides another explanation for the positively sloped wage-setting curve that is above the labour supply curve.

Similar arguments apply to recruitment. Recruitment in the real world is normally a difficult process, in which the firm has to compete against other firms to get the new hires it wants, and having unfilled vacancies is costly to the firm. In this case, it pays the employer to pay a higher wage in order to reduce the incidence of unfilled vacancies. A realistic assumption is that the proportion of its vacancies that it can fill depends positively on the wage it offers relative to its competitors and negatively on the tightness of the labour market.

Taking account of all of these considerations, the unemployment rate is therefore a key determinant of the 'wage-setting real wage', which lies above the labour supply curve.

In terms of nominal wages, and writing the wage setting equation in terms of employment, it is

$$W = P^E \cdot B(N, \mathbf{z}_w),$$

where P^E is the expected price level and B is a positive function of the level of employment, N, and a set of wage-push variables, $\mathbf{z}_w$. Whether wages are set by unions, employers, or through bargaining, it is the *nominal* (i.e. money) wage that is fixed. However, workers will evaluate wage offers in terms of the *real* wage it is expected to deliver—i.e. it is the nominal wage relative to the expected consumer price level that affects workers' standard of living and hence their utility.

The wage equation can be written in terms of real wages to define the upward-sloping wage-setting curve:

$$w^{WS} = \frac{W}{P^E} = B(N, \mathbf{z}_w). \qquad \text{(wage-setting real wage equation, } WS\text{)}$$

We set out a simple formal model of efficiency wage setting in Chapter 15. The basic diagram is shown in Fig. 2.10. The opportunity cost of taking a job consists of two elements: the unemployment benefit and the disutility of exerting effort when working. The wage on the wage-setting curve is above the opportunity cost of working and increases as the unemployment rate falls. As we shall see in Chapter 15, the equation for the wage-setting curve is non-linear: this is because as involuntary unemployment approaches zero, the cost of job loss goes toward zero and the wage that has to be paid to elicit effort goes toward infinity. Similarly, as the wage approaches the opportunity cost of working, it will not be possible to recruit labour. The figure highlights the presence of both voluntary and involuntary unemployment when the economy is at a medium-run equilibrium on the *WS* curve.

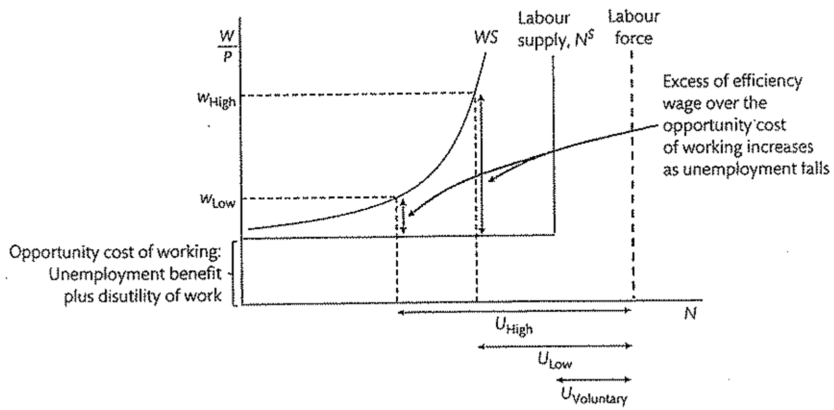


Figure 2.10 Efficiency wage setting.

To simplify the diagrams, we normally draw the *WS* curve as linear: it will be shifted by the unemployment benefit (Fig. 2.5a) and by the other wage push factors, $\mathbf{z}_w$, including the effects of unions.

When workers have market power in the labour market, then the gap between the wage on the *WS* curve and the opportunity cost of working consists of:

1. the mark-up per worker (in real terms) that has to be paid in order to get the worker to exert effort and
2. any additional mark-up associated with worker bargaining power.

Workers can have bargaining power in the labour market because they have particular skills or because they are organized into unions. In unionized workplaces, wages are set through negotiations between the employer and the union. A simplified model of union wage-setting takes the case of the so-called monopoly union, where the union can unilaterally set the wage. The union sets the wage in the interests of its members, who are concerned with both the real wage and employment. It aims to strike a balance between (i) too high a wage, which will push up the price of the firm's product and decrease demand for the firm's output and hence reduce employment; and (ii) too low a wage, which will fail to use the union's monopoly power to secure better living standards over and above the efficiency wage. A model of union wage-setting is developed in Chapter 15.

Shifts in the WS curve

The wage-push factors in the wage-setting curve (the $\mathbf{z}_w$'s) include institutional, policy, structural and shock variables.

The *WS* curve shifts down, which reduces equilibrium unemployment, when:

1. There is a fall in the level of unemployment benefits or their duration. This raises the cost of job loss.
2. There is a fall in the disutility of effort. Improvements in working conditions shift the *WS* curve downwards because they increase the cost of job loss, reducing the wage that has to be paid to get workers to work effectively.
3. Unions are given less legal protection. This reduces the union mark-up.
4. Unions are weaker, for example as measured by a lower proportion of trade union members amongst employees (lower trade union density) or when a lower proportion of employees are covered by collective bargaining agreements (lower collective bargaining coverage). (We discuss the evolution of trade union density and coverage for OECD countries over the last 50 years in Chapter 15.)
5. Unions agree to exercise bargaining restraint in the context, for example, of a wage accord, because this lowers the real wage that is acceptable to unions at a given unemployment rate.

The labour demand side and price setting

In a textbook, perfectly competitive goods market, firms are price and wage-takers: they maximize profits by taking the market price, P, and setting it equal to their marginal cost:

$$\begin{aligned} P &= MC \\ &= \frac{W}{MPL} \\ \Rightarrow \frac{W}{P} &= MPL. \end{aligned}$$

The marginal product of labour (or labour demand) curve therefore shows the employment level a competitive firm will choose at a given real wage.

By contrast, firms facing downward-sloping demand curves *set a price* to maximize profits. The mark-up on marginal cost will depend on the elasticity of demand, which is the responsiveness of output demanded to a change in price. As the elasticity of demand rises, the mark-up falls until we get to the special case of perfect competition, where the elasticity of demand is infinite and the price is the one that clears the market.

When a firm faces a downward-sloping demand curve (and assuming for simplicity it does not behave strategically towards other firms), it maximizes profits when marginal revenue is equal to marginal cost. If the (absolute value of the) elasticity of demand (η, called eta) is constant, then there is a constant mark-up of ($\frac{1}{\eta-1} = \mu$) and we have the price-setting formula (derived in the Appendix):

$$P = \left(1 + \frac{1}{\eta - 1}\right)\left(\frac{W}{MPL}\right) \equiv (1+\mu)\left(\frac{W}{MPL}\right). \qquad \text{(price-setting (mark-up) equation)}$$

For example, if the elasticity of demand is 8, $\mu = 0.14$, and the price is set at 14% above marginal cost. The Greek letter μ, mu, is pronounced 'mew'.

The next step is to derive the price-setting real wage from the price-setting equation:

$$\begin{aligned} \frac{W}{P} &= \frac{1}{(1+\mu)} MPL \qquad (2.1) \\ &= (1-\mu) MPL, \qquad \text{(price-setting real wage)} \end{aligned}$$

where to simplify the equations, we use the approximation $\frac{1}{(1+\mu)} \approx (1-\mu)$, which holds for low values of the mark-up, μ.[9]

Figure 2.11 illustrates the *PS* curve: the price-setting real wage is a fraction of the marginal product of labour. For the example of the elasticity of demand of 8, the price-setting real wage is 86% of the level of the marginal product of labour.

The excess of the real wage on the labour demand curve above that on the *PS* curve at any level of employment is the supernormal profits per worker (in real terms) associated with imperfect competition in the product market.

For simplicity in working with the macro model, we normally use a horizontal rather than a downward-sloping *PS* curve. As we have seen, switching from perfect to imperfect competition does not in itself lead to a horizontal *PS* curve. Additional assumptions are required.

[9] For the (absolute value of the) elasticity of demand above 5, the approximation is close. For other values of the elasticity, it is necessary to use the exact definition of the mark-up.

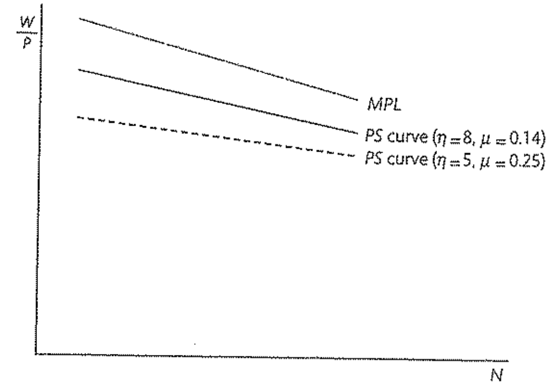


Figure 2.11 Relationship between the *MPL*, the price elasticity of demand (η), and the *PS* curve.

Two alternatives are:

1. If the marginal product of labour is constant (which implies that it is equal to the average product) and the mark-up is constant, the price-setting real wage is equal to a constant fraction of labour productivity.
2. If firms set their prices using a rule of thumb, basing their price on their average costs over the business cycle (i.e. as the economy moves from recession to boom and vice versa) the *PS* curve would also flatten. Such a 'normal cost pricing' rule might result from firms wishing to limit the extent to which they modify their prices in response to changes in costs associated with changes in demand.

When looking explicitly at different patterns of how the real wage moves over the course of the business cycle or at the implications of supply-side policies for real wages, we should consider the more general downward-sloping *PS* curve. In other cases, it is more straightforward to use a flat *PS* curve.

Taking the simplest case, to derive a flat *PS* curve, we assume constant productivity and a constant mark-up. Given these assumptions, if firms set prices to deliver a specific profit margin, then the fixed amount of output per worker is split into two parts: profits per worker and real wages per worker. The real wage implied by pricing behaviour is therefore constant and the *PS* curve is flat. Price setting can then be summarized as the marking up of unit labour costs by a fixed percentage, μ,

$$P = (1+\mu)\left(\frac{W}{\lambda}\right), \tag{2.2}$$

where unit labour costs are the cost of labour per unit of output; i.e. $W \times N$ divided by y. We define $\frac{y}{N}$ (output per worker) as λ (lambda, labour productivity) and using the same approximation as above, we have:

$$\frac{W}{P} = \frac{1}{(1+\mu)}\lambda \tag{2.3}$$

$$= (1-\mu)\lambda. \tag{2.4}$$

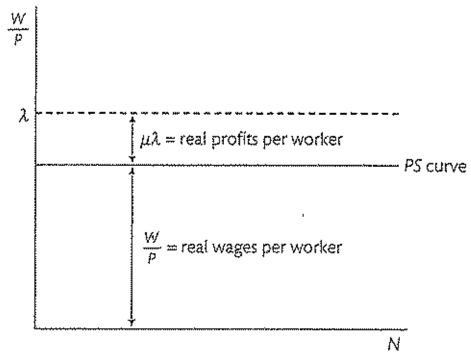


Figure 2.12 The price-setting real wage curve: *PS*.

We write the price setting real wage equation:

$$w^{PS} = \frac{W}{P} = \lambda(1-\mu). \qquad \text{(}PS\text{, price-setting real wage equation)}$$

In other words, given the mark-up, the level of labour productivity, and the nominal wage, the price level set by firms implies a specific value of the real wage.

A simple rearrangement of the *PS* equation shows how output per worker is decomposed into real profits per worker and the real wage as a result of the firm's price-setting decision and this is illustrated in Fig. 2.12:

$$\lambda = \mu\lambda + \frac{W}{P} \tag{2.5}$$

output per worker = real profits per worker + real wages per worker.

Shifts in the *PS* curve

The tax wedge: how are W and P measured?

Before providing examples of the price-shift variables, one clarification is needed. Once income taxes, labour taxes such as social security contributions and indirect taxes such as VAT are introduced, we have to be clear about what W and P measure and which measure we show on the axis in the labour market diagram. This is a matter of choosing a convention, and we find it convenient to show the real consumption wage in the labour market diagram. This entails measuring W as the post-tax money wage paid to the employee and to measure P_c as the consumer price index, i.e. inclusive of indirect taxes, t_v.

$$P_c = P(1+t_v).$$

This means that when we show $w = W/P_c$ on the axis of the labour market diagram, this is the real consumption wage—the concept relevant from the perspective of the utility of the worker.

By contrast, the real wage that is of relevance to the employer is the real product wage, which is the full cost of labour to firms—inclusive of income tax and non-wage labour costs such as social security contributions paid by employers and employees—divided by the price the firm gets for its product (i.e. excluding indirect taxes). This is called the producer price. The difference between the real consumption wage and the real product wage is called the *tax wedge*. Given the way we have defined the labour market diagram in terms of W and P_c, the wedge shows up as a price-push factor. Any increase in either direct or indirect taxation reduces the price-setting real wage and therefore shifts the *PS* curve downwards. The derivation of the *PS* curve including the tax wedge is shown in the Appendix.

Price-push factors

We incorporate the tax wedge as one of the price-push factors and write the *PS* curve compactly as:

$$w^{PS} = \lambda F(\mu, z_p), \qquad \text{(PS curve including price-push factors)}$$

where z_p is a set of price push variables including the tax wedge.

The *PS* curve shifts up, reducing equilibrium unemployment, when there is

1. a fall in the tax wedge, which is included in z_p,
2. a fall in the mark-up, μ, due, for example, to a change in competitive conditions such as tougher competition policy rules or enforcement, or
3. a rise in productivity, λ.

It is important to note that what matters for shifting the *PS* curve and therefore for affecting equilibrium unemployment is the tax wedge as a whole: a rise in income tax or in indirect tax will push equilibrium unemployment up. There is nothing special about the effect of the so-called payroll taxes, i.e. the employer and employee social security contributions. In Chapter 15, we look at the evidence about how much the equilibrium rate of unemployment is expected to increase if the tax wedge increases by ten percentage points.

Other factors included in z_p may be regulations that increase the cost of employment, such as business registration and some employment regulations. However, such regulations do not necessarily have the effect of increasing price push and therefore raising the equilibrium rate of unemployment. For example, although regulations enforcing health and safety standards impose costs on firms, they may have a compensating positive effect on productivity.

Equilibrium in the labour market

The labour market is characterized by an upward-sloping *WS* curve and a flat or downward-sloping *PS* curve. The labour market is in equilibrium where the curves cross (see Fig. 2.13):

$$w^{WS} = w^{PS}$$
$$B(N, z_w) = \lambda F(\mu, z_p), \qquad \text{(labour market equilibrium)}$$

and this defines the unique equilibrium level of employment N_e. The associated equilibrium rate of unemployment is U_e/L, where L is the labour force.

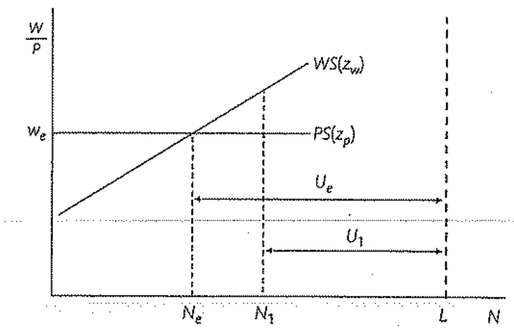


Figure 2.13 Equilibrium employment and unemployment: N_e and U_e.

Unemployment in equilibrium

The equilibrium rate of unemployment is the outcome of the structural or supply-side features of the economy that lie behind the wage-setting and price-setting curves. It can therefore in principle be changed by supply-side policies or structural changes that affect either the wage-push or price-push factors.

An increase in the degree of product market competition—as a result, say, of changes in the application of competition policy or because the internet makes it easier to compare prices—would produce a lower profit margin (μ) and a higher real wage at each level of employment (the *PS* curve would shift up). Similarly, any government policy change that affects wage- and price-setting outcomes will shift equilibrium unemployment. Policies related to the cost of job loss such as unemployment benefits, changes in taxation, labour, and product market regulation and incomes accords are all relevant. It is thus easy to imagine that international differences in policy and in institutional structures produce differences in equilibrium unemployment. We return to analyse supply-side shifts and the evidence regarding their role in explaining cross-country unemployment trends in Chapter 15.

2.2.2 Nominal rigidities, inflation and the business cycle

Demand-driven cycles

In Section 2.1 we explained that the economy will be shifted away from the medium-run equilibrium (where the *WS* and *PS* curves intersect) by fluctuations in aggregate demand. Wages and prices do not adjust spontaneously to keep the economy at equilibrium unemployment. We reported the survey evidence that wages are set periodically and that employers do not cut nominal wages when unemployment goes up.

What about prices? Fluctuations in aggregate demand lead to business cycle upswings and downswings if firms respond to changes in demand by altering output and employment. If prices and wages change but output does not change, then we would not observe demand-driven business cycles. One explanation for demand-driven cycles is that imperfectly competitive firms find it profitable to respond to shifts in demand by changing their output.[10] Price stickiness refers to the reluctance of firms to change prices in the face of changes in aggregate demand.

The demand curve faced by an imperfectly competitive firm is shifted by shocks to aggregate demand in the economy of the kind introduced in Chapter 1. It is clear from the mark-up equation (price-setting (mark-up) equation) that when profits are maximized, the price will be above marginal cost: the firm will make supernormal profits.

In general, a shift in the firm's demand curve will change the profit maximizing price and quantity. However, since its price exceeds its marginal cost, the firm can afford not to change its price, and its decision comes down to a trade-off between the costs and benefits of doing so. It seems there are costs associated with changing prices. One type of cost is the so-called menu cost. The term comes from the idea that there are costs involved in changing the prices on menus in restaurants. This is a relevant consideration for those firms that operate with posted prices like restaurants, or firms with printed price catalogues. However, these costs seem much less relevant for firms where prices can be changed at the touch of a button. However, even when the technology is available to adjust prices at low cost, a firm may worry that it will lose customers if it changes its price when other firms producing similar products do not. Given that the benefits of changing price are likely to be modest under imperfect competition, the costs do not have to be large to outweigh the benefits.

In the model of pricing we use in defining the price-setting curve, the price is set as a mark up on unit labour costs:

$$P = (1+\mu)\left(\frac{W}{\lambda}\right). \qquad (2.6)$$

This means that the price responds only to changes in labour costs and since productivity, λ, is constant, the price responds only to changes in nominal wages. In this model, firms do not change their prices at all in response to changes in aggregate demand. We can summarize by saying that in the model, prices are sticky in response to shifts in demand and flexible in response to changes in costs, that is, to changes in wages.

A large study has used firm level (i.e. micro) data sources and surveys to investigate price stickiness in the euro area.[11] The authors find that firms in the euro area change their prices infrequently, on average once a year. They also found that the main sources of price stickiness are strategic interactions between competing firms and implicit or explicit contracts with their customers, with menu costs being judged less important. The older survey evidence reported in the overview section noted that firms coordinate their price increases around industry-wide wage rounds.

Business cycles and inflation: the Phillips curve

As we saw in the previous subsection, there is a unique unemployment rate at which the labour market is in equilibrium. At this equilibrium, the *WS* and *PS* curves intersect, which means both wage and price setters are content with the prevailing real wage and have no incentive to alter their behaviour. In this section, we focus on nominal rigidities and how

[10] See Solow (1998) for further discussion of this point.

[11] See Álvarez et al. (2005).

fluctuations in aggregate demand produce business cycle upswings and downswings around the equilibrium rate of unemployment.

To simplify the modelling, we have assumed that labour productivity is constant. This implies that changes in output are reflected by one-for-one in changes in employment. When drawing diagrams, it means that the horizontal shift in output is the same as in employment. In the real world, there is not a one-for-one relationship between changes in output and employment. Okun's law refers to the empirical relationship between a change in aggregate demand, output and the unemployment rate.

Box 2.1 Okun's Law

When output rises, workers who have been kept on the pay-roll but have not been fully employed (e.g. those working shorter than normal hours) may be fully utilized, with the result that higher output does not—at least initially—entail a rise in employment. This is called labour hoarding. Also even if employment rises, unemployment does not necessarily fall if the new jobs are taken by those who were not previously in the labour force. People of working age who are neither employed nor unemployed are called economically inactive and the decision of whether or not to participate in the labour market is dependent on economic conditions.

The combination of labour hoarding and changes in the labour force mean that a 1% change in output growth above or below its trend tends to be associated with respectively a fall or rise in the unemployment rate of less than 0.5 percentage points. This empirical relationship between changes in the growth rate relative to its trend and changes in the unemployment rate is called Okun's Law (an Okun coefficient of −0.5). The responsiveness of unemployment to changes in growth is lower in countries with tighter regulations on hiring and firing (as observed in many continental European countries) and with stronger traditions of lifetime employment (as observed in Japan).

Ball, Leigh and Loungani (2012) find that Okun's Law has been a 'strong and stable relationship' in the major advanced economies since the Second World War. They argue that this relationship did not change substantially in the financial crisis of 2008–09, but that there is a large variation across countries in the coefficient in the relationship (i.e. the degree of responsiveness of the unemployment rate to output). They find an Okun's coefficient of −0.45 for the US, −0.15 for Japan and a much higher coefficient of -0.85 for Spain, where temporary employment contracts are prevalent.

A positive aggregate demand shock increases employment above the equilibrium level, and inflation rises. The timing of events is summarized as before:

Aggregate demand shock → output and employment change

Next wage round → nominal wages change

Immediately after the wage round → prices change

This behaviour is modelled by the Phillips curve. We now need to formalize that relationship to develop the model of inflation and unemployment.

The *WS* curve says that the real wage increases with employment. It simplifies the modelling to express this in terms of the *output gap*, $(y_t - y_e)$, and to write the *WS* curve in linear form:

$$w^{ws}(y_t) = (W/P)^{ws} = B + \alpha(y_t - y_e) + z_w, \qquad \text{(WS curve; linear form)}$$

where B is a constant reflecting the unemployment benefit and the disutility of work, and z_w is the set of wage-push factors. As a result of price setting always restoring the real wage to the equilibrium real wage which is equal to $w_e = B + z_w$, wage setters will attempt to increase the expected real wage by $w^{WS}(y_t) - w_{-1} = w^{WS}(y_t) - w_e = (B + \alpha(y_t - y_e) + z_w) - (B + z_w) = \alpha(y_t - y_e)$. If wage-setters expect prices to increase by $(\Delta P/P)_{t-1}$, we use the approximation $(\Delta W/W)_t - (\Delta P/P)_{t-1} \approx w^{WS}(y_t) - w_{-1}$. Thus we have $(\Delta W/W)_t - (\Delta P/P)_{t-1} = \alpha(y_t - y_e)$ or

$$(\Delta W/W)_t \approx (\Delta P/P)_{t-1} + \alpha(y_t - y_e). \qquad \text{(wage inflation)}$$

This says that wage setters set the percentage increase in the nominal wage to cover the previous period's price increase and to reflect any positive or negative output gap at the time of the wage round.

Turning to price setters, we use the price-setting rule, and noting from the timeline that firms set prices immediately after wages have been set, we have:[12]

$$P = (1 + \mu)\frac{W}{\lambda}, \text{ and}$$

$$(\Delta P/P)_t = (\Delta W/W)_t - (\Delta \lambda/\lambda)_t . \qquad \text{(price inflation)}$$

Substituting the expression for wage inflation into the equation for price inflation gives the following equation for inflation in the simple case where productivity is constant. This is called the Phillips curve:

$$(\Delta P/P)_t = (\Delta P/P)_{t-1} + \alpha(y_t - y_e) \qquad (2.7)$$

$$\underbrace{\pi_t}_{\text{current inflation}} = \underbrace{\pi_{t-1}}_{\text{lagged inflation}} + \alpha\underbrace{(y_t - y_e)}_{\text{output gap}}, \qquad \text{(Phillips curve, PC)}$$

where π_t is the rate of inflation. If P_t is today's price level and P_{t-1} is last period's price level, then the rate of inflation over the past year is π_t:

$$\pi_t \equiv \frac{P_t - P_{t-1}}{P_{t-1}} = \Delta P/P.$$

12 The conversion of the price-setting equation into the price inflation equation is carried out as follows. If you are unfamiliar with logs and their properties, see Section 8.3 in Chapter 8. We assume that μ remains constant over time. In continuous time, we find an expression for the growth rate of $P = (1 + \mu)W/\lambda$ by first taking logs, then differentiating with respect to time and using the fact that $\frac{d \log x}{dx} = \frac{1}{x}$:

$$\log P = \log(1 + \mu) + \log W - \log \lambda$$

$$\frac{d \log P}{dt} = \frac{d \log(1+\mu)}{dt} + \frac{d \log W}{dt} - \frac{d \log \lambda}{dt}.$$

By assumption, $d \log(1 + \mu)/dt = 0$.

Multiply each term by $\frac{dx}{dt}$ and use $d \log x/dx = 1/x$

$$\frac{d \log P}{dP}\frac{dP}{dt} = \frac{d \log W}{dW}\frac{dW}{dt} - \frac{d \log \lambda}{d\lambda}\frac{d\lambda}{dt}$$

$$\frac{dP/dt}{P} = \frac{dW/dt}{W} - \frac{d\lambda/dt}{\lambda}$$

$$\pi = \frac{dW/dt}{W} - \frac{d\lambda/dt}{\lambda}.$$

Writing this in discrete time, gives the expression in the text.

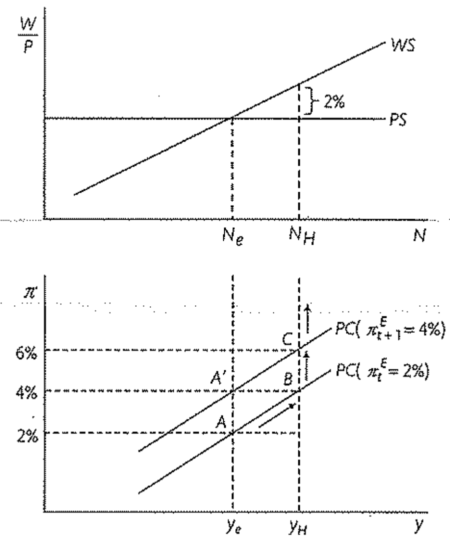


Figure 2.14 The derivation of the Phillips curve.

Graphical derivation of the Phillips curve

In Fig. 2.14, an investment boom because of more optimistic business expectations shifts the *IS* curve to the right. Output and employment rise. The fall in unemployment shifts the balance of power in the labour market towards workers as reflected in the positively sloped *WS* curve. At the next wage round, wage setters respond to the positive output gap: they will need a 2% nominal wage increase to cover last period's inflation plus an additional increase equal to α times the output gap (see the wage inflation equation). In the example shown in Fig. 2.6, this is an extra 2%. Nominal wages go up by 4%.

How do price setters respond to the higher wages? Their labour costs have risen by 4%, so they mark up this cost increase in their prices, and price inflation goes up from 2% to 4% (see the price inflation equation). We have the result that an expansion of aggregate demand that pushes output and employment above the equilibrium level is followed by a rise in wage and price inflation. By joining together points *A* and *B*, we have an upward sloping line in the inflation–output diagram. This is a Phillips curve.

Will inflation stay at 4%? Let us assume that employment remains at its high level until the next wage round. How do wage setters respond? The first thing to notice is that employees will be disappointed with their real wage over the past year: they had expected a high real wage in line with the tighter labour market. This did not eventuate because firms passed on the labour cost increase immediately in higher prices. With this behaviour generalized across the economy, the consumer price index will have risen by 4%. Wage setters will now negotiate a wage increase to make good the erosion of their expected real wage over the past year (an increase of 4%) and they will secure another 2% to take the expected real wage to its high level on the *WS* curve. Firms will follow by putting prices up by 6%. A new Phillips curve

Table 2.1 Constant, rising, and falling inflation.

Period	Output	Inflation (% per year) and employment: Lagged inflation	'Gap'	Wage inflation	Price inflation
−1	y_e	2	0	2	2
			Case 1: constant inflation		
0	y_e	2	0	2	2
1	y_e	2	0	2	2
2	y_e	2	0	2	2
			Case 2: rising inflation		
0	y_H	2	2	4	4
1	y_H	4	2	6	6
2	y_H	6	2	8	8
			Case 3: falling inflation		
0	y_L	2	−2	0	0
1	y_L	0	−2	−2	−2
2	y_L	−2	−2	−4	−4

is defined by joining up point A' and C. Each Phillips curve is labelled by lagged inflation. Case 2 in Table 2.1 shows the output and inflation outcomes following a positive demand shock in period 0.

If we take the converse case of an equal size negative demand shock that moves unemployment above the equilibrium rate and reduces output to $y_L < y_e$, the same reasoning gives the result that inflation is falling. The process will be the exact reverse of that set out above for the positive demand shock. The summary of the output and inflation outcomes for this example is shown as Case 3 in Table 2.1. This example shows that unemployment above the equilibrium is accompanied by falling inflation.

Summary of the Phillips curve (*PC*)

We have shown that the Phillips curve (*PC*) is derived from the wage-setting and price-setting curves. In the formulation of the *PC* that underlies the core 3-equation model, each Phillips curve shows a feasible set of output and inflation pairs for a given rate of lagged inflation. The Phillips curves are pinned down by lagged inflation because of the presence of lagged inflation in the wage inflation equation. Wage setters are interested in the real wage. When setting the nominal wage increase, they take a view about the way the consumer price index is likely to evolve over the course of the wage contract. A simple rule for doing this uses consumer price inflation over the previous period.

In modelling inflation and Phillips curves, it is usual to express the role of lagged inflation in terms of inflation expectations. We use this language and return in Chapter 4 to a detailed investigation of how expectations are formed. Thus, we can write:

$$\pi_t^E = \pi_{t-1}, \qquad \text{(adaptive inflation expectations)}$$

where π_t^E is expected inflation in period t and π_{t-1} is actual inflation in period $t-1$. When we model wage setters' behaviour in wage negotiations in this manner, then we say that they have *adaptive expectations*. This is because they update their expectations every period based on the out-turn for inflation in the last period. As shown in Fig. 2.14, we pin down the vertical height of the Phillips curve by expected inflation. Throughout this book, we will denote the Phillips curve by $PC(\pi_t^E = x\%)$. In the adaptive expectations case, x is simply last period's inflation rate.

The Phillips curves are upward sloping, reflecting the effect of the output gap, $(y_t - y_e)$, on wages and prices through the wage- and price-setting curves, and hence on inflation. As we saw in the previous subsection, if output is above equilibrium, then inflation will be higher than last period's inflation and vice versa. In the model, workers get a base wage increase equal to lagged inflation, and they also get an additional wage change to reflect the position of the economy relative to equilibrium. A positive output gap will result in a positive additional wage change and a negative output gap the opposite. This reflects the positively sloped wage-setting curve.

We can see from Fig. 2.14 that each adaptive expectations Phillips curve is defined by two characteristics:

(1) the lagged inflation rate (π_{t-1}), which fixes the height of the Phillips curve on a vertical line above the level of output associated with the equilibrium rate of unemployment; and

(2) the slope of the *WS* curve, which fixes its slope.[13] The Phillips curves will be steeper if the *WS* curve is steeper and vice versa.

In equation form, the Phillips curve is:

$$\pi_t = \pi_t^E + \alpha(y_t - y_e)$$

$$\underbrace{\pi_t}_{\text{current inflation}} = \underbrace{\pi_{t-1}}_{\text{lagged inflation}} + \alpha\underbrace{(y_t - y_e)}_{\text{output gap}}. \qquad \text{(adaptive expectations Phillips curve)}$$

We can see from this expression that the Phillips curve shifts up or down whenever lagged inflation changes and that its slope depends on α, which in turn reflects the slope of the *WS* curve.

Evidence on inflation dynamics in many countries over the past few decades suggests that changes in output (and employment) are followed by changes in inflation, which is summarized by saying that output leads inflation; and that inflation is persistent.[14] Consistent with this evidence, the Phillips curve states that inflation depends on past inflation, π_{t-1} and the output gap, which reflects the difference between current unemployment and the equilibrium rate of unemployment.

The assumption that prices are adjusted immediately to cost increases means that the real wage remains on the *PS* curve and is constant over the business cycle. Empirically, at the

[13] Note that if the *PS* curve was downward sloping, the slope of the Phillips curve would be steeper, reflecting the slope of both the *WS* and the *PS* curves. At output above equilibrium, for example, there would not only be the 'gap' due to the slope of the *WS* curve but there would be a second 'gap' between the existing real wage and that on the *PS* curve. Firms would push up prices further to reduce the real wage down to the lower real wage on the *PS* curve.

[14] For evidence, see Christiano et al. (2005); also Estrella and Fuhrer (2002) and Muellbauer and Nunziata (2004).

aggregate level, real wages are mildly pro-cyclical. In the model, if lags in price setting are introduced (in addition to the lag in wage-setting due to the annual wage round), then the real wage would lie in between the *WS* and *PS* curves. Real wages would then be pro-cyclical reflecting the upward-sloping wage-setting curve. Real wages would rise in business cycle upswings and fall in downswings.

2.3 Applications

Shocks in the absence of a stabilizing policy maker

We shall see in Chapter 3 that an inflation-targeting central bank will diagnose the nature of a shock and then respond by adjusting the interest rate. To motivate the introduction of a policy maker who seeks to stabilize the economy, we show what happens to the economy when it is disturbed by a demand or supply shock in the absence of such a policy maker. We look in turn at demand shocks, represented in the model by a shift of the *IS* curve or a shift along it due to a change in the interest rate, and at supply shocks, represented by shifts in the *WS* and or the *PS* curves.

An aggregate demand shock

By drawing the *IS* diagram above the labour market diagram we show the positive aggregate demand shock explicitly in Fig. 2.15a (i.e. the left hand panel of Fig. 2.15). Our assumption that there is no stabilizing policy maker is reflected in the fact that the *IS* curve remains at *IS'* following the shock and the real interest rate is kept constant at its initial level, r_e.

We assume the economy begins at equilibrium output y_e with lagged inflation of 2%. A positive aggregate demand shock shifts the *IS* curve to *IS'*. An example of such a shock would be the US economic boom associated with higher government spending (on military and non-military goods and services) that coincided with the start of the Vietnam war in 1965. US GDP growth was 6.4% in 1965 and 6.5% in 1966, compared to an average of just 3.9% for the other years in that decade.[15]

As shown in Fig. 2.15a, the shock is accompanied by ever-increasing inflation: with output above equilibrium at y_H, there is a gap between the *WS* and the *PS* every period, wages and prices are adjusted first as wage setters try to achieve the real wage w_H and second as firms push up prices to restore their profit margin (which implies the real wage is kept at w_e). The process is exactly the same as in the example used to derive the Phillips curve in the last section (see Fig. 2.14 and Case 2 in Table 2.1).

With the real interest rate kept constant at r_e, a positive demand shock is associated with higher employment and rising inflation. In Chapter 3, we address the question of why the policy maker will not be happy with a situation of ever-increasing inflation and what they could do to stabilize the demand shock and get inflation back to the initial level of 2%.

15 Figures calculated using data on real GDP from the US Bureau of Economic Analysis.

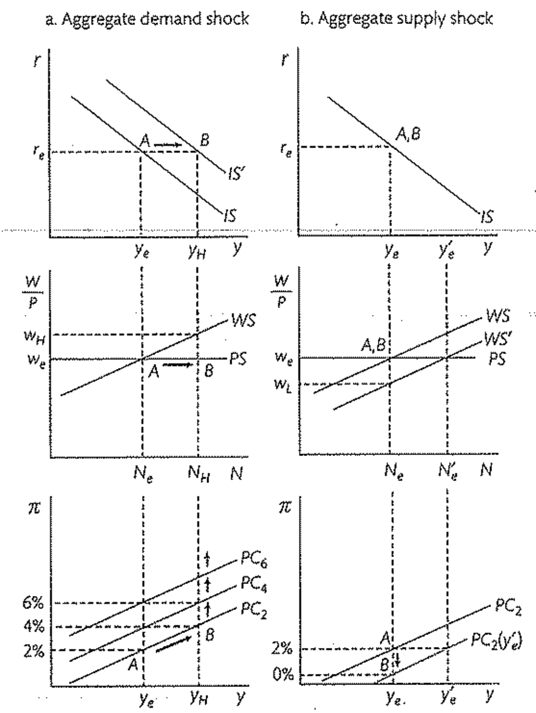


Figure 2.15 World without a stabilizing policy maker: inflationary implications of aggregate demand and supply shocks.

Note: PC_x is used as short form for $PC(\pi_t^E = x\%)$, where *x* is the level of lagged inflation and *t* is the time period.

An aggregate supply shock

We now turn to the implications of an aggregate supply shock. We can model an aggregate supply shock as a shift in

1. the production function, i.e. a technology or productivity shock (change in λ)
2. the *WS* curve, e.g. a shift in bargaining power from workers to employers or in any of the other wage-push factors (i.e. z_w)
3. the *PS* curve, e.g. more intense competition in the product market (i.e. $\downarrow \mu$) or a shift in any of the price-push factors, (i.e. z_p).

To illustrate the implications for output, employment and inflation, we take the example of a downward shift in the *WS* curve. An example of such a shock would be the Dutch supply-side reforms of the 1980s. The Netherlands was seen as one of Europe's 'employment miracles'. In 1983, unemployment stood at 8.3%, whereas the average for the 1990–2008 period was just 4.6%. The Dutch reforms included reduced unemployment benefit generosity, increased expenditure on active labour market policies to make workers more employable and closer coordination in the wage bargaining process.[16]

To illustrate the key differences in the response of the labour market to supply and demand shocks, we analyse the adjustment of the economy to the supply shock using Fig. 2.15b. We assume the economy begins at equilibrium output y_e with lagged inflation of 2%. A positive supply shock shifts the *WS* curve downward to *WS'*. This raises the equilibrium level of employment from N_e to N'_e and the equilibrium level of output from y_e to y'_e. Again, we assume there is no stabilizing policy maker, so the real interest rate is kept constant at its initial level of r_e following the shock.

At the original output level, y_e, there is now a negative gap between the *WS* and *PS* curves. Wage setters respond to the gap by reducing their real wage claim to w_L. They do this because there is now increased competition for jobs (as a result of the active labour market policies) and a higher cost of job loss (as a result of lower benefit generosity). Wages therefore rise by 0% and in order to keep their profit margins constant firms do not change their prices. Inflation falls from 2% to 0% and the Phillips curve shifts down.

In the following periods, the adjustment of the economy is similar to that of a negative demand shock (see Case 3 in Table 2.1). Inflation falls in each period. This process will carry on indefinitely, as long as output is kept below the new equilibrium—i.e. unemployment cannot be kept at a level above equilibrium without falling inflation.

A positive supply shock is defined as one that raises equilibrium output and employment. We have shown that a positive supply shock is associated with falling inflation *at the initial output level* $y = y_e$. If the supply shock is permanent, then the economy is capable of operating with lower unemployment and constant inflation. As we shall see in Chapter 3, the policy maker is likely to respond to the shock by reducing the interest rate to allow the economy to operate at the new equilibrium output level of y'_e with constant inflation.

2.4 Conclusions

This chapter has provided the second building block for the model of the macro-economy. We have set out the *WS* − *PS* model, which is used to determine the equilibrium level of output in the economy. From the *WS* − *PS* model, we derived the Phillips curve, or *PC*, which is used to model wage and price inflation and is one of three equations that underpins the 3-equation model of the macro-economy that will be set out in the next chapter.

The *WS* − *PS* model shows that wages at equilibrium unemployment are higher than workers would be willing to accept; this means there is involuntary unemployment. With monopolistic (or equivalently imperfect) competition in goods markets, firms charge a mark-up on their goods and make supernormal profits. At the supply-side equilibrium, the rate of unemployment is such that both wage and price setters are content with the prevailing real wage. The real wage is constant, which implies that wages and prices are rising at the same rate: inflation is constant. Inflation could be constant at a rate of zero, in which case nominal

[16] See Nickell and Van Ours (2000a).

wages and prices would remain unchanged. Modern economies are typically characterized by positive inflation.

By combining the analysis of aggregate demand summarized in the *IS* curve with the supply side (*WS* and *PS* curves), we can analyse current and historical episodes of unemployment. For example, we can model both the rapid growth of unemployment following the global financial crisis and the gradual increases in equilibrium unemployment that characterized many European economies from the 1970s. In the first case, there was a large negative aggregate demand shock, which pushed economies away from their constant inflation equilibrium unemployment rates and increased involuntary unemployment. In the second case, negative supply shocks pushed up equilibrium unemployment, leaving it higher, even when stable inflation was restored.

Given the economy's supply side characteristics, the Phillips curve shows the feasible set of inflation and output pairs for a given rate of lagged inflation. We can interpret the lagged inflation term either as agents forming expectations about how prices will evolve in a backward-looking manner or as agents negotiating wage increases to compensate them for the erosion of their real wages associated with last period's inflation. We shall show in the next chapter that the Phillips curve acts as a constraint on the policy maker, limiting their choice of viable inflation-output combinations in each period.

The modelling undertaken in Sections 2.2 and 2.3 has provided a framework with which we can shed light on important questions concerning the supply side of the economy. As a summary of the chapter, we return to each of the key questions posed in these sections:

1. **Supply side effects on equilibrium unemployment.** What determines the medium-run equilibrium level of output and unemployment around which the economy fluctuates? The equilibrium rate of unemployment is pinned down by the intersection of the *WS* and *PS* curves. Structural features of the economy, such as unionization, along with policy choices, such as the generosity of unemployment benefits and the stringency of competition policy, will affect equilibrium unemployment. Swings in demand will move the economy away from equilibrium, resulting in actual unemployment deviating from equilibrium (see Fig. 2.8). Unemployment below equilibrium is associated with output above equilibrium, i.e. a positive output gap ($y > y_e$). In Chapter 15, we shall consider possible feedback from actual to equilibrium unemployment, which is referred to as 'hysteresis'.

2. **Nominal rigidities, inflation and the business cycle.** What happens to inflation when the economy is away from equilibrium? When output is above equilibrium, the increased tightness of the labour market is reflected in a higher wage-setting wage. There is a lower cost of job loss and workers will need to expect higher real wages in order to exert effort. These are expected real wages because the real wage outturn depends on what happens to inflation in the economy as a whole over the period of the wage contract. In this case $WS > PS$, and workers will get a wage increase that more than compensates them for the erosion in real wage due to last period's inflation. However, firms protect their profit margins and will immediately raise prices following the wage round. This means that workers' real wage expectations are constantly frustrated when output is above equilibrium. The only way in which output can remain above equilibrium is with rising inflation, as shown in Fig. 2.14. The opposite is true for output below equilibrium. The relationship between output and inflation can be summed up by the Phillips curve: $\pi_t = \pi_{t-1} + \alpha(y_t - y_e)$. This equation shows that

the current period's inflation (π_t) depends positively on both last period's inflation (π_{t-1}) and the output gap ($y_t - y_e$).

3. **Demand shocks, supply shocks and the Phillips curve.** How does the economy respond to demand or supply shocks in a world without a policy maker who aims to stabilize the economy? The economy reacts differently to supply and demand shocks. We can highlight these differences by seeing how they are captured by different parts of the model. Demand shocks affect the *IS* curve and supply shocks affect the *WS* or *PS* curves and consequently the Phillips curve, *PC*. A positive demand shock leads to ever-increasing inflation as long as demand is kept at a level where there is a positive output gap. In contrast, a positive supply shock increases equilibrium output and is associated with falling inflation as long as output is kept at the initial equilibrium rate of output. Falling inflation signals that the economy is capable of operating at a higher equilibrium level of output.

Having a framework of the kind set out in Chapters 1 and 2 is essential in understanding how policy makers assess the state of the economy and whether intervention is required. This sets the stage for the introduction of the policy maker with stabilization objectives in the next chapter. In Chapter 3, we show why governments introduced inflation-targeting central banks in a bid to improve economic performance.

2.5 Appendix

2.5.1 The textbook model: competitive markets and complete contracts

To help bring out the way the labour market works and to clarify the notion of involuntary unemployment, it is useful to provide a comparison with a simple textbook model. In the textbook model, the intrinsic feature of the labour market of variable effort is neglected. It is assumed that the wage buys a specific amount of work.

What would a market for labour mean when employment contracts are complete? The image to have in mind is not of the labour market in a real economy but of a hiring hall where employers and workers gather. In the hall, employers shout out job offers and workers respond. Since the textbook model assumes that what the employer gets in exchange for a particular hourly wage offered is unambiguous, he will be happy to hire labour on this basis. In the hiring hall, a wage is posted and workers respond by accepting or declining to work.

As discussed in Section 2.1, we assume the labour supply curve to be quite inelastic. This reflects the offsetting income and substitution effects of a wage change, as well as the effects of changes in labour force participation. Imagine the wage offers are being broadcast outside the hiring hall and people are drawn in to participate in the labour market by higher wage offers.

Turning to the labour demand side, in the textbook model, there is a downward-sloping labour demand curve, which shows the labour demanded at a given real wage. Assuming for simplicity that the capital stock (K) is fixed, output is a positive function of the level of employment (N), i.e.

$$y = f(N; K),$$

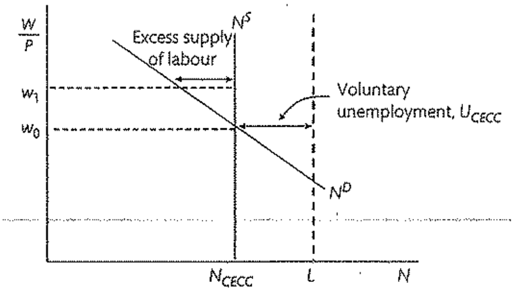


Figure 2.16 Market clearing equilibrium in a textbook labour market with complete employment contracts.

where f is the production function. It is assumed that the production function is characterized by diminishing returns, which means that as more workers are employed—holding capital constant—the extra output produced by each additional worker declines. Another way of putting this is that the marginal product of labour, which can be written as $\frac{\partial y}{\partial N}$ or *MPL*, declines as employment rises. The labour demand curve (N^D) is often referred to as the *MPL* curve since under perfect competition, firms take the real wage as given and employ labour up to the point at which the marginal product of labour is equal to the exogenously given real wage (the wage shouted out in the hiring hall).

The market clears at the real wage where the labour demand curve intersects the labour supply curve. Bringing together the labour demand and labour supply schedules as in Fig. 2.16, the labour market clears with the real wage w_0 and employment, N_{CECC} (for Competitive Equilibrium Complete Contracts).

Any temporary displacement of the economy from equilibrium is assumed to be eliminated by a movement in real wages. For example, if the real wage rises above the market-clearing level (due say, to an unexpected fall in the price level), then labour supply exceeds labour demand at that real wage (w_1). The excess supply of labour will result in falling nominal wages until the unique competitive equilibrium is re-established with the real wage at w_0 and employment at N_{CECC}. In this model, the only people who will be unemployed will be those *voluntarily* unemployed in the sense that at the going real wage, they prefer searching for a job or leisure over the goods and services obtainable through working.

In Fig. 2.16, the rate of unemployment in the market-clearing equilibrium is U_{CECC}/L, where L is the labour force, i.e. the sum of the employed and the unemployed. It is important to remember that in the textbook model, the economy is at a welfare optimum in the market-clearing equilibrium, so the voluntary unemployment that exists does not signify a problem. Rather, it reflects the choice by workers about whether and how much to work at the existing real wage. It is difficult to motivate the study of supply-side institutions and policies and of demand-side (that is, stabilization) policy in this model.

2.5.2 The mark-up and the elasticity of demand

A firm maximises profits when marginal revenue (MR) = marginal cost (MC). With labour as the single factor of production this means that

$$MC = \frac{W}{MPL}, \qquad \text{(marginal cost)}$$

where W is the nominal wage and MPL the marginal product of labour.

Total revenue is

$$R = Py, \qquad \text{(total revenue)}$$

where P is the price for which the firm sells their output, and y is the quantity of output they sell. The firm faces a standard downward-sloping demand curve. The elasticity of this demand curve, η, measures the responsiveness of output demanded to a change in the price. If there is perfect competition, the demand curve is flat and $\eta = \infty$, since under perfect competition, a marginal increase in the price will mean the firm loses all of its sales. The less elastic demand (in absolute value) is the demand curve, the less demand the firm loses by increasing its price.

Marginal revenue is

$$MR = \frac{dR}{dy} = P + y\frac{dP}{dy}. \qquad \text{(marginal revenue)}$$

If we divide each side of the equation for marginal revenue by P and substitute in the definition of η, where

$$\eta = -\frac{dy}{dP}\frac{P}{y}, \qquad \text{(elasticity of demand)}$$

we get

$$MR = P + y\frac{dP}{dy}$$

$$\frac{MR}{P} = 1 - \frac{1}{\eta}$$

$$\rightarrow MR = P\left(1 - \frac{1}{\eta}\right).$$

Then at the profit-maximizing optimum, where $MR = MC$, we have

$$P\left(1 - \frac{1}{\eta}\right) = \frac{W}{MPL}$$

or

$$P = \left(\frac{\eta}{\eta - 1}\right)\frac{W}{MPL}. \qquad \text{(pricing formula, in terms of elasticity, } \eta)$$

This can be written in terms of a mark-up μ on marginal labour costs

$$P = (1 + \mu)\frac{W}{MPL}, \qquad \text{(pricing formula, in terms of the markup, } \mu)$$

where $\mu = \frac{1}{\eta - 1}$.

In the case of perfect competition $\eta = \infty$ and $\mu = 0$. As the demand curve becomes more inelastic, the mark-up, μ, rises. If, for example, $\eta = 6$, the firm will mark up marginal costs by 20%; $\mu = \frac{1}{\eta-1} = \frac{1}{6-1} = 0.2$.

2.5.3 Deriving the *PS* curve including the tax wedge

The wage element of costs for firms is the full cost of labour to firms—i.e. the gross wage paid to the worker (which includes the income tax and social security payments that have to be made by the worker) plus the employer's social security contributions. All direct taxes are summarized in the tax rate, t_d. This is shown in the pricing equation, where P is the producer price:

$$P = (1 + \mu)\frac{W^{gross}}{\lambda} = (1 + \mu)\frac{W(1 + t_d)}{\lambda}. \qquad \text{(price equation)}$$

But to get the *PS* curve we need the price-setting equation in term of the consumption price, $P_c = (1 + t_v)P$, so that:

$$P_c = (1 + t_v)P = (1 + t_v)(1 + \mu)\frac{W(1 + t_d)}{\lambda},$$

which implies

$$\frac{W}{P_c} = \frac{1}{(1 + \mu)}\frac{\lambda}{(1 + t_d)(1 + t_v)}$$

and using our approximation that $\frac{1}{1+\mu} \approx 1 - \mu$, we have:

$$w^{PS} = \frac{W}{P_c} = \frac{\lambda(1 - \mu)}{(1 + t_d)(1 + t_v)}. \qquad \text{(PS equation including tax wedge)}$$

An approximation to the tax wedge is given by: $\frac{1}{(1+t_d)(1+t_v)} \approx (1 - t_d - t_v)$. We can see that:

1. Any fall in the wedge, for example, a fall in income tax, implies an upward shift in the *PS* curve, indicating that the real wage is higher at any level of employment since the tax take is smaller.
2. The smaller wedge means that *WS* and *PS* curves cross at higher employment: there is a lower equilibrium rate of unemployment because a higher real consumption wage on the *WS* curve is consistent with equilibrium for price setters (on the new higher *PS* curve).

2.6 Questions

2.6.1 Checklist questions

1. Explain in words and using a diagram what is meant by labour market clearing (refer to Appendix 2.5.1). Why is labour market clearing not observed in real-world labour markets?
2. Why is the wage-setting curve upward sloping? If there is a disutility of working, why are workers unhappy when they are unemployed?
3. Derive the price-setting curve. What does the equation for the *PS* curve tell us about the ability of firms to make supernormal profits? Explain in words why the decisions of firms about what price to set has implications for the real wage in the economy. Provide two different explanations for why the *PS* curve might be flat.
4. What is being assumed about the timing of wage setting and price setting that enables us to say that the economy is always on the *PS* curve but only on the *WS* curve in a medium-run

equilibrium? What timing assumptions would deliver the result that the economy is always on the *WS* curve but only on the *PS* curve in a medium-run equilibrium?

5. Explain in words the inflationary consequences of an upswing in aggregate demand. Assume the economy is initially at equilibrium and make sure you adequately explain the transmission mechanisms as well as the final result.
6. Evaluate the following statement:
   'When the economy is in equilibrium in the *WS* – *PS* model, there is only voluntary unemployment, because no agent has an incentive to change their behaviour.'
7. Equation 2.2 in Section 2.2.1 shows that firms set prices as a mark up over unit labour costs. Use this equation to show the relationship between wage growth and price growth in the economy. What assumptions do we have to make for wage changes to translate one-for-one into price changes? In this case, how will a rise in wages impact on a firm's profit margin?
8. Assess the following statements S1 and S2. Are they both true, both false or is only one true? Justify your answer.
   S1. When output is above equilibrium, wage setters will secure wage increases that reflect the tightness of the labour market.
   S2. If output is consistently above equilibrium, then wage-setters' real wage expectations are constantly frustrated.
9. Use the *WS* – *PS* model to graphically derive a set of Phillips curves. Explain the intuition behind the diagram. Provide an explanation of how a situation of deflation could occur (deflation is a situation in which inflation is negative so that prices are falling). What are you assuming about the real interest rate?
10. Giving in each case an example of what could have caused it, explain how unemployment can be above equilibrium due to the following. Use diagrams, such as those in Fig. 2.15, to show the inflationary implications of each shock (assume there is no stabilizing policy maker).
    (a) Aggregate demand being too low
    (b) A wage-setting shock

2.6.2 Problems and questions for discussion

1. Collect an inflation report from the central bank website of an OECD economy (e.g. the Bank of England, the ECB, the Federal Reserve etc.) and answer the following questions:
   (a) What factors do they view as important drivers of inflation (or the outlook for inflation)?
   (b) Divide these factors into supply-side and demand-side factors.
   (c) For each factor, provide a plausible explanation for which curve of the model they affect (e.g. *IS*, *WS* or *PS*).
   (d) For each factor, in which direction would you expect the affected curve to shift?
2. Read the following statement and then answer the questions below: 'Just by looking at real-world labour markets, it is obvious they are far too complex to be accurately modelled by either the efficiency wage model or the textbook model with complete contracts'. [Hint: the textbook model is set out in the Appendix.]
   (a) Does the statement provide a good justification for not using a model when thinking about the supply side of the economy?

   (b) What are the key predictions of the efficiency wage model and the textbook model with complete contracts?
   (c) The predictions of which model more closely match what we observe by analysing real-world data? Provide examples to justify your answer.
3. Consider two key events in recent economic history; the credit crunch that plunged the world into recession in 2008 and the Fukushima nuclear disaster that struck Japan in 2011. Find data on the path of the monthly harmonized unemployment rate from the start of 2007 onwards for both the US and Japan using OECD.Stat. Answer the following questions:
   (a) Describe (i) how US unemployment reacted to the credit crunch; and (ii) how Japanese unemployment reacted to the Fukushima nuclear disaster.
   (b) Using the material in this chapter and the previous chapter, make an assessment of which curves were likely to have been affected by these economic shocks—i.e. was it the *IS*, *WS* or *PS* curve that were affected (or some combination of them)? Make sure you justify your answer.
   (c) Given your answer to part (b), do you predict that the equilibrium level of unemployment shifted in response to these shocks? Justify your answers.
4. Collect data for GDP and the unemployment rate for OECD economies between 2007 and 2010 from OECD.Stat. Produce a scatterplot of percentage point changes in GDP and the unemployment rate (from peak to trough) over the course of the Great Recession. Answer the following questions:
   (a) Describe briefly what the graph shows.
   (b) Does the data support Okun's Law (i.e. an Okun coefficient of −0.5)?
   (c) If not, then suggest some factors that might account for the differences in Okun coefficients observed across the OECD economies.

3 The 3-equation model and macroeconomic policy

3.1 Overview

In Chapter 1, we focused on the way shocks to aggregate demand produce fluctuations in output and employment. In Chapter 2, we looked at how supply-side features of the economy determine the constant inflation equilibrium. Then, at the end of Chapter 2, we put these supply- and demand-side elements together. This produced a picture of an economy affected by both demand and supply-side shocks, which are amplified by the multiplier process and dampened by households' and firms' ability to tide themselves over through borrowing when their current income falls. In contrast to the generally slow-moving shifts in equilibrium unemployment, actual unemployment rises and falls in accordance with business cycles largely driven by fluctuations in aggregate demand.

In the model as laid out in Chapter 2, there was no policy intervention to prevent the gap between the real wage workers expect to secure (on the *WS* curve) and the real wage firms are prepared to pay (on the *PS* curve) from producing ever-increasing inflation when unemployment is lower than the equilibrium rate. This incompatibility between the claims of workers and firms—and hence the inflationary pressure in the economy—could in principle be resolved in two ways:

1. First, by supply-side institutional change or policies to raise equilibrium output by either shifting the *WS* curve downwards or shifting the *PS* curve upwards. Examples of this include policies to weaken the bargaining power of workers by worsening the living conditions they would face if they lost their jobs (brought about, for instance, by lowering unemployment benefits). Another example is an institutional change which increases bargaining restraint by unions. One policy which might shift the *PS* curve upwards would be the promotion of more competition, thus reducing the mark-up on product prices.
2. Second, by the use of demand management policy to reduce employment and output to the level consistent with a *WS* – *PS* equilibrium, where inflation will be constant. This policy lever takes supply-side institutions and policies as given.

In this chapter, the focus is on the second set of policies mentioned above. This means that we do not deal here with supply-side policies that aim to influence equilibrium

unemployment (that is postponed to Chapter 15, which can be read immediately after this chapter). Here, we begin with unemployment at its equilibrium rate (the *WS* – *PS* intersection), where the output gap is zero and inflation is constant. If, at this point, the economy suffers a shock, how would the policy maker respond so as to stabilize it at its constant-inflation level?

Shortly, we will look in more detail at how demand management policy is implemented to stabilize the economy (for example by raising aggregate demand in the face of a negative demand shock or dampening it in the face of a positive demand shock). But, before that, we need a little more background information about the policy regime and its recent history. During the 1990s, many countries around the world adopted a new macroeconomic policy regime known as inflation targeting. It means that the government gives responsibility to the central bank to stabilize the economy by adopting a specific target for the annual rate of inflation, the most commonly adopted rate being 2%. To help clarify the reasons why these changes in regime and policy have taken place it would be useful to answer three related background questions:

1. Why do governments often define the central bank's responsibility as the control of inflation?
2. Why, in any case do governments regard inflation as a problem which needs to be controlled?
3. Why do governments hand over the problem of keeping inflation low and stable to the makers of monetary policy (the central banks) rather than controlling inflation themselves?

Regarding the first of these questions, we have already seen that the rate of unemployment at which inflation is constant depends on structural features of the labour market and the goods market. It is these structural features which determine the positions of the *WS* and *PS* curves in the supply-side model. But, since the central bank does not have policy instruments with which it can change these supply-side economic features, it is clearly not the right policy maker to be charged with reducing equilibrium unemployment. The government will have to do this, in ways which are discussed in Chapter 15.

We shall see, however, that there is one important way in which the central bank's behaviour does affect unemployment. A central bank that is given responsibility for maintaining high employment as well as keeping inflation close to its target level would be less inclined to support policies of rapidly cutting inflation through pushing unemployment up sharply. In Chapter 13, we return to the question of how variation in the behaviour of different central banks around the world can be represented in the model.

The second preliminary question is why policy makers are concerned about inflation at all. One reason why governments might pursue low inflation is that the level of inflation is a key priority for voters. Governments might have to pay a large electoral price if they fail to control inflation. The importance the US general public gives to inflation is highlighted in a 1997 paper by Shiller, who surveys both economists and members of the public in the US about their views on inflation. Table 3.1 shows that the general public (more than economists) see controlling inflation as a national priority. We shall look in more detail at the factors that lie behind voters' concerns with inflation in the next subsection.

Table 3.1 Public attitudes on the importance of preventing high inflation.

Do you agree that preventing high inflation is an important national priority, as important as preventing drug abuse or deterioration in the quality of our schools?

	Fully agree		Undecided		Completely disagree	
US all	52%	32%	4%	8%	4%	n=117
Economists	18%	28%	11%	26%	18%	n=80

Source: Survey data from Shiller (1997).

This leaves the third preliminary question: why, in the inflation-targeting regime, is stabilization policy placed in the hands of the central bank rather than the government? This question is best broken down into two separate parts.

The first part is: why is monetary policy chosen over fiscal policy as the primary tool for stabilization policy? There are both practical and political economy reasons why monetary policy is the preferred stabilizer even when both monetary and fiscal policy are available. Firstly, changing public expenditure or taxation normally involve lengthy parliamentary processes, and there is no equivalent to the gradual adjustment possible through quarter-point changes in the interest rate at monthly intervals. Secondly, fiscal policy is inherently political since it involves the use of tax revenue, a fact which is captured by the classic refrain of struggles for democracy: 'No taxation without representation'. Monetary policy is viewed as more neutral and does not so obviously create winners and losers, making it a less contentious policy instrument for use in short-run demand management.

The second part is: why does the government cede control of monetary policy to an independent central bank? The government can potentially gain an electoral advantage by controlling monetary policy. For example, they could raise output above the equilibrium or change politically sensitive mortgage interest rates in the run-up to an election. In light of this, why have governments around the world chosen to delegate control of monetary policy to a committee of independent experts? It is because governments lack low inflation credibility, which translates into higher inflation. The political pressures to manipulate interest rates for electoral gain are the source of this credibility deficit. Independent central banks are free from these pressures so are more likely to deliver low and stable inflation. To the extent that voters care about inflation and sound macroeconomic management, this provides the incentive for the government to relinquish control of this policy lever. These arguments have been persuasive among both left- and right-wing policy makers, as illustrated by the fact that one of the very first acts of the British Labour Party when it came to power in 1997 was to make the Bank of England independent.

The discussion in Chapter 1 brought out the extent to which households are able to limit the costs to them of fluctuations in income that arise from shocks to the economy through smoothing their consumption. Nevertheless volatility remains and the policy maker aims to provide additional stabilization. We shall see in the course of this chapter how the central bank responds to a range of shocks: its interest rate decisions affect aggregate demand and the output gap, guiding the economy back to the medium-run equilibrium with low steady inflation as its objective.

3.1.1 The role of the central bank in stabilization

Since the early 1990s a growing number of central banks have been assigned the task of stabilizing the macro-economy around a low target rate of inflation. This is known as an inflation-targeting monetary policy regime. In response to a surge in inflation, we would expect to see the central bank raise the interest rate; this would lower interest-sensitive spending, such as spending on housing, machinery and equipment, due to the increase in the interest rate (in other words, in the cost of borrowing). The cut back in investment would lead in turn to lower aggregate demand and a fall in output. Unemployment would go up and inflationary pressure would diminish.

In 1997, the year in which the Bank of England was made independent, the Bank's Chief Economist, Mervyn King, who later became Governor, explained:[1]

> [M]onetary policy can be described in terms of two policy variables—a medium-term inflation target and a response of interest rates to shocks that create fluctuations in inflation and output. The overriding objective of monetary policy is to ensure that on average inflation is equal to the target. But such a target is not sufficient to define policy. There is a subordinate decision on how to respond to shocks as they occur.

Figure 3.1 provides a schematic overview of the macro model, including the policy maker. On the left hand side, the role of the demand and supply sides are summarized. The

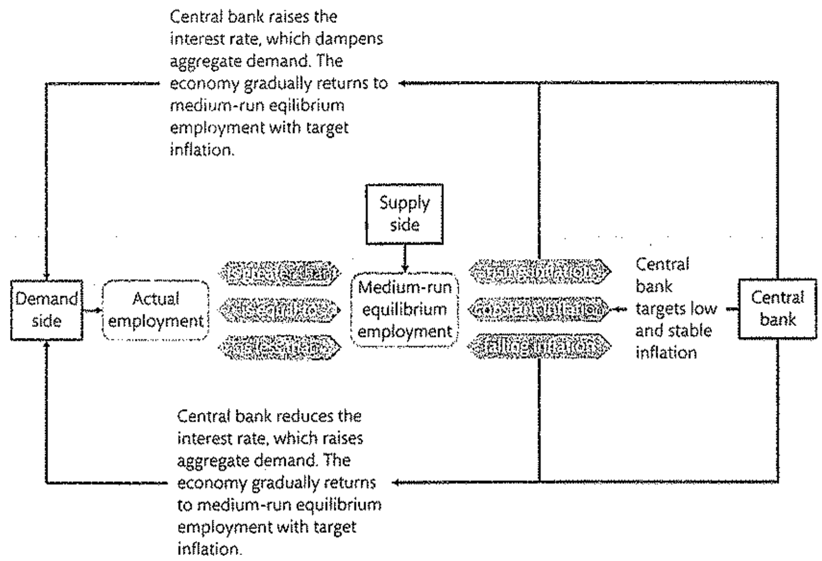


Figure 3.1 Schematic overview of the macro model.

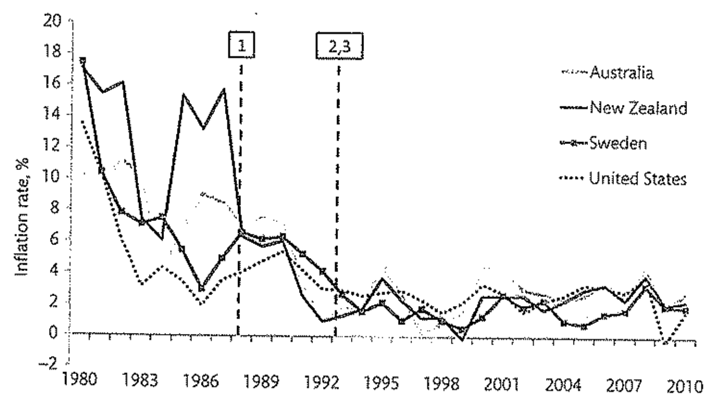
[1] Excerpt taken from: Mervyn King, 29th October 1997, *The Inflation Target 5 Years On*, speech delivered at the LSE. He became the Governor of the Bank of England in 2003.

components of aggregate demand determine the level of output and employment in the goods market equilibrium. The structural features of the economy on the supply side determine the equilibrium rate of unemployment at which inflation is constant. The third panel shows the implications for inflation: if unemployment is at the equilibrium rate, then inflation will remain constant and the central bank will do nothing. If unemployment is below the equilibrium rate, inflation is rising and vice versa for unemployment above it. An increase in inflation above the central bank's target will trigger intervention by the central bank: it will raise the interest rate. Equivalently, a fall in inflation below target will trigger an easing of monetary policy with a lower interest rate.

The outside arrows depict the feedback from the central bank's monetary policy rule to aggregate demand. This illustrates how a central bank uses a monetary policy rule to hold inflation close to a low target value.

Figure 3.2 shows the course of inflation in a selection of developed economies since 1980. The chart shows that a period of high and volatile inflation was followed by one where inflation was lower and more stable. The introduction of inflation targeting regimes was part of the evolution of the policy framework, where from the end of the 1970s, governments used macroeconomic policy to squeeze inflation out of the economy by depressing aggregate demand.

The shift from high to low inflation shown in the chart began before the large fall in the oil price in 1986 and the gradual emergence of low-cost (especially Chinese) manufacturers into world markets, which subsequently also had a significant effect on inflation. Although formal inflation targeting was not adopted by the US Federal Reserve until very recently,



Dates inflation targeting was first adopted:
1. Reserve Bank of New Zealand–April 1988
2. Sveriges Riksbank (the Swedish central bank)–January 1993
3. Reserve Bank of Australia–March 1993
Note: The US had no formal inflation target over this period, but did adopt a target of 2% in January 2012.

Figure 3.2 Inflation rates—before and after the adoption of inflation targeting: 1980–2010.
Source: IMF World Economic Outlook, September 2011.

its inflation behaviour was similar to the other countries shown in Figure 3.2, where its date of introduction in New Zealand, Australia and Sweden is shown. The move towards formal or informal inflation targeting is frequently identified with the beginning of the 'Great Moderation' (the period of unprecedented macroeconomic stability shown in Figure 1.3 in Chapter 1).

There are, however, two very important situations in which the central bank cannot be relied on to stabilize shocks by changing the interest rate as described above. The first of these situations is when the economy experiences such a large negative demand shock that even when the central bank cuts the nominal interest rate all the way to zero this is not low enough to stimulate a revival of aggregate demand. With a nominal interest rate of zero, the real interest rate, which is the one relevant to the investment decisions of firms, is not low enough to stimulate interest-sensitive spending. This problem of the *zero lower bound* is closely related to the possibility that the economy can become caught in a *deflation trap*, where it enters a vicious cycle of falling output and inflation. Once the usual channel of monetary stabilization is shut down, the central bank can try to stimulate the economy using other policies, such as quantitative easing, or it can share responsibility for stabilization with fiscal policy. The cases of the zero lower bound and the role of fiscal policy as stabilizer are addressed later in this chapter. We leave a discussion of quantitative easing until Chapter 7.

The second case in which monetary policy cannot be used to achieve stabilization is when the economy has a fixed exchange rate. This is easy to see in the setting of a common currency area such as the Eurozone: member countries of the Eurozone all use the euro as their currency and they share a single central bank, the European Central Bank (ECB). As a result, these countries do not have their own monetary policy maker which could stabilize shocks that are specific to their country. Because all Eurozone countries have the same monetary policy, any stabilization of shocks to a particular country would have to be achieved by the use of fiscal policy (a case explained in Chapter 12).

3.1.2 Inflation and deflation

Rising inflation and distributional conflict

Rising inflation reflects distributional conflict as different social groups (wage setters/employees and price setters/employers) seek to protect their interests. In an economy in which groups such as unions can influence nominal wages and where firms have price-setting power, a situation of rising inflation reflects inconsistent claims by these groups on the output per head produced in the economy. In other words, when unemployment is below the equilibrium rate, the gap between the real wage on the *WS* curve and on the *PS* curve reflects distributional conflict and causes inflation to rise. As shown in Chapter 2, as unemployment falls, workers become more powerful in the labour market and are able to secure a higher expected real wage: this is the reason the *WS* curve is upward-sloping. But firms will seek to secure their profit margin by setting prices accordingly—leaving a gap between the higher real wage workers expect on the *WS* curve and the unchanged real wage on the *PS* curve.

A similar logic applies with efficiency wage-setting: the conflict of interest between the two sides of the labour market means that the wage employers need to pay to get workers

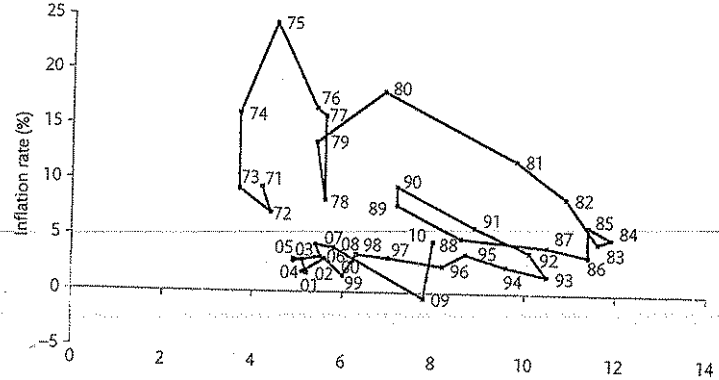


Figure 3.3 UK inflation and unemployment: 1971–2010.

Source: UK Office for National Statistics (data accessed February 2012).

to work sufficiently hard (given the state of the labour market) may differ from the wage that is consistent with the profit margin firms seek. Lower unemployment reduces the cost of job loss and raises the efficiency wage. This pushes inflation up.

Rising inflation produces social tension as frustration mounts. As we shall see, inflationary episodes of this kind have frequently been followed by painful periods of disinflation, where a period of output below equilibrium (accompanied by high unemployment) is required to bring inflation down.[2]

The empirical evidence suggests that periods of disinflation to bring down inflation from moderate rates (up to double digit rates per annum) have involved significant costs in OECD economies where this has occurred. Figure 3.3 shows the path of inflation and unemployment in the UK between the early 1970s and 2010. The high level of inflation in the 1970s was tackled by tight monetary policy from 1979. The process of getting inflation down from a higher level to the targeted one is known as disinflation. It is likely to be a costly process, involving a period of high unemployment to squeeze inflation out of the system. The costs of high unemployment include the direct waste of resources, the erosion of skills and increased vulnerability to mental illness that comes with long spells of unemployment, and the damage to the functioning of families and communities. Figure 3.3 shows that reducing inflation from rates above 20 to below 5% was associated in the British economy with a rise in unemployment from 4 to 12% during the disinflationary episode.

The benefits of low and stable inflation

The survey on public attitudes to inflation summarized in Table 3.2 shows that the main reason the US public dislike inflation is due to the erosion of living standards that are believed

[2] An early formal model of equilibrium unemployment and of the inflation process as one of conflicting distributional claims is in Rowthorn (1977).

to accompany high inflation. We can relate this to an aspect of the dynamics of the *WS-PS* model. The US public's biggest complaint about inflation accords with the outcome of the *WS-PS* model in an inflationary environment, in the sense that rising inflation leads to the persistent frustration of workers' real wage expectations.

In the model in Chapter 2, we used the simplest case regarding the timing of wage and price setting. If firms are able to adjust prices immediately after wages have been set, rising inflation reflects a situation in which the real wage wage setters require is systematically frustrated (in other words, the expected real wage of workers as indicated by the *WS* curve is above the real wage on the *PS* curve). In this simple case, the real wage outcome is on the *PS* curve, not on the *WS* curve. If there are lags in price setting as well as in wage setting, then the real wage lies between the *PS* and *WS* curves and neither wage nor price setters are fully satisfied. This case is more realistic because the empirical evidence suggests that real wages are mildly pro-cyclical, which means that when there is a positive output gap the real wage ends up between the *WS* curve and the *PS* curve (that is, higher than at the constant inflation equilibrium) and vice versa for a negative output gap.

Table 3.2 reveals a divergence between what the public and the economists view as the most significant costs of inflation. Whereas the public appear to worry about their real incomes being eroded by inflation, economists do not see this as much of a concern and place more stress on the interference that high inflation creates for the ability of prices to convey information. Periods of high inflation are also often ones with volatile inflation. To see why this matters, consider an economy with technical progress in which innovation takes place unevenly across sectors. In sectors with rapid innovation, prices will be falling relative to prices in other sectors where technology is more stagnant (think of the falling prices of computing power). These are economically significant changes in *relative prices* and should lead to a reallocation of resources in the economy. Volatile inflation can therefore distort resource allocation to the extent that it masks these relative price changes. In short, volatile inflation has adverse real effects on the economy that are hard to avoid.

When inflation is high, people also want to hold less money. This is because inflation acts as a tax on holding money balances, eroding their real value over time. The so-called inflation tax imposes inefficiencies because it distorts behaviour: people spend more time managing their financial assets incurring what are referred to as 'shoe-leather costs'. Other

Table 3.2 What is peoples' biggest gripe with inflation?

Which of the following comes closest to your biggest gripe about inflation:
1. Inflation causes a lot of inconveniences: I find it harder to comparison shop, I feel I have to avoid holding too much cash, etc.
2. Inflation hurts my real buying power, it makes me poorer.
3. Other.

	1	2	3	
US all	7%	77%	15%	n=110
Economists	49%	12%	40%	n=78

Source: Survey data from Shiller (1997).

firms incur costs as a consequence of the need for frequent price changes costs ('menu costs').

The optimal rate of inflation

Given the costs associated with high inflation can we infer that the optimal rate of inflation is zero or even negative? In thinking about the optimal inflation rate, think about the following: the return on holding notes and coins is zero so with any positive inflation rate, the return in real terms, i.e. after controlling for inflation, is negative. The negative real return leads people to waste effort economizing on their money holdings (shoe leather costs again) and this is inefficient given that it is virtually costless to produce money.

If we follow the logic of this argument, then with a positive real rate of interest, for the *nominal* interest rate to be zero, inflation would have to be negative (i.e. prices falling, which is called deflation). This was Milton Friedman's view of the optimal rate of inflation: the rate of deflation should equal the real rate of interest, leaving the nominal interest rate equal to zero.[3] Friedman's rule ensures that people avoid shoe leather costs, but is this sufficient to say that deflation is optimal?

The danger of deflation

An important reason why central banks target a low but positive inflation rate—most commonly 2%—is that they wish to prevent the economy from falling into a deflation trap, a problem which can emerge when weak aggregate demand leads inflation to fall and eventually become negative. When aggregate demand is very weak, the central bank will want to reduce interest rates in order to stimulate interest-sensitive spending like investment. This can push the nominal interest rate close to its lower bound of zero. But when a nominal interest rate close to zero is combined with falling prices (deflation), this implies a positive real interest rate which may be too high to stimulate private sector demand and get the economy back to equilibrium. Continued weak demand will make inflation more negative, thereby pushing the real interest rate up. This is exactly the opposite of what the central bank wants, which is to reduce the interest rate sufficiently to escape the deflation trap.

In addition, deflation increases real debt burdens. Debts are typically denominated in nominal terms, with borrowers having to pay back a fixed amount each period. If wages (and prices) are falling every period then the burden of the debt increases as a proportion of income. This will lower households' disposable income and squeeze firm's profits, exerting a drag on economic growth and slowing down recovery.

Deflation poses a third problem related to the apparent difficulty in cutting nominal wages. If workers are particularly resistant to nominal wage cuts, then a positive rate of inflation creates the flexibility needed to achieve changes in *relative* wages. For example, if, due to a fall in demand for one kind of labour, a real wage cut is required it can be achieved with an inflation rate of, say, 2% p.a. with the nominal wage left unchanged in the sector where the real wage cut is necessary. This argument is referred to as inflation's role in 'oiling the wheels of the labour market'.

3 See Friedman (1969).

Central bank independence and inflation targeting coincided with the achievement of a low and stable inflation environment as reflected in Fig. 3.2. This helped to reinforce the virtues of a 2% inflation target and central banks have not seriously considered changing to a higher target. In fact, the US Federal Reserve, which has historically avoided having an official inflation target, announced at the start of 2012 that it was joining other developed economies in targeting 2% inflation.[4]

3.1.3 Introduction to the 3-equation model

Figure 3.1 puts together the demand and supply sides of the economy and shows the role an inflation-targeting central bank can play in responding to a shock that takes the economy away from equilibrium. What the inflation-targeting central bank does is to raise the interest rate in response to inflation which is above the target rate in order to dampen aggregate demand and to lower the interest rate when inflation falls below the target rate.

A crucial characteristic of a modern monetary policy regime is that the central bank is forward looking. The central bank forecasts inflation by analysing what is going on in the economy and it must take into account the lags between changes in the interest rate it uses for policy purposes and the impact of that rate on economic activity. The importance of forecasting and lags for the work of central banks is highlighted in the monetary policy strategy document of the Sveriges Riksbank (the Swedish central bank):[5]

> It takes time before monetary policy has a full impact on inflation and the real economy. Monetary policy is therefore guided by forecasts for economic developments.

When three things are known—what the central bank is trying to achieve, how it thinks about the constraints it faces arising from the behaviour of the private sector, and how it implements its policy—this information will provide a skeletal model enabling an analysis of how the bank will respond to a variety of shocks to the economy. These responses to shocks can then be summarized in the form of a simple monetary policy rule. This rule sits at the heart of the 3-equation model, which is produced by adding the Monetary Rule (MR), curve to the *IS* curve from the demand side and the Phillips curve (*PC*) from the supply side.

We now turn to each of the three questions in turn to pin down how monetary policy is conducted within the 3-equation framework:

1. What is the central bank trying to achieve? It is assumed that its aim is to use monetary policy to stabilize the economy, which means keeping the economy close to equilibrium output and keeping inflation close to its targeted rate. The central bank may be penalized if it fails to meet its inflation target. In the UK, for example, if the inflation rate diverges more than one percentage point above or below the 2% target then the Governor of the Bank of England is obliged to write an open letter to the Chancellor of the Exchequer explaining the reasons for the bank's failure and a plan for returning inflation to target. This is, at the least, embarrassing for the Governor and costly in terms of the central bank's

4 Federal Reserve FOMC statement of longer-run goals and policy strategy, January 25th 2012.
5 Excerpt taken from: Sveriges Riksbank, 2010, *Monetary policy in Sweden*.

reputation. The closer is the economy to having output at equilibrium and inflation on its target, the lower is the reputational cost to the bank.

2. What forces prevent the central bank from achieving its target? This is essential information in order to assess the prospects of economic stabilization. As seen in Chapters 1 and 2, the economy is affected by shocks to demand and supply, which affect output and inflation. An unforeseen boom, for example, which takes output above equilibrium, will increase inflation above the target as the position of workers in the labour market strengthens. This aspect of wage behaviour is captured in the Phillips curve (*PC*). The central bank has to take into account behaviour, including the persistence of inflation, when it designs its response to the initial shock to the economy.
3. How does the central bank translate its objectives into monetary policy? It uses a monetary policy rule. This can be represented in the same diagram as the Phillips curves, that is, in a diagram with output on the horizontal axis and inflation on the vertical axis. This will demonstrate how the central bank will choose its preferred policy response given the constraint it faces from the behaviour of wage and price setters, which is captured by a particular Phillips curve.

To implement its preferred policy response in practice, the central bank diagnoses the shock and its forecast effect on inflation and output. It uses this, together with its preferences for stabilization, to estimate the output gap it is trying to achieve. The *MR* curve illustrates the central bank's best response to the shock. It then uses the relationship between the interest rate and output in the *IS* curve to implement that choice.

An example—a consumption boom

The example of a consumption boom is used to show how the model works. Figure 3.4 shows how the central bank analyses a shock of this kind.

The shock has its initial positive impact on output and employment (shown by the rightward shift of the *IS* curve and the movement from *A* to *B* in the *IS* diagram). By following the diagrams in Fig. 3.4 from top to bottom, we can see that the labour market impact of the consumption boom is reflected in a disturbance of the initial constant-inflation equilibrium. At the first wage-setting round following the fall in unemployment, wage and price inflation will rise from the target inflation rate of 2%. As illustrated in the wage-setting and price-setting diagram, the shock has opened up a gap of 1% between the prevailing (low) real wage and the (high) real wage consistent with the tighter labour market following the consumption boom. This means that workers will get a 2% wage rise to compensate them for the erosion of their real wages due to last period's inflation and an additional 1% increase to bring their real wage up to the level indicated by the wage-setting curve. In this case, wages rise by 3% and firms automatically raise their prices by 3% to protect their profit margins; in other words, inflation rises to 3%.

The increase in inflation from 2% to 3% is shown by the movement along the Phillips curve from *A* to *B*. Here, we need to take into account the fact that the higher inflation of 3% will become embedded in the expectations of wage setters. For the next wage-setting round—the one the central bank must anticipate in taking its stabilization decision—expected inflation of 3% will mean that the relevant Phillips curve facing the central bank is the one labelled '*PC* (expected inflation = 3%)'. Given this forecast of inflationary behaviour, the central bank

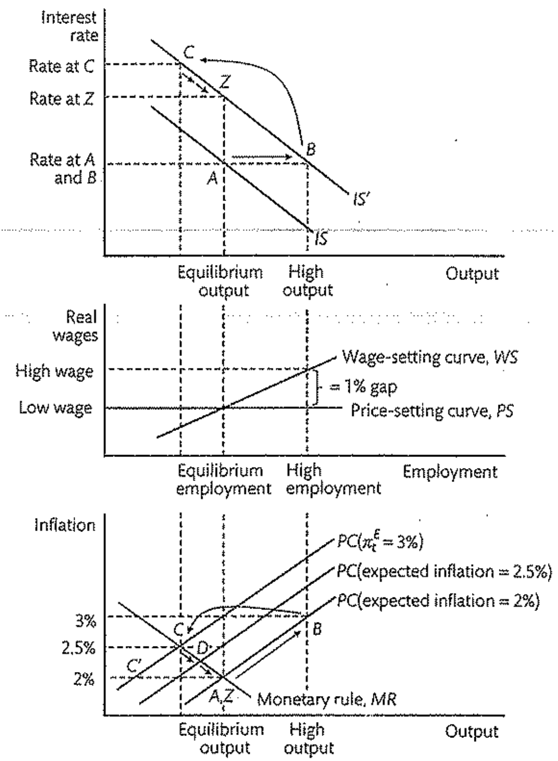


Figure 3.4 The 3-equation model: the adjustment of the economy to a permanent demand shock.

chooses its best response to the situation—the place on the new *PC* where it would like to locate. Its preference is likely to be for a balanced response—making progress in getting inflation back toward target but not imposing a harsh recession—such as at point *C*. This is a point on the *MR*, which shows the central bank's optimal output-inflation pair for any Phillips curve it faces.

In the top panel of the diagram is the *IS* curve. If the central bank takes the view that in the absence of any response on its part the consumption boom would persist, it will need to choose the interest rate shown by point *C* in order to dampen aggregate demand sufficiently to move the economy onto the path which leads to the inflation target. Once the economy is on the *MR* curve at point *C*, inflation will come down as forecast. Lower inflation will then shift expected inflation down. We can see this in the shift of the Phillips curve downwards to '*PC*(expected inflation = 2.5%)'. The central bank will choose point *D* as its best response to the new inflation environment it faces and will adjust the interest rate down from its peak. In the subsequent periods, the central bank will continue to guide the economy along the *MR* curve by reducing the interest rate as inflation returns to the target rate. The adjustment process finishes when the economy is back at equilibrium output and target inflation (point *Z*.).

There will, however, be a higher interest rate at the new equilibrium, as shown in the *IS* diagram (the top panel of Fig. 3.4). A reduced level of investment is needed to offset the persistent consumption boom.

A summary of modern inflation-targeting central banking

Central bank thinking on monetary policy in the era of inflation targeting was succinctly summarized by Glenn Stevens, the Assistant Governor of the Reserve Bank of Australia in a speech in 1999, in which he lists the six key things that have been learnt about monetary policy after 200 years of thought by economists and others:[6]

1. monetary policy affects principally, or only, prices [and not economic activity] in the medium term;
2. it affects activity in the short term;
3. because of lags, policy has to look forward;
4. but the future is uncertain, as is the impact of policy changes on the economy;
5. expectations matter, so giving people some idea of what you are trying to do, and acting consistently, is useful;
6. and an adequate degree of operational independence for the central bank [from the government] in the conduct of monetary policy is important.

3.2 Modelling

In the previous two chapters, we have built simple models of the demand and supply sides of the economy. The primary aim of this chapter is to add the policy maker to the model of the macro-economy. This gives the three curves in the graphical model based on the three equations that give the model its name:

1. *IS* curve.
2. *PC* curve.
3. *MR* curve.

This section extends the introduction to the 3-equation model by setting out in more detail how the central bank goes about achieving its objective of keeping inflation at target. This gives the *monetary rule*, or *MR* curve, which determines the output gap the central bank should set in order to stabilize the economy following an economic shock. We go on to use the 3-equation model to analyse a range of shocks in Section 3.3.

3.2.1 The 3-equation model

The first task of this section is to derive the *MR* curve, which shows the chosen output gap of the central bank in response to any economic shock. The *MR* curve shows the path along which the central bank seeks to guide the economy back to target inflation (and equilibrium output). The mathematical derivation of the model is shown in the Appendix.

The basic method for deriving a monetary policy rule involves the following steps:

1. Define the central bank's preferences in terms of a utility (or loss) function to capture the costs it incurs of being away from the inflation target and from equilibrium output. This produces the policy maker's indifference curves in output-inflation space and shows what the policy maker would like to do—i.e. to be close to the inflation target at equilibrium output.
2. Define the constraints faced by the policy maker from the supply side of the economy: these are the Phillips curves, which are also shown in output-inflation space. These show the 'objective' trade-off between inflation and unemployment in the short run and pin down what it is feasible for the central bank to achieve.
3. Derive the best response *monetary rule* in output-inflation space: this is the *MR* curve. For a given Phillips curve that it faces, the *MR* shows the central bank's desired output-inflation combination.
4. Once the central bank knows where it wants to be by using the *MR* curve, it uses the *IS* curve to implement that choice, since the *IS* curve shows the interest rate that will deliver the central bank's chosen level of output. The interest rate is the central bank's policy lever for influencing aggregate demand.

The central bank's preferences

Where do the central bank's preferences come from? A pragmatic way of thinking about how to model this is to infer their preferences from their behaviour. From the empirical analysis of the behaviour of the US central bank, the Federal Reserve, the economist John Taylor inferred a monetary policy rule. The famous Taylor Rule can be derived from a model in which the central bank minimizes fluctuations from the inflation target and the size of the output gap. Since our aim is to throw light on the way central banks behave, we use this loss function.

In this subsection, we use the central bank's loss function (i.e their utility function) to derive indifference curves representing the trade-off in its preferences between inflation being away from its target and output being away from equilibrium. Looking first at inflation, we assume that it wants to minimize fluctuations around the inflation target π^T:

$$(\pi_t - \pi^T)^2.$$

A loss function is just like a utility function except that the higher the loss, the worse it is for the central bank. This loss function has two implications. First, the central bank is as concerned to avoid inflation below its target as it is inflation above its target. If $\pi^T = 2\%$ the loss from $\pi_t = 4\%$ is the same as the loss from $\pi_t = 0\%$. In both cases $(\pi_t - \pi^T)^2 = 4$. Second, it attaches increased importance to bringing inflation back to its target the further it is away from π^T; the loss from $\pi_t = 6\%$ is 16, compared to the loss of 4 from $\pi_t = 4\%$.

We turn now to the central bank's second concern of keeping output close to equilibrium. We assume the central bank seeks to minimize the gap between y_t and y_e—remembering that it has no way of controlling y_e itself—in order to aid it in achieving its inflation target.

[6] Source: Mr G.R. Stevens, Assistant Governor of the Reserve Bank of Australia, 20 April 1999, *Six Years of Inflation Targeting*, speech to the Economic Society of Australia, Sydney. Square brackets indicate additions to the text made by the authors.

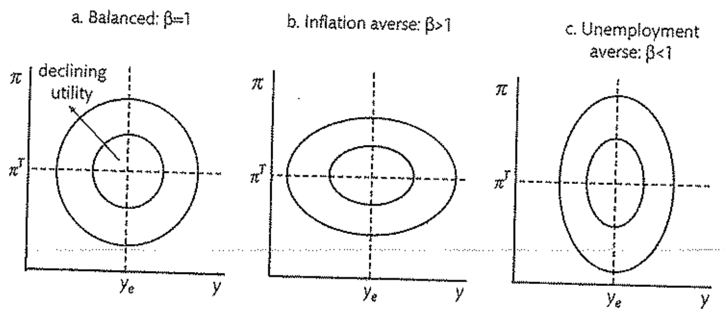


Figure 3.5 Central bank loss functions: utility declines with distance from the 'bliss point'.

The central bank's loss as a result of output being different from its target of y_e is $(y_t - y_e)^2$.

Note that this loss function also assumes a symmetrical attitude to positive and negative deviations—in this case, from the equilibrium level of output. The most straightforward way of thinking about this is that the central bank understands the model and realizes that inflation is only constant at $y_t = y_e$. If $y_t < y_e$ then this represents unnecessary unemployment that should be eliminated. If $y_t > y_e$, this is unsustainable and will require costly increases in unemployment to bring the associated inflation back down. Whenever the economy is disturbed, the central bank sees its task as steering the economy back to this constant-inflation output level.

Adding the two loss functions together, the central bank's loss function is:

$$L = (y_t - y_e)^2 + \beta(\pi_t - \pi^T)^2, \qquad \text{(central bank loss function)}$$

where β is the relative weight attached to the loss from inflation. This is a critical parameter: a $\beta > 1$ will characterize a central bank that places less weight on deviations in employment from its target than on deviations in inflation, and vice versa. A more inflation-averse central bank is characterized by a higher β.[7]

With $\beta = 1$, the central bank is equally concerned about inflation and output deviations from its targets. The loss function is simple to draw: each indifference curve is a circle with (y_e, π^T) at its centre (see Fig. 3.5a). The loss declines as the circle gets smaller. When $\pi_t = \pi^T$ and $y_t = y_e$, the circle shrinks to a single point (called the 'bliss point') and the loss is at a minimum, which is zero. With $\beta = 1$, the central bank is indifferent between inflation 1% above (or below) π^T and output 1% below (or above) y_e. They are on the same loss circle.

If $\beta > 1$, the central bank is called inflation-averse: it is indifferent between (say) inflation 1% above (or below) π^T and output 2% above (or below) y_e. They are on the same loss curve. This makes the indifference curves ellipsoid as in Fig. 3.5b. They are flat because the central bank is willing to trade off a small fall in inflation for a large rise in unemployment above equilibrium. A central bank with less aversion to inflation ($\beta < 1$) will have ellipsoid indifference curves with a vertical rather than a horizontal orientation (Fig. 3.5c). In that case, the indifference curves are steep, reflecting that the central bank is only willing to trade off a given fall in inflation for a smaller fall in output than in the other two cases. If the central bank cares only about inflation then the loss ellipses become one dimensional along the line at $\pi_t = \pi^T$.

The value of β does not reflect whether the central bank focuses on achieving an inflation target or an output target. Rather, a central bank with lower β is willing to trade off a longer period during which inflation is away from target to reduce the impact on unemployment of the adjustment path back to equilibrium than would a more inflation-averse central bank with a higher β. Central bank preferences are discussed in more detail in Chapter 13.

7 The central bank's preferences can be presented in this simple way if we assume that the central bank's discount rate is infinite. This means that it only considers one period at a time when making its decision.

The Phillips curve constraint

As discussed in Chapter 2, output affects inflation via the (adaptive expectations) Phillips curve:

$$\pi_t = \pi_t^E + \alpha(y_t - y_e) \qquad (3.1)$$
$$= \pi_{t-1} + \alpha(y_t - y_e), \qquad \text{(Phillips curve, } PC\text{)}$$

where $\pi_t^E = \pi_{t-1}$.

The Phillips curve is a constraint for the central bank because it shows all the output and inflation combinations from which the central bank can choose for a given level of expected inflation. In other words, in any period, the central bank can only choose to locate the economy at a point on the Phillips curve it faces. In the Phillips curve used in this chapter, expected inflation is simply lagged inflation—other methods of forming inflation expectations will be discussed in Chapter 4.

This is shown in Fig. 3.6, where the upward sloping lines are the Phillips curves we worked with in Chapter 2. In the diagram, it is assumed that $\alpha = 1$, so that each Phillips curve has a slope of 45°. Each Phillips curve is labelled by a given level of expected inflation, which in the adaptive expectations Phillips curve is equal to lagged inflation. Assume that $\pi_t^E = \pi_{t-1} = \pi^T = 2$ (remember that this PC must go through point A at which $y = y_e$ and $\pi = 2$). The central bank is in the happy position of being able to choose bliss point A or (π^T, y_e) at which its loss is zero.

What happens if there has been a shock to inflation and it is not equal to the inflation target? Suppose, for example, that lagged inflation was 4%. The central bank is faced with the constraint of the Phillips curve shown by $PC(\pi_t^E = 4)$ and can only choose among points along it. The bliss point is no longer obtainable. The central bank faces a trade-off: if it wants a level of output of y_e next period, then it has to accept an inflation rate above its target, i.e. $\pi_t = 4 \neq \pi^T$ (i.e. point B). On the other hand, if it wishes to hit the inflation target next period, it must accept a much lower level of output next period (point C). Point B corresponds to a fully accommodating monetary policy in which the objective is purely to hit the output target ($\beta = 0$); and point C corresponds to a completely non-accommodating policy, in which the objective is purely to hit the inflation target.

It is clear from Fig. 3.6 that given its preferences, if the central bank is faced by $PC(\pi_t^E = 4)$, then it can do better (achieve a loss circle closer to A) than either point B or point C. It minimizes its loss function by choosing point D, where the $PC(\pi_t^E = 4)$ line is tangential to

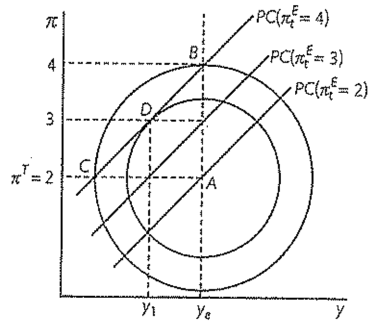


Figure 3.6 Loss circles and Phillips curves.

the indifference curve of the loss function closest to the bliss point. Thus if is it on $PC(\pi_t^E = 4)$ it will choose an output level y_1 which will in turn imply an inflation rate of 3% and a Phillips curve the following period of $PC(\pi_t^E = 3)$.

Deriving the monetary rule (*MR*) curve graphically

The *MR* curve shows the central bank's preferred output-inflation combination for any Phillips curve it faces. It can be derived graphically by finding the points of tangency between the Phillips curves and the loss circles. As shown in Fig. 3.7a, points *A*, *B* and *C* all minimize the loss function of the central bank for their given Phillips curve. For example, take point *B*, which is on $PC(\pi_t^E = 3)$. If the central bank were to choose any other point on that Phillips curve, then the economy would be on a loss circle further from the bliss point and hence at a lower level of utility.

If we join up the points of tangency as shown in Fig. 3.7b, this gives us the *MR* curve. This shows the output and inflation combination that the central bank will choose to minimize its loss function for any given Phillips curve it is faced with.

In the Appendix, we derive the monetary rule using the equations for the central bank's loss function and the Phillips curve. The equation for the *MR* curve shown in Fig. 3.7 is:

$$(y_t - y_e) = -\alpha\beta(\pi_t - \pi^T). \qquad \text{(Monetary rule, } MR\text{)}$$

The *MR* tells the central bank what output gap it should choose when it observes that inflation is away from its target. The monetary policy rules used by central banks are often described as Taylor Rules. The difference between the *MR* and a Taylor Rule is that the latter is expressed in terms of the interest rate the central bank should choose to implement its chosen output gap. Taylor Rules used by central banks are discussed in Chapter 13. To find out the interest rate the central bank should choose once it has decided on its preferred output gap using the *MR*, for example, to achieve point *C* at an output level of y_1 in Fig. 3.7, we need to introduce the *IS* curve. In addition to the Phillips curve (*PC*) and the monetary rule (*MR*), the *IS* is the third equation in the 3-equation model.

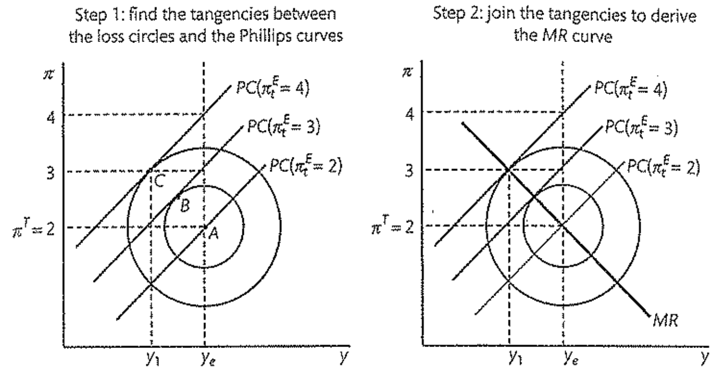


Figure 3.7 Deriving the *MR* curve.

Implementing central bank policy—the *IS* curve

In the previous subsection, we set out the process by which the central bank pins down its preferred output gap in response to an economic shock. The instrument the central bank uses to implement its policy is the real interest rate, *r*. The fact that the central bank must adjust the *nominal* interest rate that it sets in order to achieve a particular *real* interest rate on the *IS* curve is called the Taylor Principle.

The real interest rate is chosen to secure the appropriate level of aggregate demand and hence output. The central bank chooses the best point along the Phillips curve that it faces and in order to deliver the right level of aggregate demand, it must set the interest rate at the level shown by the *IS* curve. The *IS* curve shows the effect of changes in the interest rate on aggregate demand and was introduced in the following form in Chapter 1:

$$\begin{aligned} y &= k(c_0 + a_0 + G) - ka_1 r && (3.2) \\ &= A - ar, && (\textit{IS}\text{ curve}) \end{aligned}$$

where $A \equiv k(c_0 + a_0 + G)$, $a \equiv ka_1$ and k is the multiplier. The term *A* includes the multiplier and the demand shift variables such as government spending and both the autonomous and forward-looking components of consumption and investment.

In the 3-equation model, we will use a 'dynamic' *IS* curve to represent the demand side, which captures the fact that aggregate demand responds negatively to the real interest rate (*r*) with a one period lag. This is in line with the Bank of England's estimates that the impact of interest rates on output takes up to (a maximum of) one year.[8]

The dynamic *IS* curve is defined as follows,

$$y_t = A_t - ar_{t-1}. \qquad \text{(dynamic } \textit{IS}\text{ curve)}$$

[8] Compare the *How Monetary Policy Works* section of the Bank of England's website.

3.2.2 Using the 3-equation model

This section shows how the 3-equation model can be used to explain the response of the central bank to a shock. The economy is initially at the central bank's 'bliss point'—output is at equilibrium and inflation is at target. This is the medium-run equilibrium, y_e, where the *WS* and *PS* curves intersect and there is therefore no pressure on inflation to change. The interest rate that is associated with equilibrium is known as the *stabilizing rate of interest*, or r_S.

Inflation shock

An inflation shock is the simplest to analyse and serves to highlight the dynamic behaviour of the model. The term inflation shock is used to refer to an exogenous shift in the Phillips curve. This could be caused, for example, by a natural disaster such as a drought that reduces agricultural output and raises food prices, or by a temporary burst of union militancy, which pushes up wages.

In the 3-equation model framework, the central bank follows two main steps to stabilize the economy after an inflation shock:

1. The inflation shock shifts the Phillips curve upwards. The central bank must choose the position on the new Phillips curve that minimizes their loss function. This will be where the *MR* curve intersects the new Phillips curve. As inflation is above target, the central bank will have to reduce output below equilibrium to squeeze inflation out of the system.

2. The central bank uses the *IS* curve to find the increase in the real interest rate required to get the economy back onto the *MR* curve. The higher interest rate dampens output and inflation starts to fall. The central bank then gradually reduces the interest rate until output rises back to equilibrium and inflation falls back at target.

Figure 3.8 provides a summary of the dynamic adjustment of the economy to an inflation shock (start at the top and follow round). Figure 3.9a shows the typical fashion in which shocks can be modelled graphically in the 3-equation model, with the *IS* diagram at the top and the *PC* – *MR* diagram directly below it. This allows the position of the economy to be shown simultaneously on both diagrams. As shown in Fig. 3.9a, the adjustment of the economy to the inflation shock is as follows:

Period 0 The economy starts at *A*, where the central bank's utility is highest. The economy is then hit by an inflation shock which shifts the *PC* to *PC*(inflation shock) in the lower panel and the economy moves from *A* to *B*. Point *B* is not on the central bank's *MR* curve. The central bank forecasts the *PC* in the next period. In this case, the *PC* will stay in its post shock position, as output has not deviated from equilibrium as a result of the shock. We denote next period's *PC* as $PC(\pi_1^E = \pi_0)$. Faced with this *PC*, the central bank would like to locate at point *C*, back on the *MR* curve. They therefore set the interest rate at r_0. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 0 with inflation at π_0, output at y_e and the interest rate at r_0.

Period 1 The new interest rate has had time to reduce aggregate demand by dampening investment. The economy moves to point *C*, with output below equilibrium at y_1 and inflation at π_1. The central bank forecasts the *PC* in the next period. The Phillips curve shifts when

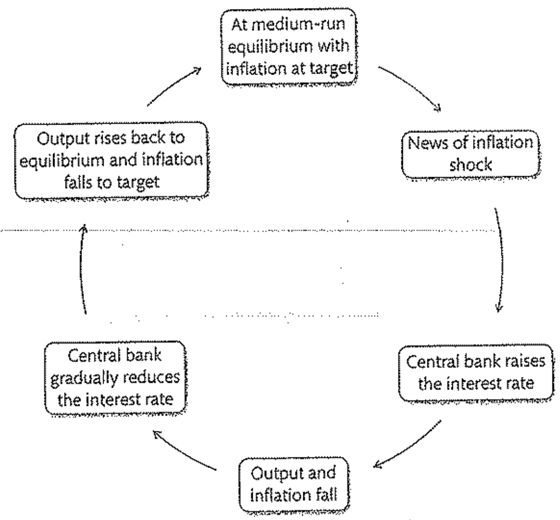


Figure 3.8 Dynamic adjustment to an inflation shock.

inflation expectations are updated: hence the *PC* will change in the next period, moving to $PC(\pi_2^E = \pi_1)$. Faced with this *PC*, the central bank would like to locate at point *D*, back on their *MR* curve. They therefore reduce the interest rate to r_1. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 1 with inflation at π_1, output at y_1 and the interest rate at r_1.

Period 2 onwards In period 2, the economy moves to point *D*, as the lower interest rate stimulates demand. This increases output to y_2 and inflation falls to π_2. The same process now repeats itself until the economy is back at equilibrium at *Z*. This involves the central bank forecasting next period's *PC* and setting the interest rate in the current period to ensure they stay on the *MR* curve in the following period. The adjustment from *D* to *Z* will take a number of periods. The economy will move gradually down the *MR* curve, as the central bank slowly adjusts the interest rate down from r_1 to r_S. This causes output to rise slowly from y_2 to y_e and inflation to fall slowly from π_2 to π^T. The economic adjustment to the inflation shock ends when the economy is back at point *Z*, with output at y_e, inflation at π^T and the interest rate at r_S.

Figure 3.9b shows the path of the key macroeconomic variables over time following the inflation shock. These types of graphs are called *impulse response functions* and are useful for visualizing the adjustment path of the economy following an economic shock. In this case, we can see that inflation has risen after the shock and is then slowly brought back to target. The interest rate follows a similar path, rising as soon as the shock takes place and then slowly being reduced to target. The path of output is slightly different, in that it does not change until one period after the shock. This is because the higher interest rate takes one period to affect output, as shown by the *IS* relation underlying the model, $y_t = A - ar_{t-1}$.

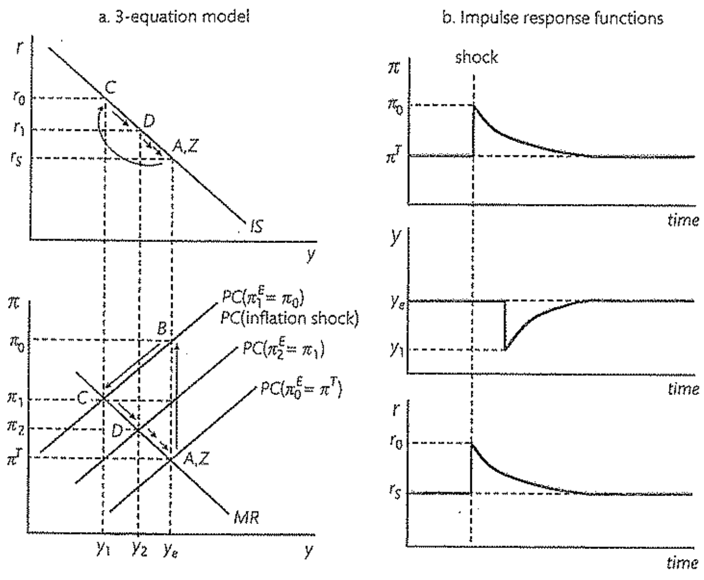


Figure 3.9 Inflation shock and the monetary rule.

The 3-equation model can be used to analyse a wide range of situations faced by the central bank. In Section 3.3, it is used to analyse demand and supply shocks and to examine the special case of a deflation trap.

3.3 Applications

In this section, we show more examples of the 3-equation model at work, focusing on demand- and supply-side shocks. We highlight the integral role played by central bank forecasting in identifying the best response interest rate to ensure medium-run economic stability. We go on to analyse the special case of a deflation trap, where a sizeable negative demand shock pushes the economy into a vicious cycle of falling inflation and output. We provide some escape routes from a deflation trap, whilst noting their potential pitfalls. We end this section by looking at a supply shock, which is a shift in either the wage- or price-setting curve. In this case, the central bank acts to stabilize the economy at the new equilibrium level of output fixed by the new intersecton of the *WS* and *PS* curves.

3.3.1 A temporary demand shock

We assume that the economy starts off in equilibrium with output at the equilibrium and inflation at the target rate of π^T (see Fig. 3.10). The economy is then disturbed by a temporary positive aggregate demand shock. By a temporary aggregate demand shock, we mean that the shock shifts the *IS* curve to *IS'*, but it only remains at *IS'* for one period. The economy is

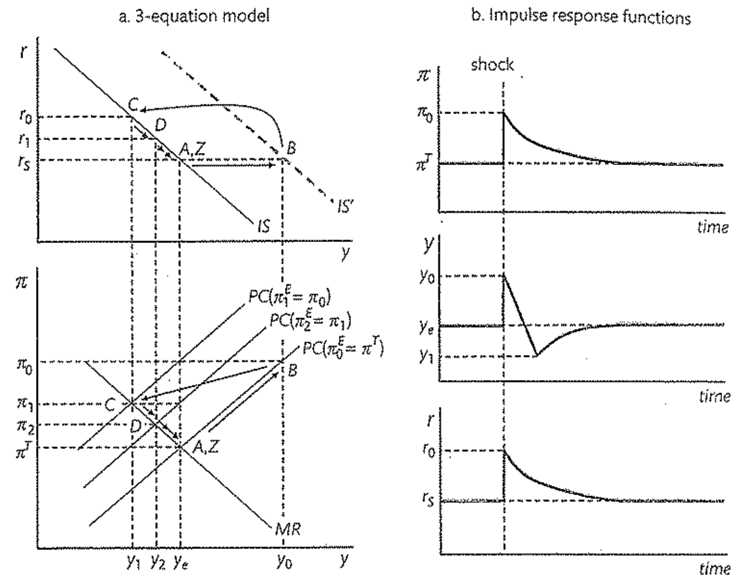


Figure 3.10 Adjustment of the economy to a temporary positive aggregate demand shock.

shifted from *A* to *B* in Fig. 3.10 as a result of the aggregate demand shock. This rise in output builds a rise in inflation above target into the system. Because of inflation inertia, this can only be eliminated by pushing output below and (unemployment above) the equilibrium. The central bank therefore raises the interest rate in response to the aggregate demand shock because it can work out the consequences for inflation. It must raise the interest rate in order to depress interest-sensitive demand and reduce output. The central bank is forward looking and takes all available information into account.

The adjustment to the shock is shown in Fig. 3.10.

Period 0 The economy starts at *A*—the central bank's bliss point. Following the demand shock, the economy moves from *A* to *B*. The shock has increased both output ($y_0 > y_e$) and inflation ($\pi_0 > \pi^T$). Point *B* is not on the central bank's *MR* curve. The central bank forecasts the *PC* in the next period. The Phillips curve shifts when inflation expectations are updated: hence the *PC* will change in the next period, moving to $PC(\pi_1^E = \pi_0)$. Faced with this *PC*, the central bank would like to locate at point *C*, back on their *MR* curve. They therefore set the interest rate at r_0. The central bank expects the demand shock to be temporary, so that the *IS* curve will shift back in the next period. This is the reason they use the original *IS* curve to set the interest rate. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 0 with inflation at π_0, output at y_0 and the interest rate at r_0.

Period 1 The new interest rate has had time to affect aggregate demand. The higher rate of interest dampens investment. This reduces output and the economy moves to point *C*, with output below equilibrium at y_1 and inflation at π_1. The central bank forecasts the *PC* in the next period: the *PC* will move to $PC(\pi_2^E = \pi_1)$. Faced with this *PC*, the central bank would like to locate at point *D*, on their *MR* curve. They therefore reduce the interest rate to r_1. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 1 with inflation at π_1, output at y_1 and the interest rate at r_1.

Period 2 onwards In period 2, the economy moves to point *D*, as the lower interest rate stimulates demand. This increases output to y_2 and inflation falls to π_2. The same process now repeats itself until the economy is back at equilibrium at *Z*. The economic adjustment to the temporary demand shock ends when the economy is back at point *Z*, with output at y_e, inflation at π^T and the interest rate at r_S.

We can use a Reserve Bank of Australia press release from March 2005 to provide an example of a case where an inflation-targeting central bank increased interest rates by 25 basis points to 5.5% in response to rising demand pressures:[9]

> In Australia, there are high levels of confidence in both the business and household sectors, credit growth is providing ample support for spending, employment is growing strongly and national income and spending will continue to be boosted this year by the rising terms of trade.

3.3.2 Forecasting and lags

The example of the central bank's reaction to a demand shock can be used to highlight the role played by forecasting and lags in the effect of monetary policy on aggregate demand and output. In the previous subsection, we discussed the adjustment of the economy to a temporary positive demand shock (Fig. 3.10). We now use Fig. 3.11 to compare this to the case of a permanent demand shock. It is left as an exercise for the reader to set out the detailed period by period adjustment for the permanent case.

The examples highlight two important points about the central bank's reaction to an aggregate demand shock:

1. The central bank has to forecast both the *PC* and *IS* curves. The forecasting of the *IS* curve means predicting the length of the shock—is it temporary or permanent?
2. The reason that the central bank must forecast the *IS* is that the persistence of the shock affects the central bank's preferred reaction. In Fig. 3.11, where the shock is permanent, the *IS* curve remains at *IS'*. The initial increase in the interest rate (i.e. from r_S to r_0) is much greater in the case of a permanent shock. In addition, a permanent shock leads to a higher stabilizing rate of interest in the new equilibrium—r_S'. This is because the higher autonomous aggregate demand (for example, due to improved consumer confidence) needs to be offset by lower interest-sensitive aggregate demand (i.e. investment) if output is to return to its equilibrium level y_e.

What happens if the central bank is uncertain about whether the shock is temporary or permanent? In this case, the central bank would be likely to set an interest rate somewhere in

[9] Excerpt taken from: Mr Ian Macfarlane, 2nd March 2005, *Statement by the Governor, Mr Ian Macfarlane: Monetary Policy*.

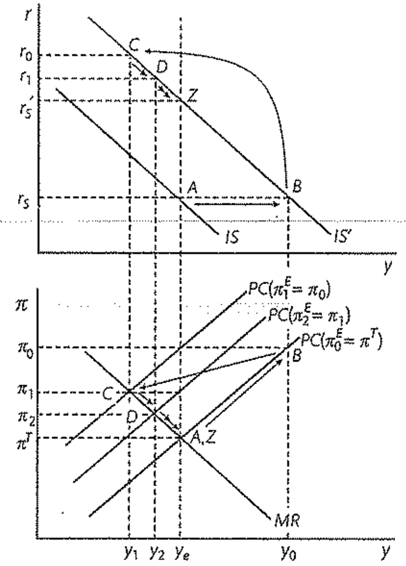


Figure 3.11 Adjustment of the economy to a permanent positive aggregate demand shock.

between the interest rates it would set in the cases presented in Figs. 3.10 and 3.11. In reality, central banks are unlikely to be able to tell (in the period of the shock) whether shocks will be temporary or permanent, so use their best judgement to respond to shocks and often have to adjust interest rates as new information arrives.

In the benchmark 3-equation model, the interest rate affects output with a one period lag: $y_t = A - ar_{t-1}$. However, it is interesting to see what happens if the central bank could affect output immediately, i.e. if $y_t = A - ar_t$. In this case, as soon as the *IS* shock is diagnosed, the central bank would raise the interest rate to r_S'. The economy then goes directly from *A* to *Z* in the *IS* diagram (in Fig. 3.11) and it remains at *A* in the *PC*–*MR* diagram, i.e. points *A* and *Z* coincide. Since the aggregate demand shock is fully and immediately offset by the change in the interest rate, there is no chance for inflation to rise.

This underlines the crucial role of lags and hence of forecasting for the central bank: the more timely and accurate are forecasts of shifts in aggregate demand (and of other kinds of shock), the greater is the chance that the central bank can offset them and limit their impact on inflation. Once inflation has been affected, the presence of inflation inertia means that the central bank must change the interest rate and get the economy onto the *MR* curve in order to steer it back to the inflation target.

The Bank of England's first strategic priority for 2012/13 was to 'keep inflation on track to meet the Government's 2% target'. One of the key actions the Bank committed to take to achieve this objective was to 'exploit the new suite of models in forecasting and analysis'.[10]

[10] These excerpts are taken from *The Bank's Strategy* section of the Bank of England's website.

This shows the focus and resources modern central banks apply to forecasting and modelling to make better informed monetary policy decisions.

3.3.3 The deflation trap

Problems arise for using monetary policy along the lines of the 3-equation model when the real rate of interest needed to stabilize demand cannot be achieved because the nominal interest rate cannot be reduced further.

The zero lower bound on nominal interest rates

In Chapter 1, we set out the Fisher equation, which shows the relationship between the real and nominal interest rates and the expected rate of inflation:

$$i = r + \pi^E. \qquad \text{(Fisher equation)}$$

When responding to an economic shock, the central bank adjusts the nominal interest rate (i), in order to affect the real interest rate (r) which in turn affects aggregate demand through the *IS* relation. To do this, it must take into account expected inflation as shown by the Fisher equation. In this regard, we can model the central bank as setting the real interest rate.

We saw in Fig. 3.1, that the central bank will respond to falling inflation by reducing the interest rate to stimulate aggregate demand. There is, however, a limit to the extent to which the central bank can reduce the nominal interest rate. The lowest nominal interest rate the central bank can set is zero (see the box below). From the Fisher equation, we can see that if the real interest rate that is needed to achieve the central bank's chosen output gap on the *MR* curve is, for example, 0.75%, and expected inflation is −1%, then a nominal interest rate of −0.25% would be required. With a *zero lower bound (or ZLB)* on the nominal interest rate, the minimum real interest rate that can be achieved is set by:

$$\min r \geq -\pi^E. \qquad (3.3)$$

This means that in our example, the minimum real interest rate achievable is 1%, which is not low enough get the economy on to the *MR* and on the path back to equilibrium. This condition is shown in Fig. 3.12 where the stabilizing real interest rate is below the minimum feasible rate of 1%. Given the depressed state of aggregate demand depicted by the position of the *IS* curve, if inflation has fallen to −1%, then it will be impossible to achieve the equilibrium level of output using conventional monetary policy. The economy is stuck at point *A* on the *IS* curve.

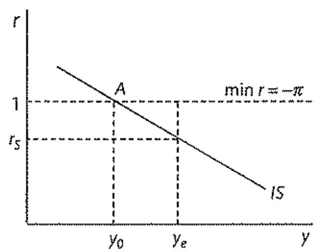


Figure 3.12 The zero lower bound on the nominal interest rate.

Box 3.1 Why can't nominal interest rates fall below zero?

There is no technical reason why nominal interest rates cannot be negative. The central bank could, for example, set the interest rate at -5%. In this environment, an individual who put \$100 in the bank at the start of the year could withdraw \$95 at the end of the year. This is not beneficial for savers, but this could be the interest rate required to get the economy onto the *MR* curve and back to equilibrium. It would certainly be expected to stimulate borrowing, by making the cost of obtaining credit cheaper.

There is, however, a major practical implication to negative interest rates; the presence of currency. When nominal interest rates are negative, there is nothing to stop people simply holding their savings in cash, which has a nominal interest rate of zero.

The existence of high denomination notes, such as the \$100 bill and the €500 note, reduces the cost associated with carrying large sums of money in cash. If only very small denomination notes were left in circulation then this cost of carrying cash might become so large that accepting negative interest rates would be preferable. However, monetary authorities have shown no intention of moving towards this system so currency still acts as the limiting factor on setting nominal interest rates.[a]

[a] For a more thorough discussion on negative interest rates see Willem Buiter's Maverecon blog column for the Financial Times entitled 'Negative interest rates: when are they coming to a central bank near you?' (May, 2009).

The deflation trap and the 3-equation model

We can use the 3-equation model to provide a period by period explanation of how a negative aggregate demand shock can lead to the economy becoming stuck in a deflation trap, where output and inflation are falling without limit. We assume that the economy starts off with output at equilibrium and inflation at the target rate of π^T as shown in Fig. 3.13. The economic adjustment to the shock is as follows:

Period 0 The economy starts at *A*—the central bank's bliss point. The economy is hit by a large permanent negative aggregate demand shock which shifts the *IS* curve to *IS′* and the economy moves from *A* to *B*. The shock has reduced both output ($y_0 < y_e$) and inflation ($\pi_0 < \pi^T$). In fact, inflation has become negative—i.e. there is deflation. Point *B* is not on the central bank's *MR* curve. The central bank forecasts the *PC* in the next period. Inflation is currently below equilibrium at π_0, which means the *PC* will shift in the next period, moving to $PC(\pi_1^E = \pi_0)$. Faced with this *PC*, the central bank would like to locate at point *C′*, back on their *MR* curve. This would require setting an interest rate of r_0'. However, this is below the minimum real interest rate the central bank can achieve (by setting nominal interest rates to zero). The lowest interest rate they can achieve is $r_0 = -\pi_0$. The central bank expects the demand shock to be permanent, so that the *IS* curve will stay at *IS′* in the next period. This is the reason they use the new *IS* curve to set the interest rate. The interest rate can only

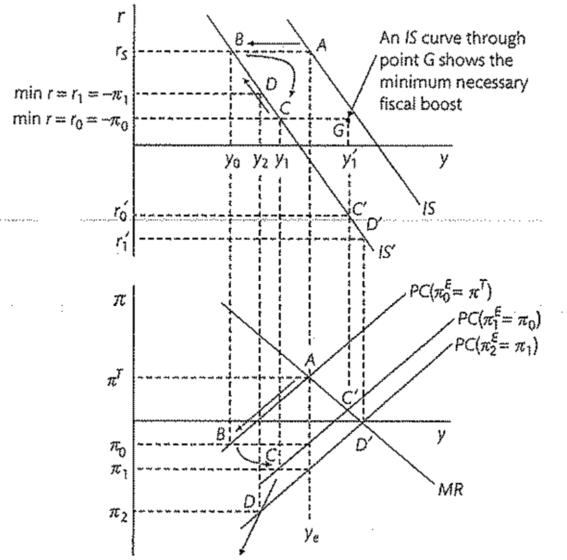


Figure 3.13 How a large negative permanent aggregate demand shock can lead to the economy entering a deflation trap.

affect economic output with a one period lag, however, so the economy ends period 0 with inflation at π_0, output at y_0 and the interest rate at r_0.

Period 1 The new interest rate has had time to affect aggregate demand. The lower rate of interest boosts investment. This increases output and the economy moves to point C, with output still below equilibrium at y_1. This level of output is far below the the central bank's best response level of y_1', which causes inflation to fall further to π_1. The central bank forecasts the PC in the next period. The PC will move to $PC(\pi_2^E = \pi_1)$. Faced with this PC, the central bank would like to locate at point D', back on their MR curve. This would require setting an interest rate of r_1'. However, this is below the minimum real interest rate the central bank can achieve (by setting nominal interest rates to zero). The lowest interest rate they can achieve is $r_1 = -\pi_1$. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 1 with inflation at π_1, output at y_1 and the interest rate at r_1.

Period 2 onwards In period 2, the economy moves to point D, as the higher interest rate dampens demand. This reduces output to y_2 and inflation falls further to π_2. The economy has entered a deflation trap—both output and inflation are falling every period and conventional monetary policy is powerless to stop them. In each future period, inflation becomes more negative, which increases the minimum real interest rate the central bank can achieve. The higher interest rate dampens demand, which further reduces inflation. In other words, the economy is caught in a vicious cycle. The arrows on Fig. 3.13 show that the economy will not revert to medium-run equilibrium in this case, but rather it will diverge, with ever-falling inflation and output.

The downward spiral of the economy in a deflation trap explains why policy makers are so keen to avoid this situation. We have shown in Fig. 3.13 that conventional monetary policy (e.g. adjusting interest rates) is insufficient to escape a deflation trap. There are, however, two policies which could be initiated to pull the economy out of this downward spiral:

1. The *IS* curve could be shifted to the right (to go through point *G*) in the period after the initial demand shock. This would allow the central bank to achieve their desired output level of y_1' by setting the minimum achievable interest rate of r_0. This positive demand shock could take place via a spontaneous recovery of autonomous investment or a recovery of autonomous consumption. Alternatively, the government can step in and implement the desired output gap shown by the *MR* curve by using fiscal policy. In this case, the *MR* becomes a more general policy rule and would be labelled *PR*, for policy rule.

2. The creation of more positive inflation expectations. If expected inflation becomes less negative, the min *r* line shifts down and the *PC* curve shifts up. This may allow the central bank to use the interest rate based monetary rule in the usual way to move the economy to the south-east along the *IS* curve.

However, the idea of escaping from the deflation trap by creating positive inflation expectations may not work in practice. The way to create expectations of inflation in the future is to create expectations of future higher aggregate demand: if the authorities do not take measures to create the demand, there is no reason to think that people will expect higher inflation.

Another potential option for central banks stuck at the zero lower bound is to introduce unconventional monetary policies, such as quantitative easing. We leave a discussion of this policy option until Chapter 7, however, as it is easier to explain once we have introduced the banking system into the model.

3.3.4 A supply shock

At the end of Chapter 2, we looked at the implications for inflation of a positive aggregate supply shock. This was modelled as a downward shift of the *WS* curve, which could be the result of a number of factors, such as a decrease in union bargaining power or a reduction in unemployment benefits. The initial effect of this supply shock is a downward shift of the Phillips curve. We will now go on to explain how the central bank stabilizes the economy in the event of a supply shock and the characteristics of the new medium-run equilibrium (see Fig. 3.14).

The stabilizing interest rate is lower (at r_S') in the new medium-run equilibrium. Since equilibrium output is higher as a consequence of the supply-side shift, a lower real interest rate is required to provide the appropriate level of aggregate demand. This example highlights that supply shocks differ from inflation and demand shocks in two key ways:

1. Firstly, a supply shock changes the equilibrium level of output (i.e. y_e' in Fig. 3.14).
2. Secondly, a supply shock shifts the *MR* schedule so that it goes through the point where inflation is at target and output at the new equilibrium (i.e. *MR'* in Fig. 3.14)

The period by period adjustment of the economy to a positive supply shock and the central bank's intervention is illustrated in Fig. 3.14 and described below.

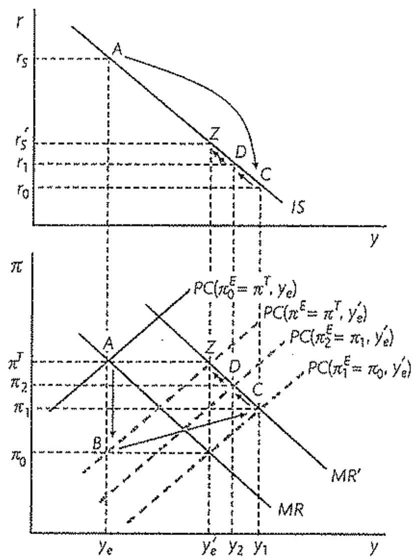


Figure 3.14 The adjustment of the economy to a positive permanent aggregate supply shock.

Period 0 The economy starts at A—the central bank's bliss point. The economy is hit by a permanent positive aggregate supply shock. This fundamentally changes the equilibrium level of output in the economy, increasing it from y_e to y_e'. At A, the economy is now below equilibrium output and there is pressure on inflation to fall (as $WS < PS$, see Chapter 2). The Phillips curve shifts down from $PC(\pi_0^E = \pi^T, y_e)$ to $PC(\pi_0^E = \pi^T, y_e')$, reflecting the change in the equilibrium level of output. The economy moves from A to B, where output is still at its original level, but inflation has fallen to π_0.

The central bank predicts that the shock is a permanent supply shock, so they shift the MR curve outwards to ensure their bliss point is now where output is equal to y_e' and inflation is equal to π^T. Point B is not on the central bank's new MR curve, MR'. The central bank forecasts the PC in the next period. Output is currently below the new equilibrium level of y_e', which means the PC will shift in the next period, moving to $PC(\pi_1^E = \pi_0, y_e')$. Faced with this PC, the central bank would like to locate at point C, on their new MR curve, MR'. They therefore set the interest rate at r_0. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 0 at with inflation at π_0, output at y_e and interest rates at r_0.

Period 1 The new interest rate has had time to affect aggregate demand. The lower rate of interest boosts investment. This increases output and the economy moves to point C, with output above the new equilibrium at y_1 and inflation at π_1. The central bank forecasts the PC in the following period. Output is still away from the new equilibrium, so the PC will change again in the next period, moving to $PC(\pi_2^E = \pi_1, y_e')$. Faced with this PC, the central bank would like to locate at point D, on the MR' curve. They therefore increase the interest rate to r_1. The interest rate can only affect economic output with a one period lag, however, so the economy ends period 1 with inflation at π_1, output at y_1 and interest rates at r_1.

Period 2 onwards In period 2, the economy moves to point D, as the higher interest rate dampens demand. This reduces output to y_2 and inflation rises to π_2. The same process now repeats itself in the usual way until the economy is back at equilibrium at Z. Adjustment to the supply shock ends when the economy is back at point Z, with output at y_e', inflation at π^T and the interest rate at r_S'.[11]

A positive supply shock can be the result of a shift in the WS or PS curves. We have shown the example of a downward shift in the WS curve above, but equally a positive supply shock could reflect an upward shift of the PS curve, due to increased product market competition or productivity gains. This latter case applies to the United States in the late 1990s, where the ICT revolution spurred productivity increases. The Federal Reserve correctly identified this as a positive supply shock and took into account its role in moderating inflation:[12]

> Responding to the availability of new technologies at increasingly attractive prices, firms have been investing heavily in new capital equipment; this investment has boosted productivity and living standards while holding down the rise in costs and prices.

3.4 Conclusions

This chapter has provided the final building block for the 3-equation model of the macroeconomy by setting out the monetary rule, or MR curve, which shows the central bank's chosen combination of output gap and inflation rate relative to target for any given Phillips curve they face. The equations of the 3-equation model are:

1. IS curve: $y_t = A - ar_{t-1}$.
2. PC curve: $\pi_t = \pi_{t-1} + \alpha(y_t - y_e)$.
3. MR curve: $(y_t - y_e) = -\alpha\beta(\pi_t - \pi^T)$.

The policy maker is modelled as an inflation-targeting central bank not because this is necessarily the best policy-making arrangement, but because it most closely resembles how modern stabilization policy is undertaken. We noted the economic and political reasons that have been used to explain why, in the more than two decades running up to the global financial crisis, stabilization policy was put in the hands of the monetary policy maker.

In Chapter 1, we introduced the IS curve to model the demand side of the economy. In Chapter 2, we introduced the Phillips curve (PC) to model the supply side of the economy.

11 Note that, the Phillips curve in the new equilibrium is denoted by $PC(\pi^E = \pi^T, y_e')$. This does not include a time subscript because the Phillips curve is in this position in two periods—in period 0 straight after the supply shock and then in the period when the economy completes its adjustment to the new equilibrium at point Z.

12 Excerpt taken from: The Federal Reserve Board, July 22nd, 1999, *Monetary Policy Report submitted to the Congress*.

In this chapter, we have introduced the *MR* curve to model the policy maker. These are the three components of the 3-equation model. The modelling section of this chapter set out the 3-equation model in full; the Appendix shows the mathematical derivation. Putting them together provides a framework in which to answer a variety of questions about the macro-economy and stabilization policy in particular:

1. Why target low and stable inflation? Central banks target low and stable inflation to minimize the negative effects of inflation on the economy—such as price distortions, shoe leather costs, menu costs—and to avoid high and rising (or volatile) inflation. Does that mean it is optimal to target 0% inflation? No, because central banks also want to guard against the threat of deflation (i.e. falling prices). Deflation is particularly problematic for the economy due to the possibility of entering a deflation trap. In light of this, the majority of developed economy modern central banks have chosen to target 2% inflation.

2. How do modern central banks go about achieving their inflation target? How do they react to economic shocks? The central bank's best response position—their 'bliss point'—is to be in medium-run equilibrium, with inflation equal to target (π^T) and output equal to equilibrium (y_e). The *monetary rule* (or *MR*) curve of a central bank shows the output and inflation combinations that minimize their *loss function* (i.e. maximize their utility) for any given Phillips curve they face. In response to an economic shock, the central bank finds the relevant Phillips curve and uses the *MR* curve to find their preferred output gap. The real interest rate (r) is the policy lever they use to achieve their desired output gap. It affects aggregate demand via the *IS* relation. The central bank then continues to adjust the interest rate to guide the economy to the new medium-run equilibrium. The use of the interest rate to implement monetary policy is sometimes called a Taylor rule based monetary policy. Demand shocks will cause the economy to diverge from its initial equilibrium level of output for a number of periods, whereas permanent supply shocks will cause the economy to move to a new equilibrium level of output.

3. What role do lags and forecasting play in monetary policy? The model reflects the real-world phenomenon that it takes time for the real interest rate to affect aggregate demand. This means that forecasting plays a large role in the central bank's setting of interest rates. For a shock in period zero, the central bank has to forecast where the Phillips curve will be next period and then set the appropriate interest rate (using the *MR* curve) to minimize their loss function. The Phillips curve also contains a backward-looking component, which means that if inflation is above target, a period of output below equilibrium (i.e. high unemployment) is required to reduce inflation back to target (and vice versa).

An important component of the core 3-equation model is the equilibrium rate of unemployment: it is possible to go straight from here to the analysis of the supply side institutions and policies that determine unemployment in Chapter 15. In this part of the book, we go on to consider a number of extensions to the model. In Chapter 4, we look in more depth at the way in which expectations about the future are formed. The role played by the banking system and how this can lead to instability and financial crises is integrated in the model in Chapters 5 and 6. The 3-equation model is extended to the open economy in Chapter 9.

3.5 Appendix

3.5.1 The 3-equation model in more detail

In the chapter, graphical analysis was used to provide a simple and intuitive explanation of the 3-equation model. In this section, we set out more carefully how the model works, focusing on the mathematical derivation of the *MR* curve. We start by deriving a more general form of the central bank's monetary rule. In setting out the equations that form the basis of the 3-equation model, we need to make explicit the timing structure. By choosing the interest rate in period zero, the central bank affects output and inflation in period 1. We assume it is only concerned with what happens in period 1. This is the reason that its loss function is defined in terms of y_1 and π_1. If we let β and α take any positive values, the central bank chooses y to minimize

$$L = (y_1 - y_e)^2 + \beta(\pi_1 - \pi^T)^2, \qquad \text{(central bank loss function)}$$

subject to

$$\pi_1 = \pi_0 + \alpha(y_1 - y_e). \qquad \text{(Phillips curve)}$$

By substituting the equation for the Phillips curve into the central bank loss function we can rewrite the loss function as:

$$L = (y_1 - y_e)^2 + \beta((\pi_0 + \alpha(y_1 - y_e)) - \pi^T)^2.$$

If we now differentiate this with respect to y_1 (since this is the variable the central bank can control via its choice of the interest rate), we have:

$$\frac{\partial L}{\partial y_1} = (y_1 - y_e) + \alpha\beta(\pi_0 + \alpha(y_1 - y_e) - \pi^T) = 0. \qquad (3.4)$$

Rearranging the Phillips curve to find π_0 and substituting this back into equation 3.4 gives:

$$(y_1 - y_e) = -\alpha\beta(\pi_1 - \pi^T). \qquad \text{(monetary rule, } MR\text{)}$$

The *monetary rule* shows the central bank's best response to a shock; it is the relationship between the inflation rate chosen indirectly and the level of output chosen directly by the central bank to maximize its utility (minimize its loss) given its preferences and the constraints it faces. The monetary rule is an inverse relation between π and y with a negative slope, which shows that the central bank must reduce aggregate demand and output, y, below y_e so as to reduce π below π^T.

In the general form of the *MR* curve shown above, it can be seen directly that the larger is α (i.e. the more responsive are wages to employment) or the larger is β (i.e. the more inflation averse is the central bank), the flatter will be the slope of the monetary rule.

In the first case this is because any reduction in aggregate demand achieves a bigger cut in inflation, i.e. whatever its preferences, the central bank gets a 'bigger bang (i.e. fall in inflation) for its buck (i.e. fall in aggregate demand)'.

In the second case, this is because, whatever the labour market it faces, a more inflation-averse central bank will wish to reduce inflation by more than a less 'hard-nosed' one.

Box 3.2 Summary of the 3-equation model

In formally deriving the *MR* curve, we now have all the components to set out the central bank's decision-making process:

1. The central bank minimizes its loss function, which expresses its objective of keeping inflation close to target, π^T

$$L = (y_t - y_e)^2 + \beta(\pi_t - \pi^T)^2. \qquad \text{central bank loss function}$$

2. subject to the constraint from the supply side, which is the Phillips curve (*PC*).

$$\pi_t = \pi_{t-1} + \alpha(y_t - y_e). \qquad \text{Phillips curve, } PC$$

3. This produces the monetary rule function (*MR*), which fixes the best response output gap $(y_t - y_e)$.

$$(y_t - y_e) = -\alpha\beta(\pi_t - \pi^T). \qquad \text{monetary rule, } MR$$

4. which is implemented through the central bank's choice of *r* using the dynamic *IS* equation, incorporating the lag from the interest rate to output

$$y_t = A - ar_{t-1}. \qquad \text{dynamic } IS \text{ curve}$$

Figure 3.15 shows the 3-equation model when the economy is in medium-run equilibrium—i.e. output is at equilibrium and inflation is constant at target. In the $PC - MR$ diagram, the economy is at the central bank's 'bliss point'—the levels of output and inflation that minimize the loss function (i.e. y_e and π^T). In the *IS* diagram, we can use the *IS* curve to read off the interest rate the central bank must set to keep inflation constant and output at y_e. This is the *stabilizing rate of interest*, r_S.

In the case when output is at equilibrium (i.e. y_e) and the interest rate is at its stabilizing level (i.e. r_S), we can rewrite the dynamic *IS* curve as:

$$y_e = A - ar_S. \qquad (IS \text{ curve, in medium-run equilibrium})$$

We can use the two versions of the *IS* equation to find a relationship between deviations of output from equilibrium and the interest rate from its stabilizing level. We do this by subtracting one from the other, such that:

$$y_t - y_e = -a(r_{t-1} - r_S). \qquad (IS \text{ equation, in deviations})$$

In period zero, this becomes,

$$y_1 - y_e = -a(r_0 - r_S), \qquad (IS \text{ equation, in deviations, in period zero})$$

which takes into account the fact that interest rate changes in period zero do not affect output until period one (i.e. the lag structure of the model). It can be useful to think of the *IS* equation in deviations form in the 3-equation model and particularly for deriving interest rate rules, which will be discussed in detail in Chapter 13.

Figure 3.15 can also give us an insight into how varying the underlying targets or parameters will affect the 3-equation model. There are six key parameters that affect the

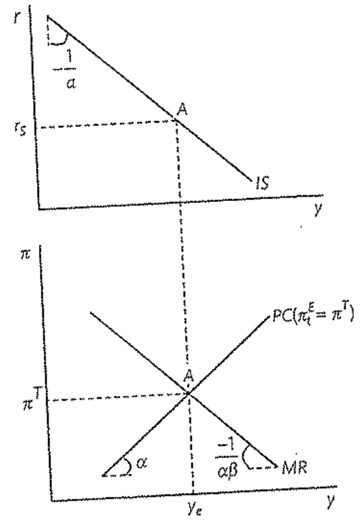


Figure 3.15 The 3-equation model in medium-run equilibrium.

underlying curves in the 3-equation model and the central bank's best response to economic shocks:[13]

1. the central bank's inflation target, π^T: this affects the position of the *MR* line;
2. the central bank's preferences, β: this determines the shape of the loss ellipses and affects the slope of the *MR* line;
3. the slope of the Phillips curve, α: this also affects the slope of the *MR* line;
4. the interest sensitivity of aggregate demand, a: this determines the slope of the *IS* curve;
5. the equilibrium level of output, y_e: this determines the level of output at which there is no pressure on inflation to change and affects the position of the *MR* line;
6. the stabilizing interest rate, r_S: the central bank adjusts the interest rate relative to r_S so it must always analyse whether this has shifted, e.g. as a result of a shift in the *IS* or due to a change in the equilibrium level of output, y_e.

3.6 Questions

3.6.1 Checklist questions

1. Why is monetary policy chosen over fiscal policy as the preferred tool for stabilization policy? What does the government gain from controlling monetary policy? Why would they choose to delegate responsibility for monetary policy to an independent central bank?

13 The Macroeconomic Simulator available from the Carlin and Soskice website can be used to test the effects of varying key parameters on the adjustment path of the economy following an economic shock. See Questions 3 and 4 in Section 3.6.2.

2. 'If the economy has high but stable inflation, the government has much to lose and little to gain by reducing inflation to a low rate.' Explain and assess this statement.
3. What are the advantages and disadvantages of a target inflation rate of 4% as compared with one of 0% per annum?
4. Explain what is meant by the central bank's loss function. How are the central bank's preferences reflected in the loss function? Draw the loss 'circles' for the cases where

   (a) $\beta = 1$;
   (b) $\beta < 1$;
   (c) $\beta > 1$.

   In which of the three cases will the central bank reduce inflation back to target quickest after an inflation shock? Is there any downside to adopting this policy stance?
5. Assume that $\alpha = \beta = 1$, derive the *MR* curve graphically using the tangencies between the loss circles and the Phillips curves. With reference to the diagrams, explain the effect of the following (in each case, assume all other parameters are held constant):

   (a) An increase in the slope of the Phillips curve, α
   (b) An increase in central bank's inflation aversion, β.
6. Following an inflation shock, explain why unemployment goes up before the economy returns to medium-run equilibrium.
7. Draw the 3-equation model and give a detailed period by period description of the adjustment process for the case where the economy is hit by a permanent negative aggregate demand shock.
8. With reference to the scenario in Question 7, explain the behaviour of the central bank and the economy in a situation where there is no lag in the *IS* curve.
9. Draw the impulse response functions for output, inflation and the real interest rate after a permanent positive aggregate demand shock and a permanent positive supply shock. [Hint: the 3-equation model diagrams for these two cases are shown in Figs. 3.11 and 3.14 respectively.]
10. Use the 3-equation model diagrams to show how the economy can fall into a deflation trap. Explain, with reference to the diagram, how the central bank/government can intervene to escape the trap. Show the relevant *IS* curve and re-label the *MR* as the *PR* to indicate that fiscal policy is being used. Are there any reasons why these policies might not work?

3.6.2 Problems and questions for discussion

1. Pick one developed economy and one emerging or developing economy. Use the latest version of the IMF World Economic Outlook Database to download inflation data for each of the countries from 1980 to the latest available data. Plot the data in a graph. Answer the following:

   (a) Describe the evolution of inflation in each country.
   (b) Do the countries have independent central banks?
   (c) If the country does have an independent central bank, did inflation fall when the central bank gained independence? Propose some reasons why.
   (d) If the country does not have an independent central bank, have they managed to find other mechanisms to establish a low inflation monetary policy regime? Propose some reasons why.

2. Select two out of the following central banks: Bank of England, Reserve Bank of New Zealand, Bank of Canada, and the Swedish Riksbank. Each of these central banks has adopted explicit 'inflation targeting'. For each of your chosen banks, answer the following questions:

   (a) What is their target level of inflation and how do they justify choosing that level?
   (b) What actions did it take following the collapse of Lehman Brothers in September 2008 and how did it explain these actions?
   (c) Did they hit the zero lower bound on nominal interest rates during the global financial crisis? If so, how could they adjust their inflation target to reduce the likelihood of this happening again in the future? Check for evidence of a public debate about this issue.
3. This question uses the Macroeconomic Simulator available from the Carlin and Soskice website http://www.oup.com/uk/orc/carlin_soskice to model a negative temporary aggregate demand shock. Start by opening the simulator and choosing the closed economy version. Then reset all shocks by clicking the appropriate button on the left hand side of the main page. Use the simulator and the content of this chapter to work through the following:

   (a) Apply a temporary 5% negative demand shock. Save your data.
   (b) Use the impulse response functions from the simulator to help explain the path of the economy following the shock.
   (c) Draw the *IS* – *PC* – *MR* diagrams for this scenario.
   (d) Adjust the central bank preferences (i.e. β) to 0.5 and save your data. Now, adjust the central bank preferences to 1.5 and save your data again.
   (e) How has varying β affected the impulse response functions? Relate this to the effect that changing β has on the *MR* curve.
4. This question uses the Macroeconomic Simulator available from the Carlin and Soskice website http://www.oup.com/uk/orc/carlin_soskice to show how the economy can get stuck in a deflation trap and what it can do to escape it. Start by opening the simulator and choosing the closed economy version. Then reset all shocks by clicking the appropriate button on the left hand side of the main page. Use the simulator and the content of this chapter to work through the following:

   (a) Apply a temporary 8% negative demand shock. Save your data.
   (b) Use the impulse response functions to help explain the path of the economy following the shock.
   (c) Apply a temporary increase in public expenditure of 7% alongside the original demand shock. Save your data.
   (d) Comment on the changes to the impulse response functions. [Hint: it might be necessary to view the second case in isolation to accurately view the movements in the impulse response functions]
   (e) Has fiscal policy been effective at stabilizing the economy? If so, explain why using the 3-equation model framework.

4 Expectations

4.1 Introduction

In this chapter, we focus on how households, firms and policy makers form expectations about the future and how this is modelled by economists. In Chapters 1–3, we touched on the importance of uncertainty and expectations in the macro-economy. We introduced forward-looking theories of consumption and investment in Chapter 1, where expectations about the future influence decision making in the current period. In Chapter 2, we saw that wage setters need to form expectations about inflation when setting nominal wage increases.

In the last chapter, we introduced the dynamic macroeconomic model, which showed how a forward-looking central bank seeks to keep the economy close to its inflation target. The 3-equation model of the macro-economy emphasizes the importance of forecasting: in order to stabilize the economy the central bank must diagnose the shocks affecting the economy and forecast their effects on the output gap and inflation. In the light of these forecasts, the central bank judges how best to set the interest rate. It continuously updates its forecasts and adjusts its policy as new information is received.

We begin this chapter by highlighting the role of expectations in the core elements of the macro model we have developed by considering in turn the three elements: the *IS* curve, the Phillips curve, and the monetary rule.

The IS curve

In the models of consumption and investment in Chapter 1, expectations about the future influence the spending decisions of households and firms. Tobin's Q theory is a very clear example of a model of investment in which there is an institutional mechanism—the stock market—that aggregates information about the views of those outside the firm about the firm's expected future profits. When Tobin's Q is greater than one, the model predicts positive investment. The stock market is signalling 'build' not 'buy' because the expected value of the firm based on expected future profits from expanding the capital stock is greater than its replacement cost.

For the household, the permanent income hypothesis encapsulates the idea that consumption decisions today reflect all the available information about expected future income: the household prefers a smooth path of consumption so it is in its interest to aggregate the available information and—if it can—consume close to its permanent income.

The economist Nick Bloom suggests that periods of very high stock market volatility occur when economic agents are very unclear about the future. There is evidence that the effects of this type of uncertainty on fixed investment decisions can be so severe that they can

actually cause the economy to go into recession (as occurred for example, after the 9/11 terrorist attacks). Another example was the extreme *policy uncertainty* in late 2012, when the USA was described as coming to the edge of a fiscal cliff. It was unclear whether the Republican-dominated Congress and the re-elected Democrat president would be able to agree on fiscal measures that would prevent tax rises and spending cuts from automatically coming into force. As a result, private agents faced serious policy uncertainty.

Bloom (2011) finds that when agents are very unclear about the future they wait and do nothing. For example, firms uncertain about future demand do not invest or hire. Similarly, consumers do not make large purchases of consumer durables if they are uncertain about their next paycheck. In times of acute uncertainty the economy therefore comes to a standstill as agents 'wait and see'. As we shall see in the next section, where we look at the concepts of risk and uncertainty, John Maynard Keynes suggested that overcoming the paralysis caused by uncertainty can come down to 'whim, sentiment, or chance'.

The Phillips curve

In the modelling of wage- and price-setting behaviour to this point, we have used backward- rather than forward-looking expectations. Wage setters use past inflation as a guide to how they expect prices to evolve over the year ahead. As in the case of those who make spending decisions, it is highly likely that wage setters think about the future. After all they are interested in setting a particular *real* wage, which means that the future movement of the prices of goods in the consumption basket is of concern. To this point, however, we have assumed that the best that wage setters can do in their wage negotiations is to get nominal wage increases that compensate them for any erosion of the real wage that occurred due to unanticipated inflation over the previous period. They use a rule of thumb that says the best forecast for the change in the consumer price index in the year ahead is last year's inflation.[1]

We can model the behaviour of wage setters as that of rational forward-looking agents who are prevented from implementing their forecast of inflation in their wage bargain because of institutional arrangements such as wage indexation or because the costs of coming to an agreement over how inflation will evolve over the coming year are too high. Alternatively, we can model the behaviour as rule of thumb where last period's inflation outturn is the 'expected inflation' agreed on by the wage setters and built into the wage bargain. Note that the Bank of England's fan chart in Fig. 4.1 suggests that, even for one year ahead, the Bank's Monetary Policy Committee sees a very wide range of possible outcomes for inflation.

Inflation forecasts by central banks and private sector forecasters have become a common feature of developed economies and can provide insight into how individuals form inflation expectations. The macroeconomic implications of inflation forecasts are limited, however, unless we know whether agents incorporate inflation forecasts into their expectations formation and the extent to which agents act on their expectations.

Armantier et al. (2011) attempt to answer this question in a particular context by carrying out a financially incentivized investment experiment to see if agents' forecasts of inflation feed into their behaviour. The experimental subjects stand to win money if they use forecast

1 Du Caju et al. (2008) provide detailed evidence on how price increases (past, past unforeseen and expected) are incorporated through formal indexation and informal practices in the wage bargains of 23 European countries, the US and Japan.

inflation when choosing between two kinds of financial investment. The authors find that most individuals tend to act on their expectations, but that the less financially and numerically literate respondents had most trouble doing so.

The MR curve

The policy maker in the 3-equation model (the inflation-targeting central bank from Chapter 3) is a forward-looking agent. As we have seen, the central bank makes a forecast about inflation in the next period and, based on this, decides the interest rate to set. This is an example of rational expectations behaviour: we have assumed that the central bank knows the model of the macro-economy and that having diagnosed the nature of any shock, it chooses its best response interest rate.

However, as we have seen, forecasting inflation is by no means easy for the central bank, due to the inherent uncertainty surrounding macroeconomic developments (and their effects on inflation). This point was echoed by Charles Bean of the UK Monetary Policy Committee (MPC) in a speech in 2005:[2]

> Uncertainty is an ever-present feature of the economic landscape that monetary policy makers cannot escape. Broadly speaking, there are three types of uncertainty that confront us on the MPC: uncertainty about the data; uncertainty about the nature and persistence of shocks; and uncertainty about the structure of the economy [in other words, uncertainty about what is the correct model].

In the next section, we look more closely at the concept of uncertainty and contrast it with the concept of risk. We provide examples of how expectations are formed in real-world situations under conditions of risk and uncertainty and introduce the rational expectations hypothesis. Section 4.3 focuses on how different hypotheses about inflation expectations are reflected in the Phillips curve. Section 4.4 contrasts adaptive and rational expectations in the 3-equation model and discusses the influence of central bank communication on the formation of inflation expectations. In Section 4.5, we explain the Lucas critique and why poor economic outcomes can result if policy makers fail to anticipate how the private sector will respond to changes in policy. This discussion is extended in Section 4.6 where the concept of inflation bias is introduced.

4.2 Risk, uncertainty and expectations

4.2.1 Risk and uncertainty

Forming expectations about the future is a vital part of economic life. People form expectations about what they think will happen in the future when they decide how much to spend or to save, whether or not to stay a further year in higher education, whether to rent or buy an apartment, whether this is the right time to have a child, to retire and so on. A business pays a lot of attention to how it believes the market for its products will develop and how its competitors will respond when it decides on the price to set. When deciding whether to accept a

2 This is an excerpt from Charles Bean's speech to Oxonia on 22nd of February 2005 entitled *Monetary Policy in an Uncertain World*.

job offer, we take a view about how we think wages and conditions, including prospects for getting promotion, will evolve in this job compared with those in an accessible alternative. Our ability to think about the future, combined with the fact we do not know exactly what will happen in the future, makes the study of expectations of central importance to economics.

While we can never know exactly what will happen in the future, in some cases we can make a much more informed prediction than in others. This matters when thinking about how households, firms and policy makers form expectations. It is therefore necessary to define some precise terminology before we can properly discuss expectations. The essential distinction we need to make is that between *risk* and *uncertainty*.

Risk exists when individuals make decisions about the future based on *known* probabilities. In such conditions agents can work out an expected outcome. For example, a group of ten people all put \$10 into a pot and a winner is randomly selected to receive the total contents of \$100. Before the random selection is made, the expected payoff to each participant is 0.1 (their probability of winning) multiplied by \$100 (the prize for winning), which is equal to \$10. In this example there is risk, given that each individual will receive either \$0 or \$100 after the random selection is made. But they know beforehand the probability with which each possible outcome (winning or not winning) will occur.

Uncertainty differs from risk in two important ways: first, uncertainty exists where it is impossible to assign probabilities to known outcomes and, second, where there are some outcomes which may be unknown.

John Maynard Keynes, one of the most influential economists of the twentieth century, believed some things were inherently uncertain and that we could not attach any meaningful probabilities to them.[3] He summed up uncertainty in this widely cited excerpt from a journal article that he wrote in 1937:[4]

> the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system of 1970. About these matters there is no scientific basis on which to form any calculable probabilities whatever. We simply do not know.

Whether we are considering risk or uncertainty matters greatly for the way we analyse how expectations are formed. Households facing risk can accurately attach probabilities to different future states of the world, which makes it much easier for them to evaluate in the present, decisions which affect the future. But in the case of uncertainty, this is not possible, because totally unforeseen events, such as the Arab Spring of 2010–11 or a major natural disaster, may occur.[5]

Risk

Even in the case of risk, decision models are far from perfect and might attach inaccurate probabilities to the occurrence of a given event. A recent example is the risk management

3 This view was first emphasized by Frank Knight (1921) and John M. Keynes (1921).

4 Excerpt from page 214 of Keynes (1937).

5 See Peter P. Wakker's article on *Uncertainty* in The New Palgrave Dictionary of Economics: 2nd Edition (2008) for a summary of the evolution of economic thinking on risk and uncertainty. It contains a discussion of a number of decision models that can be used under unknown probabilities.

models used by investment banks just prior to the global financial crisis. The models grossly underestimated the probability of a nationwide fall in house prices in the US. When house prices did fall, this triggered the worst financial crisis since the Great Depression (see Chapter 7).

Up to a point risk can be understood and compensated for. The academic discipline of economics has become more sophisticated and mathematical over time. Econometrics, particularly regression analysis, can shed light on relationships between economic variables. The results of regression analyses are frequently represented in regression tables showing the standard errors associated with the estimates of the constant term and the coefficients.

Soyer and Hogarth (2012) report the results of a survey of 257 academic economists. The survey was conducted to find out how well economists can make probabilistic inferences given different presentations of data from a linear regression analysis. The authors find that when results are presented in a regression table, economists view outcomes as more predictable than can be justified by the model due to a failure to take adequately into account the error term. Respondents place 'too much' weight on the precision of the estimated coefficients. The bias is only removed when scatter plots of the data are shown to respondents without regression tables. The example shows that when faced with known probabilities (i.e. risk), individuals are not going to be capable of making fully informed decisions unless they are able to accurately interpret probability distributions. If even professional economists are guilty of misinterpreting probability distributions, then it is likely that such misinterpretion is a regular occurrence in the macro-economy in general.

Uncertainty

There are few aspects of the macroeconomic future that can truly be reduced to known probabilities. In some scenarios, however, meaningful reduction to probabilities seems more possible than in others; we are not operating in a world of total uncertainty. Economic forecasters and policy makers frequently have to straddle the divide between risk and uncertainty. They cannot accurately attach probabilities to future outcomes, but can attach subjective probabilities (that is, their own opinions) to different scenarios.

This approach is exemplified by the Bank of England's behaviour in relation to its target, which is to achieve an annual rate of inflation of 2% two years in the future. Figure 4.1 is a fan chart for the annual change in the Consumer Price Index (CPI), the Bank's chosen measure of inflation. It is taken from the Bank of England Inflation Report, which is published four times a year. The fan chart shows the (subjective) probability of various outcomes for inflation in the future. The dotted vertical line on the chart marks two years from the date of publication, which is the inflation rate the Monetary Policy Committee (MPC) is mandated to target. The shaded bands refer to the likelihood of observing inflation of that level if the economic conditions at the time of publication were to prevail on 100 different occasions. The Bank of England MPC's best collective prediction is that inflation will fall within the darkest central band of the fan on 10 out of 100 occasions.

The probability mass of each identically shaded pair of bands on either side of the central band sum to 10%. In any given quarter, inflation over the forecast period is expected to lie within the fan on 90 out of 100 occasions and anywhere outside the fan on 10 out of 100 occasions. The 10% probability that inflation will fall outside the fan reflects uncertainty. In

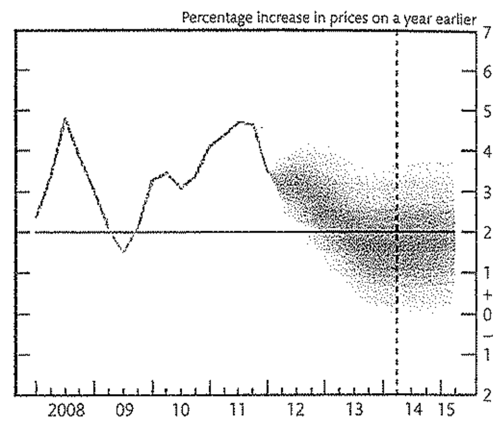


Figure 4.1 Bank of England's CPI inflation projection based on market interest rate expectations and £325 billion asset purchases.

Source: Bank of England *Inflation Report*, May 2012.

other words, there is a 10% chance that something unexpected will happen—e.g. a dramatic spike in world oil prices due to a war in the Middle East—that will move inflation outside the forecast fan. The problem of grappling with risk and uncertainty is highlighted in the way the Bank of England handled the Eurozone crisis when creating its fan charts. It stated in its August 2012 Inflation Report (p.39):

> the MPC's fan charts exclude the most extreme outcomes associated with developments in the euro area, but the possibility of such outcomes crystallizing is expected to continue to weigh on asset markets and confidence, and these influences are included in the fan charts.

The global financial crisis was not foreseen by the majority of economists, central bankers or investment bankers. In this respect, it was an uncertain event; before the crisis, virtually no probability was being attached to a future state of the world which included a near collapse of the global financial system and a global recession.

The unexpected behaviour of the British economy is reflected in the poor quality of macroeconomic forecasts during this period. For example, in the 12 quarterly inflation reports between August 2007 and May 2010, the Bank of England's average forecast for inflation a year ahead was 1.9%, and outturn inflation was 3.2% per annum. This represents an average error of +1.3 percentage points. This compares with an average error of just +0.1 percentage points between August 2001 and May 2004 (which can be considered 'normal times').[6]

A second example of increasing uncertainty in the wake of the financial crisis is the activity of the the Office for Budget Responsibility (OBR), a fiscal policy council set up by the UK government. The OBR is responsible for providing independent macroeconomic and fiscal forecasts to underpin the government's annual budget. The OBR forecasts are also used to

[6] See Bourne (2011).

determine whether the government is on target to meet its fiscal objectives; they rely on the OBR's estimate of potential output (that is, in our notation, y_e). Potential output cannot be observed, even for historical periods, so is inherently uncertain and subjective. Nevertheless, an estimate of potential output is crucial to whether the government will meet its fiscal objectives. An incorrect estimate of the output gap could contribute to the pursuit of an entirely different path of government spending over the following five years. The estimate clearly has large consequences for the real macro-economy. The example highlights that important policy decisions, with significant economic consequences, are based on forecasts that are deeply uncertain in turbulent times.

Examples of the pervasiveness of uncertainty in the macro-economy do not mean that households and firms fail to benefit from forming expectations about the future. Decisions have to be made in the present and it is integral to such decision making that people invest time and effort in forming expectations. They do this as well as they can, taking into account that unanticipated events are bound to occur. The complicated nature of forming expectations in the face of uncertainty is well articulated by Keynes:[7]

> We are merely reminding ourselves that human decisions affecting the future, whether personal or political or economic, cannot depend on strict mathematical expectation, since the basis for making such calculations does not exist; and that it is our innate urge to activity which makes the wheels go round, our rational selves choosing between the alternatives as best we are able, calculating where we can, but often falling back for our motive on whim or sentiment or chance.

4.2.2 Expectations formation in real-world situations

This section looks at some examples of how agents form expectations in real-world situations. The examples highlight that expectations are a key input into economic decision making in a wide range of contexts, and that in the real world, even the same agents may not form expectations in the same way when making decisions in different situations.

For instance, students use textbooks. Some will buy them new, some will buy them second-hand and some will not buy them at all, preferring to use a copy from the library. We might not always think of expectations playing a major role in the decision to purchase something as commonplace as a textbook, but the evidence suggests that they do. Chevalier and Goolsbee (2009) investigated the textbook market in the US for the ten semesters between 1997 and 2001. They find that students are less willing to pay for (or buy at all) books which have a short projected life. The reason for this is that when a new edition is released, older editions cannot be resold to the campus bookstore or online. The study suggests that publishers cannot consistently raise revenue by shortening revision cycles and that the market for textbooks in the US is characterized by forward-looking, rational behaviour.

Chevalier and Goolsbee's research illustrates one situation in which students' behaviour was shown to be consistent with what is called rational expectations behaviour: they were using all the information available about the costs and benefits of different courses of action to make a reasoned decision about whether to purchase a textbook, including a model of

[7] See Chapter 12: *The State of Long-Term Expectation* in Keynes (1936).

publishers' behaviour. But in other contexts students have been shown to form expectations in a biased manner. Foster and Frijters (2012) analyse how students form expectations concerning their own final course grade. The authors examined a large cross section of students at two Australian universities, collecting data on students' expectations at mid-semester of their final grade and their actual final grade. The results show that students typically over-estimate their final grade, even though it should be fairly predictable based on the information available to them, such as past performance, ability and effort levels. The authors find statistical evidence that this bias arises from the direct utility students gain from holding high expectations. The high self-esteem is associated with higher expectations.

In financial markets, profits and ultimately livelihoods can depend on having accurate expectations. The trader who accurately picks the stock which outperforms the general index is likely to reap a financial reward from that decision. But how do money managers form expectations? Psychoanalyst David Tuckett does not believe that the behaviour of money managers can be described as rational since they operate in a world of 'radical uncertainty' and ambiguous information. He conducted a detailed survey of money managers and found that they are able to commit to making decisions under conditions of radical uncertainty by creating narratives, or trading stories, about the fundamental facts thought likely to influence the price of financial assets.[8] In this world, it is not that agents do not expend effort and take into account all available information when forming expectations. Rather, because future asset prices are extremely unpredictable in the buying and selling of stocks, they do not behave completely rationally.

4.2.3 The rational expectations hypothesis (REH)

The previous section showed that in some circumstances, agents make decisions in a rational forward-looking manner (e.g. the US textbook market). In order to formalize this approach to forming expectations, we set out the *rational expectations hypothesis (REH)*:

> The rational expectations hypothesis refers to a choice by the economist about how to model behaviour. A rational expectations model is one where the *agents in the model* use the model and all available information to forecast and therefore do not make systematic errors. The term model-consistent expectations can be used instead of rational expectations.

How does this compare to the approach taken in the 3-equation model in Chapter 3? Up to this point, we have assumed that wage setters form inflation expectations in a backward-looking manner (i.e. adaptively). In a period of rising inflation, agents using adaptive expectations make systematic errors, because inflation turns out every period to be higher than had been expected.

This is illustrated graphically in a situation of ever-increasing inflation when output is above equilibrium (see Table 2.1 and Fig. 2.15a in Chapter 2). As inflation increased from 2% to 4% to 6%, wage setters were repeatedly wrong in their anticipation of a higher real wage (on the *WS* curve) when the output gap was positive. By maintaining the assumption that expectations were formed adaptively, we did not allow agents to learn from their mistakes

[8] See Tuckett (2012).

and *change the rule* they used. Their expectations formation behaviour is therefore not model consistent.

No one can know the correct model of the real world so no one can be fully rational in the way that agents in a model can be. We use models to help us understand the world. In the world, forming expectations rationally as agents can do in a model is not possible. Apart from the presence of uncertainty, and the fact that unlike agents in a model, those in the world do not know the model, forecasting involves time and cognitive effort. These considerations mean that the usefulness of a rational expectations-based model for understanding particular situations depends on the setting.

In our example of students and the textbook market, the evidence suggests that students were able to figure out the model that was being used for the frequency with which new editions were being introduced, and they responded by only buying in the cases where a longer edition cycle was in place. In this case, the behaviour of the students fitted the prediction of a model with agents forming rational expectations. In the case of students forecasting their grades, however, their ability to make rational judgements was impaired because of the utility they (irrationally) derived from forming over-optimistic expectations.

In the world, we can observe that some agents devote considerable resources to forecasting. For example, central banks build, test and refine their models of the economy. Their models and forecasts are subject to extensive external scrutiny and it is unlikely that the central bank would make systematic mistakes without being forced to reconsider its methods. But making systematic mistakes is quite different from being repeatedly wrong in its forecasts.

The earlier quote from Charles Bean of the Bank of England highlights the difficulties that the central bank faces in spite of all of the resources it has available to understand the nature of the shocks and the structure of the economy. Households face larger obstacles. Economics students spend years trying the understand the 'predictions of the relevant economic theory'. Rules of thumb (such as adaptive expectations) may provide a better description than rational expectations of the behaviour of a range of private sector agents.

As we shall see, models with rational expectations provide powerful insights about the dangers of neglecting how individuals and households respond to changes in the economic or policy environment. No rule of thumb is of much use if the environment changes in such a way as to render it obsolete from the point of view of the agent.

4.3 Phillips curves, expectations and inflation

This section investigates in more detail how economists have modelled the way agents in the macro-economy form inflation expectations. In the 3-equation model of Chapter 3, we used a specific assumption about inflation expectations, which produced the Phillips curve that we used in that model. As we shall see, the Phillips curve will take different forms depending on the way inflation expectations are modelled.

Notation and timing

We start by setting out the notation that will be used throughout the remainder of this chapter. We are primarily concerned with how agents at the beginning of period t form expectations of the current period's inflation. The current period is denoted by a t subscript. These expectations will be formed on the basis of the information available at the beginning of period t—we say expected inflation in period t, conditional on the information set (denoted by Ω) dated $t-1$, is therefore:

$$E_t(\pi_t \mid \Omega_{t-1}) \equiv \pi_t^E.$$

We use π_t^E to denote expected inflation in period t (conditional on the information set in period $t-1$) throughout the rest of the chapter as it makes the notation simpler. The information set in period t includes the outturn of all the macroeconomic variables up to and including period $t-1$. This means that, in period t, agents know the actual rate of inflation for every period up to and including period $t-1$ (i.e. the last period). As we shall see, when agents have rational expectations, then apart from a random shock that occurs in period t itself, they also know inflation in period t because they can work out the implications for inflation of the shocks that have occurred up to and including period $t-1$. The only thing that could make inflation in period t different from what they calculate is an unpredictable shock to inflation within period t itself.

Phillips curves

The term Phillips curve is used for the relationship between inflation and output arising from the supply side of the economy. There is no single form for the Phillips curve, but rather, it depends on the way we model inflation expectations. To allow us to discuss different hypotheses about expectations, we write the Phillips curve like this:

$$\pi_t = \pi_t^E + \alpha(y_t - y_e). \quad \text{(standard Phillips curve)}$$

We shall call this a standard Phillips curve.

In this section, we set out two ways in which inflation expectations are modelled.

1. *Adaptive* expectations—agents expect inflation to be what it was in the previous period. This is a simple form of what is called error correction behaviour. Agents set aside their forecast of inflation last period and take last period's outturn as their best guess of inflation this period. This is the model we use most of the time in this book.
2. *Rational* expectations—agents in the model use the model to form their forecasts. This implies expectations are forward looking, all available information is used and systematic errors are avoided.

The rest of the section is divided into two parts. First, we look at inflation in the UK since the end of the Second World War and relate this to the modelling framework. Second, we set out the adaptive and rational expectations Phillips curves and discuss their implications for the trade-off between output and inflation.[9]

[9] We do not discuss the New Keynesian Phillips curve in this chapter. It is based on the assumption that firms are fully forward-looking with rational expectations but are constrained from flexibly adjusting their prices. This case is examined in Chapter 16.

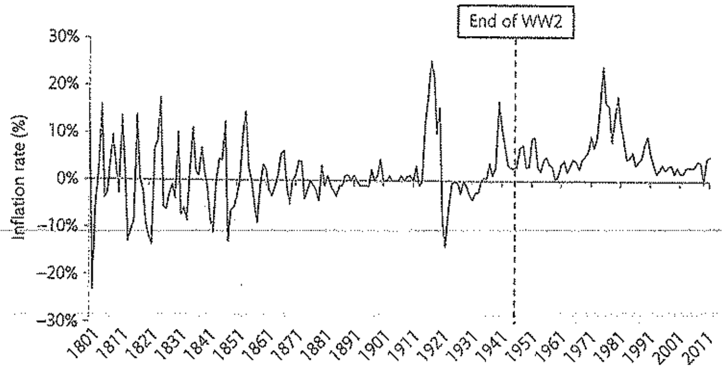


Figure 4.2 UK inflation: 1801–2011.

Source: UK Office for National Statistics (accessed June 2012).

Note: Variable shown is the annual percentage change of the long term indicator of prices of consumer goods and services.

Inflation in the post-war period

We can see from Fig. 4.2, that inflation in the UK fluctuated around zero between 1801 and the end of World War II (WWII). In the 1950s and 1960s, however, inflation was positive in almost every year but there was no trend. This was followed by a period of high inflation from the end of the 1960s through to the early 1980s. Disinflation followed.

One interpretation of the period of stable but positive inflation in the 1950s and 1960s is that governments were pursuing activist demand management policies and were happy to achieve low rates of unemployment even if there was some positive inflation. This can be understood against the background of the Great Depression in the 1930s and the awareness of policy makers of the dangers of deflation. At the time, policy makers feared unemployment and deflation more than inflation. Their policies reflected this, yielding consistently positive, although fairly stable, inflation.

However, toward the end of the 1960s, inflation began to rise year after year. One part of the explanation is that important developments on the supply side of the economy in the late 1960s and the 1970s led to a rise in equilibrium unemployment, which meant that conventional measures of 'high employment' implied more inflationary pressure. In this period there was a shift in bargaining power toward workers reflected in a wave of strikes across Europe referred to as the 'hot autumn' of 1968 (an upward shift in the *WS* curve) and a fall in the rate of productivity growth (which implies a downward shift in the *PS* curve, assuming wage-setting behaviour does not adjust quickly to the slower productivity growth).[10] As well as this, there was a series of commodity price shocks (which imply downward shifts

10 Productivity growth slowed for two reasons. First because the gains from the widespread adoption of so-called Fordist mass production methods in the catching up countries of Western Europe and Japan were becoming exhausted. And second, the limits to further productivity gains based on Fordism were being reached even in the technology leader, the USA.

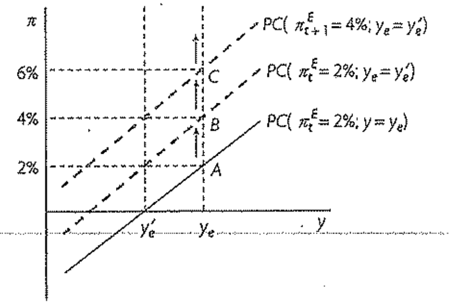


Figure 4.3 The government attempts to maintain low unemployment after a negative supply-side shock.

in *PS* curve) preceding the first oil shock of 1973. The oil shocks are discussed in detail in Chapter 11.

We can interpret the new inflationary behaviour using the model of Chapter 2. As an example, we take an economy initially at point *A* in Fig. 4.3 with inflation stable at 2% and output at equilibrium, y_e. The supply-side shocks reduce equilibrium output to y'_e. If the government continues to pursue an unchanged level of output, inflation will rise continuously.

Inflation expectations are updated to correct for the erosion of the expected real wage.[11] Under this hypothesis, wage setters update their expectations based on the outturn of inflation. Workers do this to safeguard their real wages, which are otherwise eroded in every period where inflation is above its expected level. The *PC* therefore shifts upwards as a result of agents updating their expectations (Fig. 4.3).

As long as the government keeps $y > y'_e$, inflation expectations will adjust upwards as wage setters see the expected real wage eroded. With $y = y_e$, the shifted *PC* results in inflation of 4%. The only outcome of the government continuing to pursue an output target above equilibrium will be ever-increasing inflation, as shown by the movement from *B* to *C* (and onwards) in Fig. 4.3. A government policy of this kind is referred to as an accommodating policy, effectively, there is a vertical *MR* curve through points *A*, *B* and *C*. It contrasts sharply with the behaviour of an inflation-targeting central bank as set out in the 3-equation model.

Expectations and the output–inflation trade-off

In the adaptive expectations framework, workers respond to their forecasting errors—i.e. they update their expectations based on the outturn for inflation last period. We modify the standard Phillips curve to get the adaptive expectations Phillips curve as follows:

$$\pi_t = \pi^E_t + \alpha(y_t - y_e)$$
$$= \pi_{t-1} + \alpha(y_t - y_e), \qquad \text{(adaptive expectations Phillips curve)}$$

where $\pi^E_t = \pi_{t-1}$.

11 The pioneering work on adaptive expectations was carried out by Nobel Prize winners Edmund Phelps and Milton Friedman in the late 1960s and early 1970s. See, for example, Phelps (1967) and Friedman (1968).

The equation can be rearranged to show how inflation changes over time:

$$\pi_t - \pi_{t-1} = \Delta\pi_t = \alpha(y_t - y_e). \tag{4.1}$$

Equation 4.1 shows that inflation will continue to increase as long as output remains above equilibrium, which matches the outcome shown in Fig. 4.3.

The model predicts there is no long-run trade-off between inflation and unemployment. The presence of a trade-off would mean that policy makers could choose to operate the economy at a level of output above equilibrium if they were willing to accept a higher level of inflation. In other words, they could choose their preferred point on a single Phillips curve. In the model, this is not possible, because there is always pressure on inflation to rise (and the *PC* to shift upwards) if output is above equilibrium.

Using this model, we could say that the low and stable inflation of the 1950s and 1960s shown in Fig. 4.2 only held as long as the government did not systematically try to run the economy at a level of activity higher than the equilibrium. The supply shocks of the late 1960s and the 1970s led to governments running the economy with levels of aggregate demand that produced levels of output above equilibrium.

Nobel prize winner Milton Friedman described the prevailing postwar policy stance in his famous 1968 Presidential Address:[12]

> Today, primacy is assigned to the promotion of full employment, with the prevention of inflation a continuing but definitely secondary objective.

He then neatly sums up the inflationary consequences of pursuing this policy and acknowledges that there is no long-run trade-off between unemployment and inflation:[13]

> In order to keep unemployment at its target level of three per cent [which in this hypothetical example is below the 'natural' rate], the monetary authority would have to raise monetary growth still more.... the 'market' rate [of unemployment] can be kept below the 'natural' rate only by inflation...
>
> To state this conclusion differently, there is always a temporary trade-off between inflation and unemployment; there is no permanent trade-off.

Figure 4.4a uses US data to illustrate the reasonably stable relationship between inflation and output in the 1950s and 1960s. The relationship completely fell apart during the following 15 years (see Fig. 4.4b).

In many developed economies, the 1970s were characterized by 'stagflation'—a combination of high unemployment and high inflation—which destroyed any notion of a long-run trade-off between inflation and output. The experiences of this decade marked the end of a period where macroeconomic policy had been guided by the principles of Keynesian demand management which accorded priority to avoiding deflation. In macroeconomics, it ushered in the era where models were built based on agents using rational expectations to form their forecasts and guide their actions.

12 These excerpts are from Friedman (1968).

13 Friedman uses slightly different terminology than we have so far in this chapter. The 'natural rate' of unemployment is equivalent to the equilibrium rate of unemployment. The 'market' rate is equivalent to the actual unemployment rate.

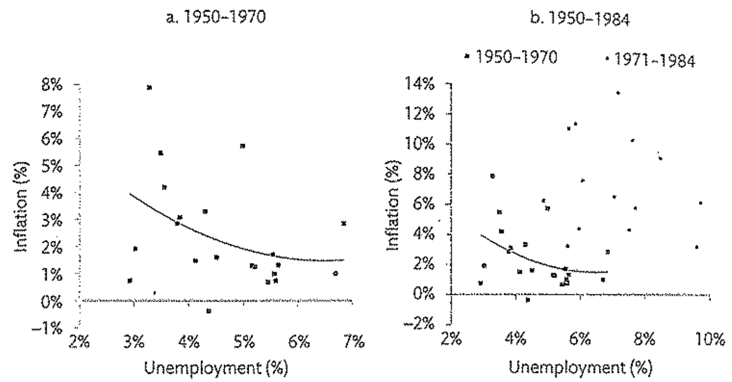


Figure 4.4 US empirical Phillips curves:
a. 1950–1970
b. 1950–1970 and 1971–1984.
Source: US Bureau of Labor Statistics (accessed June 2012).
Note: The black curves on the graphs are quadratic regression lines.

The adaptive expectations hypothesis states that wage setters respond to their forecasting errors but they do this in a purely mechanical way. This involves them making *systematic errors* because as long as $y > y_e$, the real wage outcome is below what had been expected since the inflation outturn is higher than expected (see Fig. 4.3). This suggests we investigate the implications of wage setters not just updating their *expectations* but changing their *forecasting rule*. This is the basis of the next expectations hypothesis, rational expectations.

In the rational expectations framework, it is only unanticipated shocks to inflation and the output gap that result in inflation being different from its expected value. If we apply the rational expectations hypothesis (REH) to the standard Phillips curve it becomes:

$$\pi_t = \pi_t^E + \alpha(y_t - y_e) + \epsilon_t, \tag{4.2}$$

where

$$\pi_t^E = \pi_t$$

$$\rightarrow y_t = y_e - \frac{\epsilon_t}{\alpha} \quad \text{(rational expectations Phillips curve)}$$

and ϵ_t is a random shock term. In this model of behaviour, expectations are correct apart from the random shock term. It is crucial that the error process is of a particular kind, namely 'mean zero white noise' in order to rule out systematic errors. A mean zero white noise shock is sometimes positive and sometimes negative, but on average it is zero. We can see that applying rational expectations to the standard Phillips curve produces the result that there is no longer a trade-off at all between inflation and the output gap.

4.4 Expectations and the 3-equation model

In this section, we look at how inflation expectations influence the 3-equation model. We contrast the 3-equation model with backward-looking inflation expectations introduced in the last chapter with the 3-equation model with rational inflation expectations. This allows us to show how the key predictions of the 3-equation model are altered by the use of rational inflation expectations. We end the section with a discussion of how modern central banks use their communication strategies to try to anchor inflation expectations close to the inflation target.

4.4.1 A graphical example

To see the implications of rational expectations in the 3-equation model, we take a simple example in which the central bank adopts a new lower inflation target, π^T. In Fig. 4.5a (i.e. the left-hand panel of Fig. 4.5), the economy is initially at point A, with output at equilibrium, but inflation above the new target. To minimize their loss function the central bank would like to reduce inflation to π^T, moving the economy from A to Z. This implies a new MR curve, labelled MR'. We assume they announce this policy change at the start of period $t+1$. Under rational expectations, and assuming the announcement is credible, the economy will jump immediately (i.e. within period $t+1$) from A to Z. We can see this using the rational expectations PC in Equation 4.2. The disinflation is *costless*, in the sense that it has not involved any rise in unemployment.

As we have seen in Chapter 3, if expectations are formed adaptively, then the disinflation would not be costless. In this case, the policy announcement would be coupled with action by the central bank to tighten monetary policy by raising the interest rate so as to move the economy to point B in Fig. 4.5b. It would then take a further $x-1$ periods of active adjustment of the interest rate before the economy reached point Z and the Phillips curve had shifted to the new equilibrium at $PC(\pi^E_{t+x} = \pi^T)$. This process would result in unemployment being above equilibrium for some time—i.e. it would be costly.

4.4.2 Comparison of adaptive and rational inflation expectations

To help make the difference between the adaptive and rational expectations assumptions clear, it is useful to apply rational expectations to the 3-equation model and compare the results with the familiar behaviour of the model with adaptive expectations. In the 3-equation model set out in Chapter 3, we assumed wage setters set the nominal wage increase according to the WS curve and price setters followed immediately by marking up their prices. To summarize:

$$(\Delta W/W)_t = \pi^E_t + \alpha(y_t - y_e),$$

where

$$\pi^E_t = \pi_{t-1}, \qquad (4.3)$$

$$\pi_t = (\Delta W/W)_t, \qquad (4.4)$$

$$= \pi_{t-1} + \alpha(y_t - y_e). \qquad \text{(PC in adaptive expectations 3-equation model)}$$

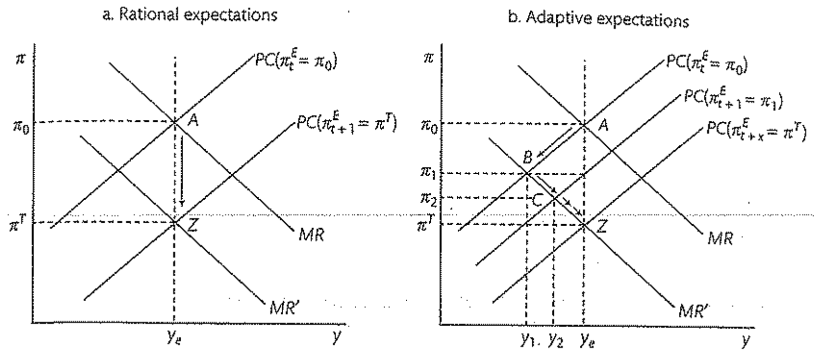


Figure 4.5 The effect on output and inflation of the central bank reducing the inflation target from π_0 to π^T.
a. Rational expectations.
b. Adaptive expectations.

If instead of the adaptive expectations assumption, $(\pi^E_t = \pi_{t-1})$, wage and price setters use rational expectations, then for expectations of inflation to be fulfilled, i.e.

$$\pi_t = \pi^E_t,$$

output must be at equilibrium, i.e. $(y_t = y_e)$. The implication when $y_t \neq y_e$ is that there is no value of inflation that satisfies the equation for fulfilled expectations, $\pi_t = \pi^E_t$; hence inflation will be infinite.

$$\pi_t = \pi^E_t + \alpha(y_t - y_e) + \epsilon_t$$

$$\Longrightarrow \pi^E_t < \infty \text{ if and only if } y_t = y_e \text{ and } \pi^E_t = \infty \text{ if and only if } y_t \neq y_e.$$

To see why inflation must be infinite when $y_t \neq y_e$, we take the case where output is slightly above equilibrium output, so that the term $\alpha(y_t - y_e) = \varepsilon$, a small positive number. Under the rational expectations hypothesis, both sides of the Phillips curve have to be equal in expectation, hence $\pi_t = \pi^E_t + \varepsilon$. The only value of inflation that satisfies this equation and where inflation expectations are fulfilled (i.e. $\pi_t = \pi^E_t$) is when $\pi_t = \pi^E_t = \infty$.

By assuming rational expectations behaviour of wage and price setters in the 3-equation model, it is clear that the central bank must set the real interest rate at the start of period t to deliver $y_t = y_e$ in expectation. Otherwise inflation is expected to be infinite. This in turn implies that the IS curve must take the form:

$$y_t = A - ar_t, \qquad \text{(IS curve in the REH 3-equation model; no lag)}$$

where there is no lag between interest rate setting and output. Households and firms are assumed to know the model and to behave as such: this eliminates the lags in the IS equation because they know that output must be at equilibrium for inflation to be other than infinite.

The next step is to pin down the level of inflation expected by wage and price setters. The core idea is that rational expectations behaviour on the part of wage and price setters means that they understand how the model works. Hence, they know that the central bank chooses y_t where the *MR* curve intersects the Phillips curve, $PC\,(\pi_t^E)$. They choose their inflation expectations, π_t^E at the start of period t so those expectations are consistent with the outcome for inflation, π_t. They do not make systematic mistakes. Random shocks will occur and mean that the outturn for inflation can differ from what was expected, but the key is that these deviations of outturns from expectations are unsystematic. To save notation, we do not show the error term ϵ_t in the Phillips curve equation.

Wage and price setters know that π_t is generated by the following two equations, $PC\,(\pi_t^E)$ and *MR*:

$$\pi_t = \pi_t^E + \alpha(y_t - y_e) \qquad \text{(PC in REH 3-equation model)}$$
$$y_t - y_e = -\alpha\beta(\pi_t - \pi^T), \qquad \text{(MR in REH 3-equation model)}$$

where π^T is the inflation target of the central bank. In a diagram, this implies that the Phillips curve goes through the *MR* at $\pi = \pi^T$. In terms of equations, the wage and price setters know that

$$\pi_t = \pi_t^E - \alpha^2\beta(\pi_t - \pi^T)$$

and that for their inflation expectations to be fulfilled, i.e. $\pi_t = \pi_t^E$, they have to choose

$$\pi_t^E = \pi^T.$$

The rational agents in the economy know that the central bank is targeting inflation of π^T, so this is the level of inflation that they expect. The central bank knows that this is how inflation expectations are formed and hence that the relevant Phillips curve it faces is

$$\pi_t = \pi^T + \alpha(y_t - y_e).$$

Putting this together with the *MR*, the central bank knows that the output gap it wishes to set is

$$y_t = y_e.$$

Graphically, as noted above, this is the intersection of the $PC(\pi_t^E = \pi^T)$ and the *MR*.

Given all of this information and the fact that if there is a random shock in period t that none of the actors can know about when they make their decisions at the beginning of period t, the central bank simply uses the *IS* curve (with no lags) to set (or maintain) the interest rate, r_t, as follows:

$$y_t = A - ar_t = y_e.$$

The central bank sets the interest rate at its stabilizing rate.

The assumption of rational expectations on the part of all of the actors in the model (and the associated assumption of no lags in the *IS* curve) produces important changes in the 3-equation model:

1. The economy is—apart from random shocks—at equilibrium output with inflation at target. The dynamic behaviour of the economy in response to different kinds of shocks analysed in Chapter 3 disappears. Inflation does not get 'built into the system' and, therefore, there is no costly process of disinflation to return the economy to target inflation and equilibrium output.
2. The job of the central bank is much simpler. It does not have to take account of the lags in the response of aggregate demand to its interest rate decisions (in the *IS* equation) to work out its interest rate response. In contrast to the 3-equation model with adaptive expectations, there is no role for a stabilizing central bank that guides the economy along the *MR* curve to equilibrium.
3. Since wage and price setters are forward looking, the central bank can influence expectations directly. By contrast, with adaptive expectations, actual inflation has to fall before it influences expected inflation.

4.4.3 Central bank communication and anchoring inflation expectations

The modern, inflation-targeting central bank spends a large amount of time and effort on communication. It is hard to explain this if the world is well-explained by the rational expectations model of the previous section. Central banks do not limit their communication to informing the public and the financial markets of their interest rate decisions. They will typically also provide: macroeconomic commentary and forecasts, a monthly or quarterly inflation/monetary policy report, which sets out the detailed reasoning behind interest rate decisions, minutes from committee meetings, transcripts of speeches by prominent members of the central bank, a statistics hub containing data on historic interest rates and exchange rates and the financial sector, in-house economic research on monetary economics and financial markets; and games (e.g. ECB) and competitions (e.g. Bank of England) involving monetary policy.

The aim is to keep inflation expectations anchored at the inflation target. If the central bank's commitment to the inflation target is perfectly credible and inflation expectations are firmly anchored, then an inflation shock would just be a one-period shock to inflation—i.e. there would be no need for a painful increase in unemployment to bring inflation back to target. In this case, in the period following the shock, the Phillips curve would revert to the one indexed by π^T and as a result, the central bank has to do nothing.

We can capture this idea in a simple way by modifying the adaptive expectations Phillips curve to model expected inflation as a weighted average of the inflation target and lagged inflation, where the weight on the inflation target is called χ (chi, for credibility):

$$\pi_t = \left[\chi\pi^T + (1-\chi)\pi_{t-1}\right] + \alpha(y_t - y_e) \qquad (4.5)$$

where

$$\pi_t^E = \chi\pi^T + (1-\chi)\pi_{t-1}. \qquad \text{(partially anchored expectations)}$$

For a fully credible central bank, $\chi = 1$ and the Phillips curve is anchored to π^T. For $\chi = 0$, inflation expectations are entirely backward looking. In Fig. 3.9 in Chapter 3, we set out the case when $\chi = 0$, as inflation inertia is a key assumption of the standard 3-equation model. We can use the same kind of diagram to set out two other cases (see Fig. 4.6), firstly where expectations are firmly anchored to the inflation target (i.e. $\chi = 1$) and secondly where expectations are only partially backward looking (i.e. $\chi = 0.5$).

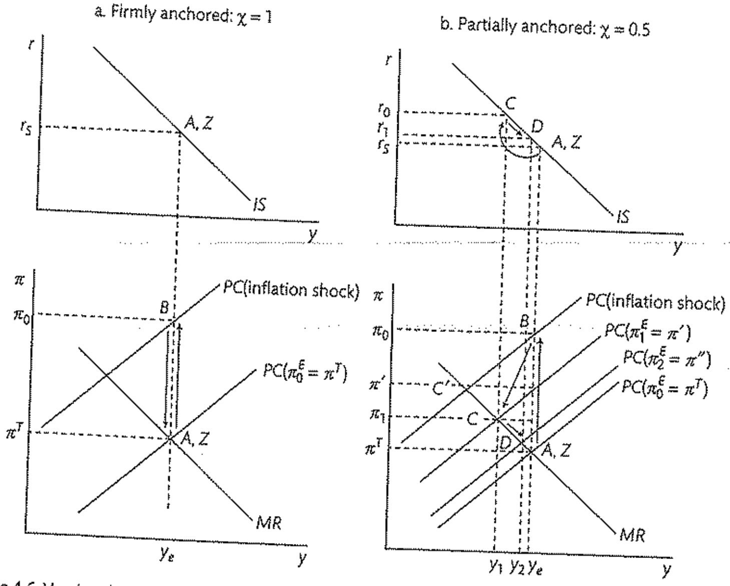


Figure 4.6 Varying the level of central bank credibility, using the example of an inflation shock:
a. Firmly anchored: $\chi = 1$
b. Partially anchored: $\chi = 0.5$.

Figure 3.9 in Chapter 3 and Fig. 4.6 show the three benchmark cases following an inflation shock and allow us to compare the adjustment of the economy back to medium-run equilibrium:

$\chi = 0$: *fully backward-looking*

After the inflation shock, the central bank must raise interest rates to squeeze inflation out of the system. It requires a number of periods where output is below equilibrium (and unemployment is above equilibrium) to get the economy back to the initial equilibrium (i.e. the central bank's bliss point). This case is shown in Fig. 3.9 in Chapter 3.

$\chi = 1$: *firmly anchored at* π^T

After the inflation shock, there is just one period of high inflation. In the second period, the Phillips curve reverts to its original position (i.e. $PC(\pi_0^E = \pi^T)$, as inflation expectations are firmly anchored to the target. This means that the central bank does not have to change interest rates and hence output does not deviate from equilibrium—the disinflation is costless.

$\chi = 0.5$: *partially anchored at* π^T

After the inflation shock, the central bank has to raise interest rates, but to a lesser extent than when expectations are fully backward looking. The fall in output is lower, and adjustment back to the initial equilibrium is quicker than when $\chi = 0$. If $\chi = 0.5$, then in the period following the shock, the Phillips curve moves down to intersect the y_e line at a level of inflation which is an average of inflation last period (i.e. π_0) and target inflation (i.e. π^T). In Fig. 4.6b this is shown as the inflation level π', and on this new Phillips curve the central bank chooses a level of inflation π_1 at point C. In contrast, point C′ in the lower panel of Fig. 4.6b marks the inflation–output combination associated with $\chi = 0$.

This modification of the 3-equation model highlights the role played by ability of the central bank to affect inflation expectations, which is referred to as its credibility in inflation targeting. More firmly anchored inflation expectations will reduce the negative impact of stabilization policy on the economy by restoring output and inflation to target more rapidly. The more credibility a central bank has, the lower the costs associated with disinflationary policies, thus central banks place a lot of emphasis on transparency and communication.

Two interrelated elements of credibility are important:

1. Will the central bank stick to its policy objective?
2. Can the central bank help shape expectations of inflation?

The communications strategies of central banks appear to address both. In relation to the first, they emphasize the independence of monetary policy decisions from political pressure and seek to explain that the bank's work will not be affected, for example, by the pressures on the government from an upcoming election. In terms of the second element, the central bank knows that if by helping the public to believe in the inflation target, it can keep the Phillips curves close to the inflation target, it will incur lower losses.

Independence is not the only driver of central bank credibility, transparency also plays a role. A transparent central bank is one that lets the public and financial markets into their decision-making process. The more the central bank communicates what they are doing and why, and their view of the future path of the economy and interest rates, the more easily the general public and the private sector can form inflation expectations. These concerns are illustrated by the Monetary Policy Framework of the Bank of Canada:

> The inflation-control target has helped to make the Bank's monetary policy actions more readily understandable to financial markets and the public. The target also provides a clear measure of the effectiveness of monetary policy. One of the most important benefits of a clear inflation target is its role in anchoring expectations of future inflation. This, in turn, leads to the kind of economic decision making—by individuals, businesses, and governments—that brings about non-inflationary growth in the economy.

The understanding of how monetary policy should be operated evolved rapidly in the wake of the failure of attempts to use money supply targets to bring inflation down in the late 1970s and early 1980s. By the end of the 1990s, central banks across the developed world had realized the importance of independence, transparency and credibility for successfully maintaining low and stable inflation. Figure 4.7 shows the relative levels of transparency and independence of global central banks in the 2000s.

Evidence

The academic literature on central bank communication has two distinct strands. The first strand uses constructed indices of transparency to test the empirical relationships between transparency and macroeconomic variables, such as inflation and output. The second strand

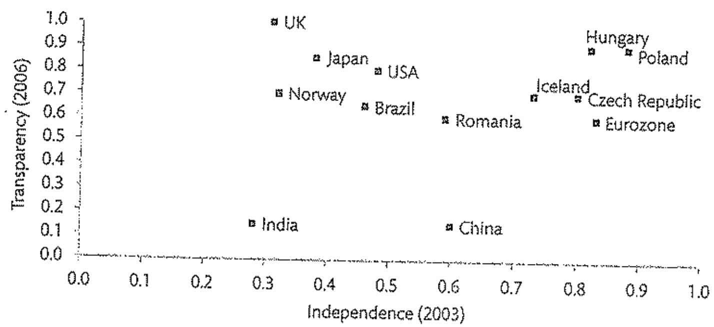


Figure 4.7 Central bank independence (in 2003) and transparency (in 2006).

Source: Crowe and Meade (2008).

focuses more explicitly on communication, trying to assess whether communication leads to better predictions of central bank behaviour and therefore has a desirable effect on economic stability.

Demertzis and Hughes Hallett (2007) and Dincer and Eichengreen (2009) are two papers from the first strand. They both find that central bank transparency reduces inflation variability, but has no statistically significant effect on the level of inflation.

Sturm and De Haan (2011) is a paper in the second strand, which identifies five indicators of European Central Bank communication that are based on the ECB President's introductory statement at the press conference following the ECB policy meetings. The authors find that although the indicators are often quite different from one another, they add information that helps predict the next policy decision of the ECB. This finding is backed up by Blinder et al. (2008), which provides a review of the available literature on central bank communication. The paper comes to the conclusion that more and better central bank communication improves the predictability of monetary policy decisions.

4.5 The Lucas critique

Macroeconomic policy throughout the post-war period has been heavily influenced by innovations in academic macroeconomics. Equally, macroeconomic theory and modelling has evolved partly in response to developments in the real economy. For example, the Keynesian doctrine was strongly shaped by the policy mistakes of the Great Depression and came to dominate policy making in the three decades after WWII.

In a similar fashion, the stagflation of the 1970s coincided with the development of *New Classical Macroeconomics*. This school of economics has not only been very influential in policy-making circles since the 1980s, but has also fundamentally changed the course of macroeconomic modelling. The underlying principle of New Classical Economics is that macroeconomic models should be based on rigorous microeconomic foundations, and in particular, on forward-looking optimizing agents with model-consistent expectations. The real business cycle model was developed in this tradition. It is set out in Chapter 16.

Such models were thought necessary to address the so-called *Lucas critique*, which highlighted the dangers of forecasting by using models that rely on the relationships found in historical data. Such forecasting methods will be unreliable if those relationships are conditional on the policy regime that was in place during the historical period being analysed. If the policy regime is different in the future, then the relationships found in historical data could well break down. Why? Because economic agents will change their behaviour in response to the new policy regime.[14]

The Lucas critique was directed at the use of the large econometric models that had been developed for forecasting. Such models dominated governments and central banks from the 1950s to the mid-1970s. A well-known example was the MPS model designed by Modigliani and his collaborators in the late 1960s. This model was used for quantitative forecasting and macroeconomic analysis by the Federal Reserve Board.[15] Lucas was very sceptical that these models, which were based on relationships from historical data, could provide any meaningful forecasts. He made a powerful critique of the policy mistakes that can be arise if wage and price setters use rational expectations, and policy makers interpret data as though they are not using rational expectations.

To explain this, we use the 3-equation model with rational expectations. To begin, we take an example where the government uses fiscal policy (tax or spending giveaways) to get a pre-election boost in its popularity. It may be tempted to allow a small increase in inflation before the election in order to let output and employment increase.

Suppose the government sneaks in such a spending increase after the central bank has set the interest rate for the period. Hence from the central bank's point of view there has been an unexpected permanent increase in aggregate demand, engineered by fiscal policy so that $y_t > y_e$. However, the central bank cannot do anything about it until next period.

What now happens to inflation during the current period? Assume that wage and price setters set prices with the same information as the central bank, that is before the government had increased expenditure. Since wage and price setters have rational expectations, they continue to assume that $\pi_t^E = \pi^T$ and they believe that $y_t = y_e$, so inflation remains at target. The next period the central bank reacts to the higher level of government spending and sets r so $y_{t+1} = A - ar_{t+1} = y_e$. Inflation is therefore at π^T in the next period and output falls back to equilibrium.

Hence if wage and price setters have rational expectations, the best that the government can do is to get a one-period gain in output. Nevertheless *in the data*, one would observe a one period rise in output without a rise in inflation above target.

The next step is to introduce a government that is suspicious of independent central banks. It believes, for example, that central banks are 'conservative' and are simply concerned to keep output and employment low. The central bank defends its behaviour and, using the Phillips curve with rational expectations, explains that increased government expenditure raises inflation if output is above equilibrium; and if wage and price setters have rational expectations and output is maintained above equilibrium, there will be great increases in inflation.

[14] See Lucas (1976). For a definition of the *Lucas critique* see Lars Ljungqvist's definition in The New Palgrave Dictionary of Economics: 2nd Edition (2008).

[15] See Brayton et al. (1997).

In spite of the advice of the central bank, the evidence the new government uses is that the old government increased government expenditure and output increased, without any change in inflation. Hence the new government thinks the central bank is simply conservative. It therefore takes away the independence of the central bank and says that the central bank must not raise interest rates when it increases government expenditure. The new government then increases government expenditure. Since the central bank is forbidden from reacting by raising interest rates, inflation will jump upward in the period after the shock and will remain elevated thereafter if the government continues to pursue the policy. If wage setters have rational expectations, they can only be 'fooled' by the government for one period, after which they will build the higher level output into their wage requests. Hence, if the higher output is not offset by the central bank in the period after the surprise shock, then the boost in government spending will result in higher inflation. As we saw above, with rational expectations, inflation would be expected to be infinite.

This example illustrates how the use of historical data showing the association of a fiscal expansion, higher output and no increase in inflation to forecast the outcome of a fiscal policy intervention can produce a poor economic outcome. In this example, the correct model of behaviour is one with rational expectations of wage and price setters: this can account both for the data observed from the pre-election episode and for the failure of the new government's policy of abolishing central bank independence and pursuing fiscal expansion.

Note also however, that a poor economic outcome can also result if a policy maker assumes that wage setters have rational expectations when they do not. Assuming the REH, a policy maker would expect that the announcement of a lower inflation target would produce a costless shift to lower inflation at equilibrium output. As Fig. 4.5 illustrated, if behaviour is better captured by adaptive expectations, then the anticipated costless disinflation would not take place. An inflation-targeting central bank would have to raise the interest rate and depress output below equilibrium in order to bring inflation down to the new target.

4.6 Expectations hypotheses, inflation bias and time inconsistency

In this section, we return to the 3-equation model where the policy maker engages in inflation targeting. The focus is on the problems that can arise when the *government* is able to exert control over monetary policy. What happens if the government has an inflation target but also an output target, which is a lower unemployment rate than is consistent with a zero output gap? The result of this policy stance is that the economy ends up at equilibrium output but with a higher rate of inflation. This outcome is called 'inflation bias'. Since inflation is above target and output is at equilibrium, the outcome is unambiguously worse than the case where the monetary authority targets a zero output gap. Inflation bias highlights the danger of government control of monetary policy and provides a justification for tying the government's hands through the delegation of monetary policy to an independent central bank.

The speed with which inflation reaches its higher equilibrium level after the new output target is announced depends on whether inflation expectations are adaptive or rational. We use the 3-equation model and begin with adaptive expectations where the economy adjusts over time to the new equilibrium.

We shall see in this section why inflation 'goes to infinity' under rational expectations in the case discussed under the Lucas critique and why there is a *finite* inflation bias in the case discussed here. Inflation only goes to infinity if the government or the central bank continue to pursue a level of output above equilibrium regardless of the consequences for inflation. The reason this does not happen in the inflation bias case analysed in this section is that although the government has an output target above y_e, it also has an inflation target. Hence there is a downward-sloping *MR* curve. We shall see that at the medium-run equilibrium with inflation bias, the economy is on the *MR* curve: the policy maker does not raise output because it would move them off their *MR* curve, which shows their optimal position given the Phillips curve they face.

4.6.1 Adaptive expections and inflation bias

A myopic government or one with an upcoming election may prefer to use monetary policy to target a level of output above equilibrium—i.e. $y_H > y_e$. We illustrate this case in Fig. 4.8, where the central bank's *MR* curve has shifted out as a result of their more expansionary policy stance. Now, the *MR* goes through the new output target, y_H, where the new loss circles are centred. The central bank's bliss point is now at *A'*, where output is at y_H and inflation is at target.

The adjustment of the economy after the change in policy stance is as follows:

1. The central bank forecasts that the Phillips curve will not move next period, because actual inflation has not changed as a result of the change in output target and agents form expectations in a backward-looking manner (e.g. $\pi_t = \pi_{t-1}$). The central bank therefore minimizes their loss function by lowering the interest rate, which stimulates aggregate demand and moves the economy to point *B*. Given the new *MR* curve, point *B* is the central bank's optimal point on the original Phillips curve. At point *B*, the government is closer to their new bliss point than they were at point *A*; they are willing to accept the rise in inflation, because output is closer to their new output target.
2. At point *B*, however, output is above equilibrium, so there is pressure on inflation to change. Inflation rises, the Phillips curve shifts upward and the central bank reoptimizes by raising the interest rate and moving the economy to point C.
3. At point C, output is still above equilibrium, so there is still pressure on inflation to change. Again, inflation rises, the Phillips curve shifts upward and the central bank reoptimizes by raising the interest rate. This adjustment process continues over a number of periods until the economy reaches the new medium-run equilibrium at point Z.

Figure 4.8 shows the inflation bias result. When the government tries to target a level of output above equilibrium the only result of the policy is that inflation is higher than its target. Output in the medium-run always has to be at the equilibrium level of output, as this is the only output level at which there is no pressure on inflation to change. On the new *MR* curve in Fig. 4.8, the medium run equilibrium has to be at point Z. At point Z, the government is further from its bliss point than if it had stuck with the original output target of y_e and

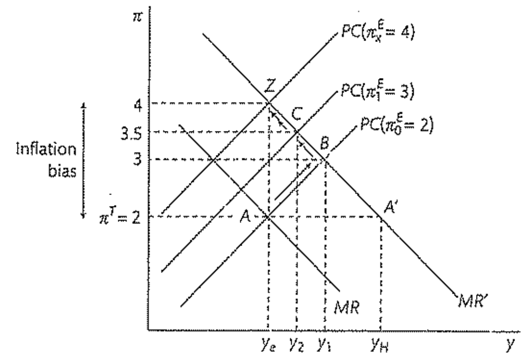


Figure 4.8 The inflation bias.

remained at point *A*. The example highlights the futility of governments targeting a level of output above equilibrium.

We can quantify the inflation bias by taking the difference between the equilibrium rate of inflation and target inflation: the inflation bias is equal to 4% − 2% = 2%.

The next step is to derive the same result mathematically and pin down the determinants of the size of the inflation bias. The loss function the central bank now wants to minimize is

$$L = (y_t - y_H)^2 + \beta(\pi_t - \pi^T)^2, \tag{4.6}$$

where $y^H > y_e$. This is subject as before to the Phillips curve,

$$\pi_t = \pi_{t-1} + \alpha(y_t - y_e). \tag{4.7}$$

Minimizing the central bank's loss function—Equation 4.6—subject to the Phillips curve—Equation 4.7—implies the first order condition:

$$y_t - y_H + \alpha\beta(\pi_{t-1} + \alpha(y_t - y_e) - \pi^T) = 0$$
$$y_t - y_H + \alpha\beta(\pi_t - \pi^T) = 0.$$

So the new monetary rule is:

$$y_t - y_H = -\alpha\beta(\pi_t - \pi^T).$$

This equation indeed goes through (π^T, y_H). Since equilibrium requires that there is no pressure on inflation to change, it must be the case that $y_t = y_e$ and that $\pi_t = \pi_{t-1}$. If we substitute this back into the monetary rule we have:

$$y_e = y_H - \alpha\beta(\pi_{t-1} - \pi^T)$$
$$\Rightarrow \pi_t = \pi_{t-1} = \pi^T + \underbrace{\frac{(y_H - y_e)}{\alpha\beta}}_{\text{inflation bias}}. \qquad \text{(inflation bias)}$$

In equilibrium, inflation will exceed the target by $\frac{(y_H - y_e)}{\alpha\beta}$. This is called the inflation bias and it will be positive whenever $y_H > y_e$.[16] The steeper is the central bank's monetary rule (i.e. the less inflation averse it is, lower β), the greater will be the inflation bias. A lower α also raises the inflation bias. A lower α implies that inflation is less responsive to changes in output. Therefore, any given reduction in inflation is more expensive in lost output; so, in cost-benefit terms for the central bank, it pays to allow a little more inflation and a little less output loss.

4.6.2 Rational expectations, inflation bias and time inconsistency

When wage and price setters have rational expectations, inflation bias also arises. Unlike adaptive expectations, however, the adjustment process is instantaneous. We can use Fig. 4.8 to illustrate the adjustment process under rational expectations. When the policy is announced, the economy jumps straight from point *A* to point *Z*. This is because the forward-looking, rational agents in the economy can see that given the new *MR* curve, the new medium-run equilibrium will be at point *Z*.

We can also derive the mathematical equation for the size of the inflation bias under rational expectations. Suppose that the goal of the government is to minimize the loss function:

$$L = (y_t - y_H)^2 + \beta(\pi_t - \pi^T)^2,$$

where $y_H > y_e$. All the actors can work out that if the central bank minimizes this loss function subject to the rational expectations Phillips curve, $\pi_t = \pi_t^E + \alpha(y_t - y_e)$, this leads to the new *MR* curve:

$$y_t - y_H = -\alpha\beta(\pi_t - \pi^T).$$

When the central bank chooses its optimal y_t, it has got to be on both the new *MR* and the *PC*. But if the wage and price setters use rational expectations so that $\pi_t = \pi_t^E$, then the central bank has to set $y_t = y_e$ because, as we have seen above, if it did not, inflation would soar (to infinity) and even a pro-employment policy maker would want inflation to be finite. Hence it has no alternative but to set $y_t = y_e$.

In that case the left hand side of the *MR* equation can be written

$$y_t - y_H = y_t - y_e + y_e - y_H = y_e - y_H = -(y_H - y_e) \text{ so that the } MR \text{ equation is now:}$$
$$y_H - y_e = \alpha\beta(\pi_t - \pi^T)$$
$$\Rightarrow \pi_t = \pi_t^E = \pi^T + \underbrace{\frac{(y_H - y_e)}{\alpha\beta}}_{\text{Inflation bias}}.$$

The determinants of inflation bias are the same as before.

The situation described is an odd one for a rational central bank or government to find themselves in. Behaving rationally, but with an output target of y_H, they end up with

16 For an early model of inflation bias with backward-looking inflation expectations, see Phelps (1967).

$$\pi_t = \pi^T + \frac{(y_H - y_e)}{\alpha\beta}$$
$$y_t = y_e$$

whilst if they targeted y_e, they would have ended up with

$$\pi_t = \pi^T$$
$$y_t = y_e.$$

Paradoxically, they would have been better off aiming for the *lower* output level y_e as the output target—a target they did not want.

Given the inflation bias result, can government that chooses the monetary policy stance commit to an output target of y_e? No, because of the problem of time inconsistency. A government that prefers $y_H > y_e$ always has an incentive to use monetary policy to boost output after rational wage and price setters have set prices. Using this approach, the government can achieve a one period boost in output without a rise in inflation (exactly as in the Lucas critique example), which moves them closer to their bliss point. In this scenario, rational wage and price setters can foresee that the government will adjust its policy stance after they have set prices, so will preempt the government and set prices consistent with the government pursuing an output target of y_H. The economy jumps to point Z in Fig. 4.8.

In other words, a government with a preference for output above equilibrium cannot credibly commit to targeting output of y_e. The government therefore ends up with inflation bias built into the system; output remains at equilibrium but inflation is persistently above target.

4.6.3 Approaches to mitigate inflation bias

Delegation

The problem of inflation bias can be solved by the monetary policy maker giving up their over-ambitious output target. This logic lies behind the idea of delegating monetary policy to an independent central bank, which has no incentive to target output above equilibrium. Put another way, the independent central bank can credibly commit to targeting output of y_e. Delegation is socially beneficial; welfare is higher if a body with different preferences from the government (and the private sector who elect them) has control over monetary policy.

In the standard central bank loss function, the policy maker cares only about minimizing the loss in the current period, i.e. there is extreme myopia. Whereas governments are likely to be characterized by myopia as a consequence of the electoral cycle, an independent monetary policy committee will take account of the consequences of inflation this period for next period's inflation (and their chances of meeting the inflation target). In order to encourage independent central bankers to take a medium-term view of policy, chairmen and committee members typically have long tenures that extend over multiple electoral cycles. The simplest way of representing this in the one-period loss function is to place a greater weight on the inflation deviation, i.e. to use a higher value of β. In other words, the longer term view of independent central bankers makes them more inflation averse than the government.

The empirical evidence supports the argument that delegation reduces inflation bias. Klomp and Haan (2010) use 59 previous studies that examine the relationship between inflation and central bank independence to carry out a meta-regression analysis. The authors find that the overall body of empirical evidence shows that central bank independence exerts a statistically significant negative effect on inflation (i.e. reduces it). They find a significant 'true effect' of central bank independence on inflation, even once they control for a significant publication bias, which makes it harder to publish results that do not show a statistically significant effect. The literature also suggests that adopting an inflation-targeting regime with an *explicit* inflation target improves macroeconomic performance in terms of both inflation and output stability by anchoring the public's inflation expectations to the central bank's objectives.[17]

In many OECD economies, inflation bias is no longer a significant problem, because control of monetary policy has been delegated to an independent central bank that is run by officials who are motivated by concern about their professional reputations. This point is summarized neatly by the economist Peter Howitt, who highlights the role of foreign exchange market operators in containing any tendency of domestic policy makers to try to boost output above equilibrium. We return to the role of the foreign exchange market in Chapter 9:[18]

> The 'temptation' to raise the level of economic activity with some surprise inflation might exist if society were indeed locked into expectations. In reality, however, the temptation just doesn't arise, as practitioners of central banking have long maintained. Central bankers are keenly aware that although there are long and variable lags between monetary stimulus and any resulting rise in the level of economic activity, there are no lags at all between such stimulus and the currency depreciation and capital flight that will occur if the stimulus is taken by investors as a signal of future weakness in the currency. Because of this, there is no reason for believing that discretionary central banks have the inflationary bias that the game-theoretic [time-inconsistency] view attributes to them. . . .
>
> Responsible people entrusted with such important and delicate jobs as the management of a country's central bank are typically motivated by the desire to be seen as having done a good job, to have acquitted themselves well. They pursue this objective by doing everything possible to avoid major inflations, financial panics and runs on the currency, while carrying out the day-to-day job of making available the base money needed for the financial system to function.

Reputation building

Another solution to the problem of inflation bias lies with the government or central bank building a reputation for being tough on inflation. Suppose that the government has delegated monetary policy to the central bank but wage setters remain unsure of just how independent the central bank is. They only know that there is a probability p that the central bank is independent and a probability $(1 - p)$ that it is a puppet of the government. The only

17 For example, Orphanides and Williams (2005).

18 Howitt (2001). Howitt refers to the useful paper by Mervyn King, then Deputy Governor of the Bank of England; from 2003, Governor of the Bank of England: King (1997). Another useful source is the short book of three lectures by Alan Blinder reflecting on how he used academic research when he was Vice-Chairman of the Federal Reserve Board: Blinder (1999).

way that they can find out is by observing the decisions taken by the central bank. If this is the case, how should the central bank behave? This problem can be analysed in detail using game theory.[19] Here we convey the flavour of the solution.

The situation is one in which the central bank interacts with wage setters more than once. Will a 'weak' central bank with an output target above the equilibrium find it rational to behave as if it were tough—i.e. with an output target closer to the equilibrium? If so, then we can say that it is possible to build a reputation for toughness as a method of solving the inflation bias problem.

Let us begin with the case in which the interaction between the central bank and wage setters occurs twice: in period one, wage setters choose π_1^E with no knowledge of whether the central bank is weak or tough (but they know there is a probability of p that it is tough); the central bank then chooses output in period one, y_1 knowing π_1^E. In period two, the wage setters choose π_2^E knowing y_1; the central bank then chooses y_2 knowing π_2^E.

The result is that a weak central bank will choose to act like a tough one in the first period, which will establish a low expected inflation rate in the second period, thereby providing bigger gains from boosting output in the second period. The central bank gains because in the first period, the outcome is inflation at its target (no inflation bias) and output at the equilibrium (instead of the time inconsistency outcome of inflation above the target and output at equilibrium) whilst in the second period, it can gain by setting output above the equilibrium (i.e. by exploiting the short-run trade-off between inflation and unemployment by a surprise increase in inflation).

When the game is extended from two to many periods, the benefits to the central bank from behaving as if it were tough increase. This is because the situation in period one (where inflation bias is avoided) is repeated again and again until the last period. This type of model provides an explanation for the process by which a reputation for toughness can be built in the face of public scepticism.

4.7 Conclusions

This chapter has taken a closer look at the role of risk, uncertainty and expectations in the macro-economy. We have shown that *risk* (known probabilities) and *uncertainty* (unknown probabilities and unknown events) are pervasive in the economy, which means that economic agents (e.g. households, firms, the government, etc.) must form *expectations* about the future if they are to make the best decisions possible.

For example, workers form expectations about inflation, which they use as part of the wage setting process to ensure their real wage expectations are met. We set out the two ways in which inflation expectations are commonly modelled: *adaptive* expectations, where expected inflation is equal to last period's inflation; and *rational* expectations, where agents form their forecasts using the model and taking into account all available information.

19 See Vickers (1986). For a simplified version of the central bank signalling game presented in Vickers' paper see Carlin and Soskice (2006), Chapter 16.

We also looked at the impact on the 3-equation model and its predictions when inflation expectations are modelled in different ways. This allows us to answer some interesting questions about the part played by expectations in the macro-economy:

1. When inflation expectations are formed rationally, how does this change the role of the central bank in the economy? The central bank's job is easier under rational expectations. They can directly influence the public's inflation expectations, so do not need to incur the employment cost of pushing down inflation before inflation expectations can be lowered. Disinflation is therefore costless. The introduction of rational expectations removes the lag between interest rate changes and output changes, meaning the central bank no longer has to gradually adjust the interest rate to guide the economy along the *MR* curve. After a rate change, the economy immediately jumps to the new equilibrium. Central banking is simpler if inflation expectations are formed rationally and consequently the economy is at equilibrium apart from random (i.e. unanticipated) shocks.

2. Why do modern central banks communicate to the public and financial markets about monetary policy? Central bank communication is directly related to transparency and credibility. In the absence of rational expectations, it is more likely that inflation expectations will stay close to target if the public are better informed about monetary policy. This was one of the key reasons for central banks introducing explicit inflation targets in the first place.

3. What is the Lucas critique? What dangers does it highlight for economic policy making? The Lucas critique concerns the problems of basing policy decisions on forecasting models that rely on relationships found in historical data. If the relationships in the models are conditional on the historical policy regime, then the models are rendered obsolete when the policy regime changes. There is a danger that constructing policy using these models will fail to produce the intended policy response due to the fact that economic agents change their behaviour when the new policy regime is introduced.

4. What happens when the monetary policy maker with an inflation target, targets output above equilibrium? How do expectations affect this process? When a monetary policy maker targets output above equilibrium, an inflation bias is built into the system; output is at equilibrium but inflation is persistently above target. The economy ends up at equilibrium output because there is always pressure on inflation to rise when output is above equilibrium. Whether expectations are adaptive or rational does not influence the eventual equilibrium, only the pace of adjustment. Under adaptive expectations the process of adjustment takes time, under rational expectations the economy jumps straight to the new medium-run equilibrium. Inflation bias is most likely to occur when the government exerts control over monetary policy. A government with an over-ambitious output target would actually be better off targeting y_e, but cannot credibly commit to that output target. The problem can be alleviated by delegating control of monetary policy to an independent central bank.

The modelling in this chapter has built on the core 3-equation model introduced in the last chapter. We have shown the impact on the macro-model of assuming the private sector as well as the central bank has rational expectations. It is very useful to understand the rational expectations hypothesis and its implication for the 3-equation model, because of the large impact the hypothesis has had on the development of academic macroeconomics and economic policy making.

Until the last chapter of the book, which introduces the real business cycle model and the New Keynesian DSGE model, we use as the base case a hybrid model where the central bank is forward-looking and has model-consistent expectations. The behaviour of wage setters is modelled by a Phillips curve with lagged inflation (where fully anchored expectations is a special case) and there is a lag in the *IS* curve. These choices reflect features of real-world economies (disinflations are costly in terms of unemployment and there are lags in monetary transmission) and provide a reasonable match to the way central banks describe their decision-making. Nevertheless, the Lucas critique needs to be kept in mind when designing and evaluating economic policy.

4.8 Questions

4.8.1 Checklist questions

1. What role do expectations play in the *IS* curve that underpins the 3-equation model? Provide an example of a situation where a change in expectations of the future can influence households' behaviour in the current period and shift the *IS* curve.
2. Describe the difference between risk and uncertainty. Provide an example of each case.
3. This question focuses on the discussion of Chevalier and Goolsbee (2009) in Section 4.2.2. Make sure you re-read this discussion before answering the following questions. Chevalier and Goolsbee (2009) suggests that the market for textbooks in the United States is rational and students are forward looking. If expectations were not rational and forward looking, how would you expect the example to differ? Would publishers be able to raise revenue by speeding up revision cycles?
4. Assess the following statements S1 and S2. Are they both true, both false or is only one true? Justify your answer.

   S1. Rational expectations means agents do not to make systematic errors.

   S2. Rational, forward-looking central banks' forecasts are often wrong.
5. When we add rational expectations to the 3-equation model, how does that change the predictions of the model?
6. Explain what is meant by the following statement: 'disinflation can be costless if the central bank is perfectly credible'. Draw the impulse response functions for output, inflation and the real interest rate following an inflation shock and interpret your results, when:
   (a) inflation expectations are fully backward looking
   (b) inflation expectations are firmly anchored to the inflation target.
7. Assess the following statements S1 and S2. Are they both true, both false or is only one true? Justify your answer.

   S1. Better communication by central banks can influence the path the economy takes after an economic shock.

   S2. Better communication by central banks does not affect how economic agents react to interest rate changes (which is the main tool used by monetary policy makers to achieve their inflation target).
8. Provide a definition of the Lucas critique. What is the relevance of the critique to the stagflation of the 1970s?
9. Imagine the government is able to exert control over monetary policy. What happens under adaptive expectations when the government targets a level of output above equilibrium? Is there a short-run trade-off between inflation and unemployment? How about under rational expectations?
10. Can a government with an overly ambitious output target credibly commit to targeting equilibrium output? If not, then how can they solve the inflation bias problem?

4.8.2 Problems and questions for discussion

1. Use this chapter and your own knowledge and further research to answer the following questions:
   (a) Is Keynes' treatment of expectations consistent with the rational expectations hypothesis?
   (b) What is meant by agents making 'systematic errors' under adaptive expectations?
2. Collect data on inflation and unemployment for the UK and the US during the 1970s and 1980s using OECD.Stat. The 1980s was a period of significant disinflation for the two economies. Use the data gathered and the content of this chapter to answer the following questions:
   (a) Was the disinflation costly (in terms of rising unemployment)?
   (b) Does the data provide evidence that inflation expectations were being formed rationally during this period?
   (c) How can the concept of anchored expectations be used to suggest what the UK and US governments could have done to reduce the costs associated with disinflation.
3. Assume that inflation expectations are formed rationally. Use the 3-equation model to show the adjustment of the economy to a permanent demand and a permanent supply shock. Provide a period by period explanation of the adjustment process (as done in Chapter 3). How does the central bank reaction differ from the cases where we assumed adaptive inflation expectations (as shown in the last chapter)?
4. This question uses the Macroeconomic Simulator available from the Carlin and Soskice website http://www.oup.com/uk/orc/carlin.soskice to show how central bank credibility affects the adjustment of the economy following an inflation shock. Start by opening the simulator and choosing the closed economy version. Then reset all shocks by clicking the appropriate button on the left hand side of the main page. Use the simulator and the content of this chapter to work through the following:
   (a) Apply a 2% positive inflation shock. Save your data.
   (b) Adjust the degree of inflation inertia/credibility of the central bank to zero (i.e. full credibility, which is equivalent to setting $\chi = 1$ in Equation 4.5). Save your data.
   (c) Use the impulse response function and Section 4.4.3 to compare the adjustment path of the economy following the shock in each case. Use the 3-equation model to explain any differences in the adjustment paths.