REFLECTIONS ON MONEY

ON MONEY
REFLECTIONS

Edited by
David T. Llewellyn

Foreword by
Lord Ezra

Published in association with the
Economic Research Council

REFLECTIONS ON MONEY

Reflections on Money

Edited by
David T. Llewellyn
Professor of Money and Banking
University Banking Centre
Loughborough

Foreword by Lord Ezra
President
Economic Research Council

MACMILLAN in association with the ECONOMIC RESEARCH COUNCIL

First published 1989

Published by
THE MACMILLAN PRESS LTD
Houndmills, Basingstoke, Hampshire RG21 2XS
and London
Companies and representatives
throughout the world

Typeset by Wessex Typesetters
(Division of The Eastern Press Ltd)
Frome, Somerset

Printed and bound in Great Britain at
The Camelot Press Ltd, Southampton

British Library Cataloguing in Publication Data
Reflections on money.
1. Monetary policies
I. Llewellyn, David T. II. Economic Research Council
332.4'6
ISBN 0–333–48528–9 (hardcover)
ISBN 0–333–48529–7 (paperback)

Contents

Foreword by Lord Ezra vii

Notes on the Contributors ix

1 Introduction 1
David T. Llewellyn

2 The Development of Monetary Theory 25
C. A. E. Goodhart

3 In Defence of Monetarism 37
K. Alec Chrystal

4 Credit, Broad Money and the Economy 59
Tim Congdon

5 Monetarism and Stagflation 83
Brian Reading

6 Structural Change in the British Financial System 109
David T. Llewellyn

7 Money in International Trade 139
Alan W. Clements

8 Money, Measurement and Accounting 167
Michael Bromwich and Christopher Noke

Foreword

The continuing objective of the Economic Research Council, of which I have the honour to be President, is to promote education in the science of economics with particular reference to monetary practice. The council has been working to achieve this objective by various means since it was founded, as the Joint Council for Economic and Monetary Research, in 1943.

My predecessor as President, the late Lord Beeching, once wrote that economics as a subject has only a small core of established theory, but the effects of it spread outwards into many other fields of thought and activity. This book is about money. We have little choice but to think about money in our daily lives: and we transact business not only with money but increasingly with credit. So where, one may ask, does money stand in economic theory and practice? And has monetary theory a place in that small core of established theory?

In observing in *The Wealth of Nations* that a propensity to truck, barter, and exchange one thing for another is one of the basic ingredients of human nature, Adam Smith virtually defined economics. These activities are its essence. They have been practised from time immemorial and for a very long time without the use of money as we know it. Looking back from now, however, who can doubt that for many centuries the topic that has had more enduring significance than any other in political economy is money? Probably no other issue in economics and in the conduct of policy has been subject to more academic and policy dispute. Schools of thought range from those holding to Pigou's idea that money is but a veil with little independent influence in an economy, to those who believe that controlling the money supply is the single most important task of policy-makers.

This divergence of views is reflected in the essays brought together in this volume, which is based on four study lectures on Money arranged by the Economic Research Council in 1988, with three additional chapters.

The issues exposed and discussed here are of more than passing interest. Each of the eight distinguished contributors illuminates the multifaceted complexity of this subject with the light of his knowledge and experience. The book may not settle any of the major issues

about money. That would be too much to expect. But if it gives readers greater insight into how economic systems work, with particular reference to the role of money, the Economic Research Council will be continuing to achieve what it set out to do in 1943.

Economic Research Council LORD EZRA

Notes on the Contributors

Michael Bromwich is the Chartered Institute of Management Accountants Professor of Accounting and Financial Management at the London School of Economics. He qualified as an accountant with the Ford Motor Company before reading for the BSc(Econ) at LSE. After a short period as Lecturer in Accounting at LSE he became the first Professor of Accounting at the University of Wales Institute of Science and Technology in 1970 and in 1975 became Professor of Finance and Accounting at the University of Reading. He took up his present Chair in October 1985. In 1987–8 he was President of the Chartered Institute of Management Accountants.

K. Alec Chrystal is National Westminster Bank Professor at the City University Business School. Formerly he was Professor of Economics at the University of Sheffield and he has also held teaching posts at the Universities of Essex and Manchester, the University of California at Davis and at the Civil Service College in London. During the years 1974–6 he was an Economic Adviser at HM Treasury and in 1983–4 he was a visiting scholar at the Federal Reserve Bank of St Louis. Books published include *Controversies in Macroeconomics*, *Political Economics* (with James Alt) and *Exchange Rates and the Open Economy* (edited with Robert Sedgwick). In addition he has published several papers in learned journals in the fields of money, macroeconomics and international finance.

Alan W. Clements is Group Finance Director of ICI, territorial director for Africa, and is responsible for overseeing the activities of the Acquisitions Team of ICI. He was appointed to the Board of Imperial Chemical Industries (ICI) in 1979. Having joined ICI as a member of the Taxation and Rating Section at Head Office in 1956, he then served as Assistant Treasurer, Deputy Treasurer and Treasurer between 1966 and 1979. He is also a non-executive director of Cable and Wireless, Guinness Mahon Holdings, and Trafalgar House. He was a Lay Member of the International Stock Exchange Council from 1984 to 1988.

Tim Congdon is Economic Adviser to Gerrard & National. He has been an influential economic commentator in recent years and has

been described as 'the City's last monetarist'. Although his views are actually more complicated than this, his research and forecasting have been distinctive in their emphasis on the importance of credit and money to understanding the economy's behaviour. He contributes widely to the financial press, particularly *The Times* and *The Spectator*. His most recent work is a short book on *Monetarism Lost: and why it must be regained*, published by the Centre for Policy Studies. He was Chief Economist at L. Messel & Co., the stockbrokers, for over a decade until 1987, and was briefly Chief UK Economist for Shearson Lehman Hutton.

C. A. E. Goodhart is the Norman Sosnow Professor of Banking and Finance at the London School of Economics. Before joining LSE in 1985, he worked at the Bank of England for seventeen years as a monetary adviser, becoming a Chief Adviser in 1980. Earlier he had taught at Cambridge and LSE, and was an Economic Adviser in the Department of Economic Affairs (1965–6). He has written a number of books on monetary history; he has just revised his graduate monetary textbook, *Money, Information and Uncertainty*; and he has published a collection of papers on monetary policy, *Monetary Theory and Practice*.

David T. Llewellyn holds the Chair of Money and Banking at Loughborough University and is Chairman of the Loughborough University Banking Centre. He has previously held posts at Unilever, HM Treasury, Nottingham University and the International Monetary Fund. He has held consultancy posts at financial institutions and the OECD and World Bank. He is also Consultant Economist to money brokers Butler Harlow Ueda. He has written extensively in the areas of international finance, monetary policy and institutions and is currently researching in the area of competition, financial innovation and structural change in the financial system.

Christopher Noke is Lecturer in Accounting at the London School of Economics. He read Philosophy, Politics and Economics at Oxford before qualifying as a Chartered Accountant with Peat, Marwick Mitchell. He later took an MSc in Accounting and Finance at LSE where he has held his present position since 1976.

Brian Reading is an Economic Consultant and journalist and currently Associate Editor of the *Sunday Times*. He has lectured at Oxford

(1960–2), and held economic advisory posts at the Bank of England, NEDO and Department of Economic Affairs. He was Special Adviser to the Prime Minister, 1970–2. He was Economics Editor of *The Economist* from 1972 to 1977 and has been self-employed since then. He is a regular broadcaster and has published widely.

1 Introduction

David T. Llewellyn

Money, and its management, has been an enduring topic of theoretical and empirical analysis over the centuries. Old and basic controversies constantly re-emerge in increasingly sophisticated ways even though the fundamental issues remain constant. Over the past two decades in particular there has been a substantial evolution in monetary theory, in our thinking about the role of money in the economy, and also with respect to monetary policy regimes and techniques of monetary policy.

Since the late 1970s this has been conducted in the context of far-reaching changes in the structure of the financial system, and the pace and nature of financial innovation. Given that the money supply is traditionally measured as notes and coin plus some measure of the domestic currency deposit liabilities of banks, changes in the structure of the financial system and the pattern of financial intermediation that involve banks have implications for trends in the money supply. Similarly, if non-bank intermediaries offer deposits with similar characteristics to those offered by banks, and if the functional and operational distinctions between institutions are eroded, the question arises as to which precise set of assets should be counted as 'money' and, by extension, should be the focus of monetary policy operations. In particular, a central issue is whether financial intermediation conducted by banks should be a particular focus of monetary policy. Some argue that it should be because of the unique character of the liabilities of banks and that it is this that makes them 'money'. On the other hand, others take the view that the uniqueness of banks is greatly overstated most especially in the current environment, that policy should not focus exclusively upon banks as institutions or their liabilities, and that monetary policy operates through a wide range of channels, markets and institutions.

The contributions in this volume were commissioned by the Economic Research Council. Four of the chapters (by Goodhart, Congdon, Reading and Llewellyn) are based upon public lectures given at the London School of Economics in the spring of 1988. In

this introductory chapter, and in order to set the scene, an overview is offered of some of the issues that feature in debate about monetary policy, particularly in the context of the experience of money supply targets and the Medium Term Financial Strategy (MTFS) which became the framework for the conduct of British monetary policy after 1979. In their different ways the contributors discuss various aspects of 'money' and the conduct of monetary policy and while the approaches are different, and the conclusions reached vary, they each start from the premise that 'money' and monetary policy are important issues and have an important bearing on the management of the economy.

I SOME CENTRAL ISSUES

There are several basic issues that feature in theoretical and applied analysis of monetary policy:

(1) The role of monetary policy in the management of the economy and the weight to be given to it relative to other policies and most especially fiscal policy.

(2) Whether monetary policy should be conducted by reference to an *intermediate target* lying between the instruments of policy and the ultimate objectives, or whether instruments are to be adjusted directly by reference to observed conditions in the economy.

(3) If an intermediate target is adopted a choice has to be made as to which one to focus upon and most especially whether the reference should be to a monetary or credit aggregate, the level of interest rates, or the exchange rate. If a monetary aggregate is adopted there is a wide range of alternative concepts to target upon. Within this choice multiple targets might be adopted, in which case decisions have to be made about how they are to be set in a consistent manner and how divergent movements are to be interpreted and accommodated.

(4) What type of monetary policy strategy is to be adopted and especially whether policy is to be adjusted in the short run or whether some form of fixed rule is to be applied. Related to this is the time-period to be adopted and whether monetary policy is viewed as a means of short-term demand management or as setting a medium-term framework in which both other

policies are conducted and other decision-makers in the economy (e.g. the wage-bargaining partners) frame their own strategies.

(5) A choice also has to be made over the precise instruments of policy though this is constrained in part by the institutional environment.

(6) A central issue in the conduct of monetary policy is the role of interest rates and whether discretionary adjustments to interest rates (e.g. in the current context, the setting of the Bank of England's bill dealing rates) are regarded as an explicit instrument, or whether interest rates are viewed as an equilibrating mechanism to other policy measures as is implied by monetary base control.

(7) In the final analysis an important dimension is how conflicts or 'dilemma situations' are to be resolved. Monetary policy cannot target simultaneously on the money supply, the exchange rate or interest rates, though focusing upon one may have undesired implications for the other two. This might involve, for instance, the setting of conditional targets where the target is compromised if other variables move outside an established range.

Over the past two decades there have been significant developments in the theoretical analysis of each of these issues and how monetary policy regimes have addressed them.

II ROLE OF TARGETS

Given the way that monetary policy evolved over the 1970s and 1980s, and particularly with respect to the MTFS, the role of targets is briefly discussed. In the structure of a monetary policy regime a *target* stands between the instruments to be manipulated and the ultimate goal of policy. But there are problems with this as this approach to monetary policy is efficient only if three conditions are met: (i) there is a predictable relationship between the target and the goal; (ii) the target itself is a stable concept and does not change its meaning; and (iii) the available instruments can in fact influence the target at an acceptable cost. If the role of a target is also designed to influence market expectations, a fourth condition is that it is credible in the sense both of the government's commitment to it and how it might realistically influence behaviour.

The general case for establishing targets is threefold: (i) they may

influence behaviour (e.g. the establishment of a monetary or exchange rate target might induce more moderate wage claims or increase employers' resistance); (ii) they can give an early indication of the movement of the goal variable (e.g. the rate of inflation) and as such be a guide to policy adjustments; and (iii) they may be regarded as a 'discipline' on all agents in the economy including the government and monetary authorities. In this way the setting of targets may give markets and decision-makers a statement of intent about the conduct of policy, and to that extent it may create an element of certainty to the policy environment.

But there are also problems with the establishment of targets. First, to the extent that the two key relationships (between instrument and target and between target and objective) are variable and unpredictable, serious policy errors can arise. Secondly, the efficiency of particular targets in yielding information about conditions in the economy can frequently be distorted; this has been particularly significant with respect to the efficiency of money supply aggregates as indicators of inflationary pressure. Thirdly, there is a practical danger that, given the commitment made to it, the target in effect becomes the objective almost irrespective of its effect upon the ultimate goal variable. It is also apparent that for the same reason, policy can be adjusted so as to produce cosmetic effects upon the target without influencing the ultimate objective.

III EVOLUTION OF MONETARY POLICY

Over the past few decades, the conduct of monetary policy has been subject to major changes with respect to all the basic issues noted earlier. These changes have been a reflection of different political ideologies (with respect to both the role and techniques of monetary policy), changes in policy priorities with respect to objectives, the experience of alternative policy regimes, and changes in the institutional environment in which monetary policy operations are conducted.

During the 1960s the assignment strategy implied a secondary role for monetary policy. In the dominant Keynesian ethos, fiscal policy was assigned to short-term demand management, 'Incomes Policy' was the chosen route to deal with inflation, and interest rates were adjusted in order to ease pressure on the exchange rate in the fixed-rate regime of the time. No reference was made to the 'money supply'

even to the extent of giving a low priority to defining and collecting data for money supply concepts; £M3 was first published in 1969! The view of the Radcliffe Committee (1969), for instance, was that monetary policy has only a modest role and that monetary policy operated predominantly through interest rates (by affecting the behaviour of lenders), albeit they were not very powerful. The Radcliffe view was that what was of central importance was the overall liquidity of the economy which itself was a difficult concept to measure. Certainly, no particular significance was given to the money supply. Throughout the 1950s and 1960s the emphasis was on credit control with direct restraints (e.g. lending ceilings) placed on banks to moderate their lending. However, this was because of the primacy given to credit rather than the money supply.

Although it was not viewed in this light at the time, the Bretton Woods regime acted as an anchor with the fixed exchange rate effectively acting as an intermediate target. Chrystal notes in his chapter that, although monetary policy was neglected in the post-war Keynesian strategy of economic management, the fixed exchange rate regime in practice imposed a powerful monetary discipline. He argues that only when money supply growth accelerated sharply after the advent of floating exchange rates, was attention focused on control of the money supply as a deliberate act of policy.

Gradually during the 1970s policy began to give more emphasis to monetary policy because inflation became a more serious problem, because increasing disillusion developed about the power of fiscal policy to fine-tune the economy, but also because the previous neglect of the money supply as such was seen as being inappropriate. By the mid-1970s policy targets for the growth of the money supply had been established by most industrial countries.

In the UK this culminated with the MTFS in 1979 which was a very precise formulation of a monetary policy strategy which gave primacy to monetary policy and within that to control of the money supply. However, in practice, by 1982 policy had become considerably more pragmatic than was originally implied by the MTFS. By the end of 1986 considerable scepticism had developed over conducting policy on the basis of explicit money supply targets, and formal targets were discontinued in the 1987 budget as policy came to be focused increasingly on the exchange rate. During 1988 it appeared as if there was an explicit sterling–Deutschmark target range.

In effect, by the late 1980s policy had turned full circle. At the end of the 1970s control of the money supply had become the central

feature of economic policy within a firm rules-oriented strategy; in theory the exchange rate and the level of interest rates were to adjust to the money supply targets. A decade later money supply targets had been all but abandoned and policy had become more pragmatic, flexible and with no exclusive emphasis being given to any particular variable though the exchange rate could be regarded as prima inter pares! Indeed, by 1989 as much attention was being given to the ultimate goal of policy as to any intermediate targets.

IV THE MEDIUM-TERM FINANCIAL STRATEGY

While a money supply target was first introduced in the UK in 1976, there was a symbolic turning point in the conduct of monetary policy with the establishment of the MTFS in 1980. While subject to several adjustments, this remained the focal point of both fiscal and monetary policy for several years. Its original version had two elements: (i) a declining set of target ranges for a particular definition of 'money' (£M3) over a four-year period beginning with 7–11 per cent for 1980–1 and ending with 4–8 per cent in 1983–4; (ii) the PSBR would be steadily reduced as a proportion of GNP from 3.75 to 1.5 per cent. This overall strategy was designed to produce an element of consistency between the government's monetary and fiscal policy. The overall objective was to reduce inflation through monetary control but without the necessity of 'high' interest rates. The normal policy dilemma (that of having to choose between money supply and interest rate targets) was to be avoided through the public sector reducing its demand for credit.

In terms of the issues outlined at the outset the major elements of this strategy were: publicly announced target ranges for a specific monetary aggregate, £M3; a decisive formal move towards the rules end of the rules–discretion spectrum; a medium-term time horizon, and a formal attempt to make fiscal and monetary policy strategy consistent. The MTFS (as originally intended) represented a strategic change in all the major dimensions in the conduct of monetary policies: (i) monetary policy became more central than fiscal policy; (ii) the establishment of an intermediate target; (iii) the focus of monetary policy became control of the money supply rather than interest rates; (iv) a medium-term time horizon was established; (v) the role of fiscal policy changed substantially and away from its hitherto demand-management, counter-cyclical role towards simply being an adjunct

to monetary policy, and in particular as a means of securing monetary control at low interest rates; and (vi) the central objective of policy became the moderation of inflation.

Early in the period of the MTFS there was considerable public debate, initiated in part by the authorities, about the techniques of monetary control and especially about the wisdom and practicality of adopting a form of monetary base control, the proponents of which argued would be the only certain way of securing the precise degree of monetary control envisaged in the MTFS. The arguments were outlined in Treasury (1980). While the Bank of England announced a series of changes in the technical conduct of monetary policy operations (Llewellyn, 1989) MBC was not adopted.

The rationale of the MTFS was based ultimately upon a monetarist view of inflation and that monetary control was not only a necessary element of an anti-inflation strategy but a sufficient ingredient. In particular, not only was 'Incomes Policy' a redundant policy for moderating inflation, to the extent that it distorts the workings of the labour market, it was viewed as positively harmful to economic efficiency and growth. Secondly, there was a belief that the announcement of a medium-term strategy would have a positive effect upon wage bargainers' expectations which could lead to a moderation of wage claims and inflation without the requirement of even a temporary rise in unemployment. Thirdly, the MTFS was a reaction against the concept of an 'elastic money supply' that, the government maintained, had been a feature of the 1960s and early 1970s and which had built into the economy an inherent bias towards inflation. In brief, the rationale was that if wage settlements exceeded the rate of growth of productivity either prices would rise (if there was an accommodating increase in the money supply) or unemployment would rise if there was no monetary adjustment. In this view, trade unions 'cause' inflation only to the extent that they successfully induce government passively to adjust the money supply to offset what would otherwise be the unemployment consequences of 'excessive' wage claims. It was contended that, because of an unquestioned commitment to full employment, successive governments had conducted monetary policy on this basis and that wage bargainers had thereby been protected from the unemployment effects of their wage bargaining. In the process there was an inflation bias built into the system. The MTFS, by announcing publicly in advance what the government's monetary policy would be, was signalling that in future the money supply would not be passively adjusted. The signal was that wage bargaining was

to be conducted in the knowledge that any unemployment effects of wage bargains would not be offset and this should be taken into account by wage bargainers.

On the face of it this appeared to be a decisive shift towards a monetarist strategy though in practice the monetary targets were not met, and in Chapter 3 Chrystal contends that this was not a test of the monetarists' strategy not least because it was not applied to its full, originally intended rigour.

In the event, there was a sharp real appreciation of the exchange rate in 1981 and 1982, output (and manufacturing output in particular) declined markedly, unemployment rose sharply, though inflation decelerated substantially. Thus the strategy did not work in the way envisaged by some economists (the rational expectations school) who believed that the announcement of credible money supply targets over a medium-term time horizon would induce a change in wage bargaining that could induce a decline in the rate of inflation without even a short-run rise in unemployment. Against this it might be argued that the episode was not a true test of the strategy in that the subsequent experience of the monetary targets being substantially exceeded demonstrated that the necessary credibility of the strategy (an important ingredient) was not sustained. The debate is summarised in Minford (1979) and Llewellyn (1980).

V A CONUNDRUM

When considering the results of the MTFS there is a basic conundrum: the rate of inflation was lowered during the period of successive versions of the MTFS (from around 20 per cent in 1979 to 3 per cent) and yet the money supply targets were constantly and massively breached. For instance, taking the mid-points of the target ranges for £M3 growth, it was originally envisaged in the first MTFS statement in March 1980 that in the subsequent four years the growth of £M3 would be kept to 44 per cent. The out-turn was 78 per cent. Between March 1980 and September 1986 the growth of £M3 was 153 per cent. And yet the rate of inflation was substantially reduced.

In the face of this conundrum two alternative explanations are possible. Either the theory (linking money supply and inflation) is wrong or the statistics on the money supply are inaccurate and £M3 is an inefficient and inaccurate measure of the true stance of monetary policy. Either way statistically it implies that the velocity

of £M3 fell sharply. This was one of the major issues addressed by the Governor of the Bank of England in his forthright 1986 Loughborough Lecture (Leigh-Pemberton, 1986). It is worth quoting at length for it focuses the issue clearly:

> Despite the progress that we have made towards our objectives, it cannot be said that our experience with our chosen framework for operating monetary policy has been satisfactory. In common with other countries, that framework has been one of targeting the rate of growth of a monetary aggregate . . . targeted aggregates have been periodically redefined, and target ranges revised upwards or even suspended for a period. Only two of the past six annual target rates of growth for £M3 have been achieved and, of those two, that for 1982–83 was achieved only after the target range indicated in the previous MTFS had been raised in the 1982 Budget . . . The fact is that we have been a good deal more successful in achieving our ultimate objective over the past five or six years than we have at hitting our intermediate broad monetary targets.

The main thrust of the Governor's lecture was about the second of the two explanations we have offered. It was in this context that the Governor stated in 1986 that the time might have come 'at which we would do better to dispense with monetary targeting altogether'. The Governor was emphatic that this was not because of any change in the objectives of monetary policy, but because specific monetary aggregates were difficult to interpret in a regime where financial innovation and structural change in the financial system meant that particular monetary aggregates were giving misleading signals.

VI WHICH M?

This is perhaps the point to consider the issue of which of the many alternative definitions of money (at one time eight definitions were published in official data) should be the focus of attention in the conduct of monetary policy. In the original MTFS a single concept (£M3) was adopted and in Chapter 4 Congdon gives support to it. In Chapter 2 Goodhart discusses the concept of 'money' and Chrystal presents econometric evidence for more than one concept in Chapter 3. In later versions of the MTFS multiple targets were set (£M3 and

PSL2) and in the final version the focus was switched to the narrowest concept, M0.

Llewellyn (1985) outlines the general problems of defining any unambiguous concept of 'money' from within the broad spectrum of financial assets. All particular definitions are arbitrary and include some assets which in all probability should not be included while excluding some that should be included. In addition, the relevant concept changes with financial innovation (e.g. as when building societies or money market funds issue cheque books) and may be different at different levels of interest rates. Financial innovation means that the distinction between balances which are included and those which are excluded from 'money' becomes increasingly difficult to sustain. This is most especially the case when there is substantial movement between them.

This general problem (which is not exclusive to £M3) in turn produces a dilemma in the setting of monetary targets. The inclusion of any one aggregate in the target regime places an unwarranted reliance on the efficiency of a single measure which for various reasons can become distorted, not least (following Goodhart's Law) because it is the target! No single concept is robust enough to play the central and exclusive role demanded for money supply targets, but the setting of multiple targets is equally hazardous for it has a confusing impact on expectations. Also, if the different aggregates are growing at the same rate then no additional information is provided, but if they are growing at different rates the problem of interpretation arises as to which is giving the relevant signals. Nevertheless, as is argued by Chrystal, this is not in itself a reason for relegating monetary policy though, as argued by the Governor of the Bank of England, it might question the efficacy of specific and precise monetary targets.

There are several money supply concepts. M2 is a recent innovation and was designed to be a measure of retail transactions balances irrespective of the institutions at which the balances are held. It is a 'monetary' aggregate that is not restricted solely to deposits at banks. It has the advantage of being defined in terms of clear functional rather than institutional criteria. The uniqueness of banks in providing transactions balances is appropriately questioned in this aggregate, and bank deposits account for less than half the total. It attempts to exclude investment balances. On the other hand, it is clearly demand-determined and hence of limited value for control purposes. Nevertheless, M2 is probably the most accurate measure of transac-

tions balances and as such might serve as a useful forward indicator of spending. It was never one of the official target aggregates (partly because there is only a short time series of data), though the authorities observed its movement.

The longest standing of the target aggregates is £M3. It was included as a target aggregate since monetary targets were first introduced in 1976, and for many years it was the exclusive target aggregate. It is a broad concept and hence is more a measure of liquidity than transactions balances. But £M3 is an unfortunate compromise, being too wide for transactions balances but not sufficiently broad as a liquidity measure as it is restricted entirely to notes and coin and bank deposits. It is an arbitrary definition of liquidity which is difficult to defend conceptually. On the other hand, it has the advantage of being more stable than the narrower aggregates, as many of the portfolio adjustments between alternative monetary aggregates take place within £M3. But the main advantage, and the one that ensured its survival, is essentially statistical. It can easily be integrated with other statistical concepts and is closely related to the aggregate sterling balance sheet of the monetary sector (banks). A related advantage is that the assets counterpart to the bank component of £M3 can be easily identified which in turn means that pressure on £M3 can be analysed in terms of the demand for bank credit.

As a liquidity concept, £M3 is deficient. The two *Private Sector Liquidity* aggregates are more comprehensive, as they include £M3 (except for a minor adjustment to exclude bank deposits with a maturity in excess of two years) but also encompass a wider range of liquid assets held by the private sector. PSL2 in particular has value to the extent that the liquidity (as opposed to strictly money balances) of the private sector is important in the economy perhaps because it indicates ease of access to transactions balances or directly influences the level of expenditure. The holding of non-money liquid assets may enable agents to economise on the use of money balances and hence increase the velocity of monetary aggregates. PSL2 is more general than the monetary aggregates in that: (i) it is not restricted to deposits; and (ii) the deposits it does include are not restricted to banks. It challenges the uniqueness of banks in providing liquidity.

VII THE UNDERMINING OF THE MTFS

In practice the MTFS was not executed with the precision implicit in its formulation; the targets were consistently exceeded, the policy in practice became more flexible than was originally envisaged, and recourse was made to multiple targets. Successive 'overshoots' were consolidated into subsequent years' targets and the target ranges were raised from the 1980 version. In effect, the authorities *chose* not to control the money supply along the path of the successive targets partly because it became apparent that £M3 in particular was giving false signals about the stance of monetary policy, but also because of the evident interest rate and exchange rate implications of adhering to the precise targets.

Above all inflation was decelerating and the overall stance of policy (as measured by movements in the real economy and the exchange rate) was evidently more restrictive than suggested by the path of the monetary aggregates. In effect, the *objective* of policy was being secured even though the *intermediate target* was not.

The MTFS was very precise in the formulation of medium-term targets for the growth of the money supply, and great confidence was placed in a particular broad monetary aggregate (£M3). In practice, the conditions under which the setting of monetary targets would in itself work to produce lower inflation are strict:

(1) it is possible to identify, within the broad spectrum of financial assets, an unambiguous concept of 'money';
(2) this concept can be precisely measured;
(3) it is a stable concept;
(4) it has a predictable relationship with the rate of inflation;
(5) it can be controlled through the available instruments at an acceptable cost.

If any of these conditions are not met in practice the central strategy of the MTFS (at least in the way it was envisaged to work) collapses. In fact there is substantial doubt about all of them.

In general the experience of the MTFS suggests four basic conclusions:

(1) the strict conditions for the successful operation of the MTFS are not, in practice, available;
(2) structural change and financial innovation in the financial system

create serious problems of both interpretation of monetary data and the implementation of monetary policy;

(3) the chosen monetary aggregate within the MTFS (£M3) became a very uncertain focus;

(4) the available instruments of monetary control (most especially interest rates) in practice operate with considerable uncertainty. This means that it is difficult to secure monetary or credit control at an acceptable cost.

It is for these reasons that, in practice, the strategy was implemented more pragmatically than originally intended. Several of the contributions to this volume highlight the substantial movements in velocity though different conclusions with respect to interpretation are given. The Governor of the Bank of England noted that £M3 velocity had by 1986 declined at an average annual rate of 4 per cent. The crucial point is the implication that there was not, over the period of the MTFS, a predictable relationship between the chosen target and money national income in general or the rate of inflation in particular.

This does not mean that monetary policy is not effective. It does, however, bring into question the efficiency of focusing upon specific money supply targets and perhaps the role of intermediate targets in general. By 1988 reference was being made to the ultimate objective of policy rather than intermediate targets largely because this seemed to give less misleading indications about the appropriate stance of policy to adopt. Overall, pragmatism and flexibility in the conduct of monetary policy is an appropriate approach when the market and institutional environment is changing as substantially as it has been since 1980. Nevertheless, there is a credibility problem when the official *description* of policy (which is designed partly to influence market expectations) is consistently at variance with the *practice* of policy.

Thus by 1987 the conduct of monetary policy had become pragmatic. Public statements showed that several indicators were monitored (including developments in the real economy and the rate of inflation and the exchange rate) in a manner that could not be encapsulated in a precise formulation; indeed, the essence of pragmatism and the antithesis of a strategy based upon precise targets with respect to intermediate variables. This implied a decisive shift back along the rules–discretion spectrum. In particular, the focus shifted towards interest rates and the exchange rate. In 1987, it appeared that the sterling–Deutschmark rate was the central focus and the foreign

currency reserves doubled in the context of substantial upward pressure on sterling in the foreign exchange market. It also appeared for a while that an unofficial (or unpublished) trade-off was established between the level of interest rates and the exchange rate in order to manipulate aggregate demand in the economy. By 1988, and in the context of renewed inflationary pressure, it appeared that, albeit in the context of an unprecedentedly large current account deficit, the exchange rate was to be used as the intermediate target as the discipline against conceding 'inflationary' wage increases. Overall, this represented almost a return to the implicit discipline of the Bretton Woods fixed exchange rate regime. By 1988, although reference was still made to M0, the primacy of 'the money supply' *per se* had been replaced by reference to interest rates and the exchange rate. This was reinforced in 1989 with a sharp rise in interest rates designed in large part to prevent the exchange rate depreciating in the context of a massive current account deficit.

VIII STRUCTURAL CHANGE IN THE FINANCIAL SYSTEM

Monetary policy theorising is often conducted in an institutional vacuum, the most notable example being the dropping of dollar bills from a helicopter. The evidence, however, indicates that structural change in the financial system and financial innovation have implications both for how 'money' is to be defined, and the impact that monetary policy operations have. Structural change and financial innovation are relevant in several ways:

— If distinctions between financial institutions are blurred the question arises about the alleged uniqueness of banks.
— Similarly, it raises the question as to which institutions' liabilities are to be defined as 'money'.
— Financial innovation (new instruments, etc.) can effectively widen the concept of money.
— If financial innovation takes the form of offering interest on money balances this will have implications for the interest sensitivity of the demand for money.
— Deregulation in the financial system affects the behaviour of banks and other financial institutions.
— Changes in the competitive environment and in the efficiency of the financial system also affect the supply, demand and form of

financial intermediation and how institutions are likely to respond to monetary policy operations.

Several of the contributions to follow discuss how the velocity of money has changed over the past two decades and how this has been affected by financial innovation.

The 1980s was a decade of unprecedented structural change and financial innovation in the British financial system induced by the simultaneous pressures of competition, changes in regulation, and the impact of technology. The overriding influence was a greatly intensified competitive environment. The substantial changes in regulation, the structure of the financial system, financial innovation, and the behaviour of financial institutions had a more general impact of relevance to the conduct of monetary policy during the 1980s. In the period 1980–7, the savings ratio fell sharply, the volume of personal sector borrowing rose substantially despite the high level of real interest rates, and simultaneously the personal sector built up its holdings of liquid assets on a large scale. This simultaneous rise in personal sector assets and debt represents an increased demand for financial intermediation services not only to finance expenditure but also the acquisition of financial assets. This partly explains why the demand for credit by the personal sector seemed not to have been responsive to the rise in real interest rates that occurred after 1980. This may be a key to the explanation for the rise in debt in that it may partly reflect both increased efficiency of the financial system, and a desire on the part of the personal sector to change the structure of its total balance sheet position. They are linked with a general process of deregulation combined with a more intense competitive environment and changes in the behaviour pattern of key financial institutions including building societies.

A major element in the recent sharp pace of credit creation was a stock adjustment associated with deregulation in the financial system; the system was adjusting to a new competitive environment. There have also been more structural factors which relate to the efficiency of the financial system and institutions and changes in their behaviour, though this too was associated in part with changes in the regulatory regime. If financial institutions become more efficient, innovative and responsive to market demands, it is likely that both the supply and demand for financial intermediation will rise. This will affect the demand for both financial assets and credit. If, for instance, institutions become less passive and more asset driven, it is likely that

the volume of credit will expand through a supply-side stimulus. It is likely that in both the UK and the USA the trend towards financial deregulation and the more intensive competitive environment has had a significant impact on the total supply of credit.

To the extent that this competitive environment also made financial assets more attractive, and a narrower interest rate margin between deposit and loan rates, both the volume of personal sector debt and the acquisition of financial assets would be expected to rise. The narrower the differential, the lower is the cost of maintaining liquidity on the basis of borrowed funds and even the greater possibility of making arbitrage profits. This is especially the case to the extent that borrowed funds secure tax relief or tax-exempt assets are acquired.

Overall, the new competitive environment induced in part by imposed and internal deregulation was a significant factor inducing a strong expansion in credit and the rise in the personal sector's debt–income ratio to record levels. In practice, there is little that monetary policy can do to moderate the pace of bank lending induced in part by a phase of structural change in the financial system. Monetary policy could not be used to offset the impact on credit flows of the regulatory, structural and financial innovation changes in the financial system evident since the late 1970s. It is similar to the response made to Competition and Credit Control in the early 1970s; monetary policy was largely impotent and the pace of lending was eventually moderated via other means.

IX ASSESSMENT

A lot has been learned about the conduct and operation of monetary policy, and the constraints that are encountered, through the various phases in the conduct of UK monetary policy, especially since the mid-1970s. A brief summary inevitably cannot do justice to the complexity and nuances of the issues involved. Nevertheless, as some of these features are discussed in the contributions to follow, it might serve to set a perspective though it cannot be claimed that a personal interpretation reflects a summary of the conclusions of subsequent chapters.

(1) The formulation and execution of monetary policy is a complex process in all the dimensions outlined at the outset of this chapter; i.e. in terms of strategy, form and technique. It is

difficult to imagine that a simple, unique strategy will suffice and governments have moved away from such an approach.

(2) The conditions required for the successful implementation of a precise 'rules' strategy as envisaged in the MTFS are very demanding and not met in practice.

(3) Both (1) and (2) are in part due to the complexities and non-static nature of the financial system. The conduct and interpretation of monetary policy cannot be divorced from the business operations of financial institutions, structural change in the financial system, and financial innovation. Each of these has implications for the nature of money and movements in velocity. The competitive and efficiency conditions of the economy also have implications for credit trends and the form and structure of financial intermediation.

(4) There is no single monetary aggregate that is sufficiently robust to be the exclusive focus of precisely defined monetary targets. All particular definitions are arbitrary and include some assets which in all probability should not be included while excluding some that should be included. In addition, the relevant concept changes with financial innovation. Financial innovation means that the distinction between balances which are included and those which are excluded from 'money' becomes increasingly difficult to sustain. This is most especially the case when there is substantial movement between them. Goodhart's Law suggests that the efficiency of any aggregate is compromised when it is established as a target.

(5) In the absence of exchange control, and as the financial system becomes more complex, the range of feasible monetary policy instruments is reduced. In particular, the power of direct-control mechanisms, or any instrument directed at a narrow range of institutions, is substantially reduced.

(6) In practice this implies that the only effective instrument of monetary policy is the short-term interest rate. However, in the conditions in which it becomes the only available instrument its power may at times be limited. As the relevant consideration in the demand for money is the differential between the rate of interest on money and non-money assets, the demand for money has become less sensitive to the general level of interest rates as a wider range of money and near-money balances carry market-related interest rates. In fact a rise in short-term interest rates can, under some circumstances, be counter-productive to

the extent that it increases the attractiveness of money as an asset, albeit as an investment rather than transactions medium though data cannot identify the motive for holding particular assets. In addition, the demand for credit seems not to be very responsive to movements in short-term interest rates, although as the personal sector's debt–income ratio rose steadily during the 1980s the impact upon income available for expenditure would have increased. A rise in interest rates can also be counter-productive to the extent that at high interest rates companies are more inclined to borrow through the banking system rather than the capital market.

(7) Monetary policy operations become particularly difficult to formulate and interpret, and their effectiveness is likely to be compromised, if the financial system is also in a phase of structural change, most especially if this is due to deregulation. The resultant stock adjustment in financial behaviour is likely to overwhelm any impact that monetary policy operations might normally have.

(8) As the volume of international capital movements has increased, and exchange controls have been lifted, the international dimension to monetary policy has become more pervasive. The floating rate regime, given the way it has been operated, has not in practice allowed governments to follow monetary policy divorced from that in the rest of the world.

(9) The policy dilemmas inherent in open financial systems can be particularly acute when choices have to be made between monetary aggregates, interest rates and the exchange rate.

This suggests that monetary policy operations must be conducted on a pragmatic basis, especially during a period of financial innovation and structural change in the financial system.

The operation of monetary policy is complex and far more so than can be encapsulated in any particular aggregate. All financial aggregates are potentially significant and not just 'money', all financial institutions and markets are important and not just banks, all financial flows are significant and no unique attention should be paid to the PSBR, and no single interest rate has overriding significance. Monetary policy is important but works through complex channels and mechanisms. The size of all financial imbalances, the portfolio preferences of financial institutions and their customers, and questions of relative efficiency of alternative financial intermediation

mechanisms all have a bearing on the conduct, and effectiveness, of monetary policy.

X THE CONTRIBUTIONS

Charles Goodhart discusses three basic issues related to money: the definition of money (especially its role as a means of payment) and how this has evolved over time; the nature of the demand for money; and the transmission mechanism of monetary policy. Information is seen as the crucial issue: 'Money should essentially be perceived as an instrument that allows an increasingly widespread and anonymous economic society to deal with the inevitable resulting shortcomings in information and trust of each of the members on the others.' In passing, he contemplates the possibility of an economy without money which is a theoretical possibility resulting from developments in information technology and computer-based electronic funds transfer. As with several of the other contributors, he notes the importance of the stability and predictability of the demand for money and how, due to structural change and financial innovation in the financial system, this has shifted.

Goodhart's chapter is followed by two written from a monetarist standpoint though the approaches adopted by Tim Congdon and Alec Chrystal are different. Chrystal, in his 'In Defence of Monetarism', addresses the critiques levelled at the monetarist approach to theory and policy formulation and states at the outset that 'the most vehement critics of monetarism simply did not understand what it was they were criticising'. He argues that 'the core of monetarism is a modest and sensible set of propositions, almost all of which are strongly supported by a mass of historical evidence'. He judges that monetarism has won and that the principal debates are within the monetarist school. The chapter discusses the nature of monetarism and of the implied policy strategy. In a series of statistical tests over a long run of data Chrystal seeks to demonstrate that velocity is ultimately stable, money causes inflation, and the main effects of fiscal policy are on the composition rather than the level of output. He cautions against judging monetarism on the basis of the UK's experience in the 1980s as the strategy adopted was not a true monetarist strategy.

Two main themes are discussed by Tim Congdon: (i) the growth of money is driven by the growth of credit, and (ii) the volume of

money has a strong link with the volume of spending. In the long run the link is between the money supply and the aggregate price level which, on the basis that output is not affected by money, is via interest rates and direct spending on goods and services. By distinguishing between commodity and bank money Congdon emphasises the nature of modern money as a liability of financial institutions. Distinguishing between monetary equilibrium and disequilibrium (and the process of how equilibrium is restored), he notes that only in disequilibrium situations does money disturb behaviour. In emphasising the significance of broad money concepts he argues that narrow concepts are seldom, if ever, in disequilibrium, and hence only broad money aggregates can behave in ways which surprise people and therefore make them reassess their decisions about spending. He therefore rejects the significance of M0 (which at the time of writing seemed to be the government's preferred money supply concept) because it is totally endogenous. The essay analyses the process of monetary dynamics and how, through decisions to purchase financial or real assets and attempts to get rid of excess money balances, equilibrium is restored via interest rate and price adjustments. Congdon concludes that the Bank of England has enormous power over the economy which can be applied irresponsibly.

A different approach is adopted by Brian Reading in his chapter, 'Monetarism and Stagflation'. He analyses trends in the international economy over a long period and observes that neither the Keynesian nor monetarist approaches cover the possibility of 'stagflation': high inflation coupled with weak conditions in the real economy, which he judges to have been the recent condition of industrial countries. The origin lies in cost–push pressures: 'The root of post-war cost inflation in the major Western democracies was that collectively voters demanded that governments spent more than they were willing, as earners and tax-payers, to release from the real output they produced. Stagflation and slumpflation have been the mechanisms which have persuaded governments and voters to change this behaviour.' Reading argues that forces producing stagflation steadily declined during the 1980s, though demand–deflationary forces intensified. However, the potential recessionary consequences of this were offset by a massive increase in world credit which prevented what otherwise could have been a major recession. As for the future, Reading argues that the world economy is on a knife edge and much depends upon what happens to real interest rates.

A different perspective is offered in David Llewellyn's chapter which considers the nature and causes of financial innovation and the major structural changes in the British financial system, together with their relevance for credit trends and monetary policy. After decades when the structure was ordered and stable, the British financial system changed markedly after 1970 and especially during the 1980s. Llewellyn describes a close parallel between the process and causes of financial innovation and structural change in the financial system: both have reduced rigidities and demarcations between institutions and instruments, and both have been induced by the simultaneous pressures of competition, technology, and changes in regulation. The overwhelming influence has been the intensification of competition in the financial system. In the process the financial system has become less structured and more integrated as the historic divisions between subsectors have been eroded. These structural changes have implications for the interpretation of the substantial growth of credit during the 1980s and the simultaneous rise in personal sector debt and financial assets. The essay concludes that there is a strong presumption that financial innovation and structural change have increased structural and allocative efficiency of the system, even though they may complicate the operation of monetary policy.

Alan Clements is the Finance Director of ICI and an experienced corporate finance officer within a major multi-national company. In the chapter on 'Money in International Trade' he considers the role of monetary factors and the international financial regime in the financing of international trade and investment, and considers how a multinational company deals with the special problems involved in the international aspects of money. He observes at the outset that both international trade and especially monetary transactions have expanded at a considerably faster rate than output. It is an important aspect of the international monetary environment that international monetary flows far exceed the volume of trade-related transactions and there has been a spectacular increase in the former since the mid-1970s. This, he suggests, is due to deregulation, financial innovation and the developments in information and trading technology. Throughout the chapter Clements draws on the valuable and illuminating experience of ICI and how developments in the international monetary system have affected the company's operations, and how financial management techniques have responded to the substantial changes in the international financial environment.

In a wide-ranging historical survey Clements traces how the

international monetary regime has evolved since 1950. This period has witnessed substantial changes and shocks: a major change in the system with the demise of the Bretton Woods system of fixed exchange rates, deregulation, the development of new markets and financing instruments, oil price shocks, high and volatile inflation, the evolution of massive international financial imbalances, and volatile interest rates and exchange rates. All these have major implications for the operation and financing of multinational companies. The essay describes the nature of the problems encountered (especially exposure problems) and the techniques and management systems that have evolved to handle them. This has involved the development of what, in effect, is an 'in-house' banking operation.

In summarising how the system has evolved over the period Clements concludes:

> Instead of basic economic transactions being expressed in money, money [now] often determines economic trends and developments . . . . Whereas in the 1960s and 1970s financial conditions were dependent on economic developments, now they have largely become dependent variables which can often determine the current level, and the likely future growth rate, of real international activity in both trade and investment.

In the final chapter an entirely different approach is adopted. Michael Bromwich and Christopher Noke are accountants and they discuss the important aspect of how money and inflation are handled by the accountant. Indeed, as the authors note at the outset, money is one of the basic prerequisites for systematic accounting and hence there is a problem if the value of money is not constant. This has an important bearing on the two major aspects of accounting: 'accounting for stewardship' and the presentation of a true picture of a company and the role of accounting techniques in generating information for decision-making.

Bromwich and Noke consider in detail the measurement problems associated with conventional accrual accounting and the value and problems associated with two other approaches: the 'economic income approach' and the 'user-needs approach' which in turn is divided into the 'decision-maker' and the 'decision model' approaches. Each of these alternatives offers different perspectives of the relationship between money and accounting. The authors give special reference to inflation and whether money is the most appropriate unit of

measure. They point to the many difficulties and ambiguities in attempting to deal with inflation and price changes. They conclude with a discussion of 'inflation accounting' (a controversial issue inside and outside the accountancy profession) and argue that some form of inflation adjustment should be made.

References

Leigh-Pemberton, R. (1986) 'Financial Change and Broad Money', *Bank of England Quarterly Bulletin* (December).
Llewellyn, D. T. (1980) 'Can Monetary Targets Influence Wage Bargaining?', *The Banker* (January).
Llewellyn, D. T. (1985) 'The Difficult Concept of Money', *Economic Review* (January).
Llewellyn, D. T. (1989) 'Bank of England Operations and the Determination of Interest Rates', in S. Ghartak (ed.) *Recent Issues in Monetary Economics* (Brighton: Wheatsheaf).
Minford, A. P. L. (1979) 'A Return to Sound Money', *The Banker* (July).
Treasury, HM (1980) *Monetary Control*, Cmnd 7858 (London: HMSO).

2 The Development of Monetary Theory

C. A. E. Goodhart

THE DEFINITION OF MONEY

Two subjects regularly stay both in the forefront and in the deeper recesses of our mind. The oldest profession in the world involves the exchange of one for the other. This chapter is about the less exciting of these two topics, that is to say 'money'.

It is important to be clear, or at least as precise as possible, what is meant by 'money', since the word is typically used in a very loose fashion. For example, in phrases such as 'Mr Getty has a lot of money' or 'money is the root of all evil', money is used in a generic sense to represent total wealth, or at least total financial wealth. In the chapters in this volume, however, the word 'money' is used in its narrower, more precise form – as also does the Inland Revenue. If one should be sufficiently misguided to try to draw up one's own will, stating in it that all one's money is being left to the favourite nephew, then he would have a legal claim to the currency holdings, notes and coin, and to the bank deposits that had been left. But all the rest of the financial assets would be taken by the Inland Revenue, on the grounds that they had not been left to any specified inheritor.

Instead, the definition of money employed in this chapter, and which should generally be assumed the other chapters use also – unless otherwise stated – will be that money represents those assets that represent a means of payment. That is to say, those assets whose transfer *completes* an economic transaction, leaving neither the seller nor a third party, who may have given credit to the buyer, with a further claim on the buyer. It should be noted that this definition, a means of payment, is subtly different from, and narrower than, the definition of money as a medium of exchange. The difference is as follows: a medium of exchange comprises any financial instrument that allows an exchange to proceed. This may include credit instruments, such as trade credit and borrowing from the bank, but such

latter credit transactions still leave the creditor with an outstanding claim against the initial purchaser, so that the transaction is not finally completed.

Even though we may think at first glance that defining money as a means of payment provides a fairly tight, precise definition, the set of those assets that might play this role in the economy has shifted in the course of historical development and, partly in consequence, the appropriate definition of money has, at times, been the subject of heated debate. Thus, in the nineteenth century there was intense discussion whether bank deposits, or indeed even bank notes, should be counted as money in addition to gold and silver coin; on the one hand, one can argue that the completed exchange of a cheque payment, when it has been cleared, does finally settle a transaction. But there were those who argued that a bank deposit, or a bank note, should be regarded as no more than a credit *claim* upon true money, and that the latter should be restricted to gold and silver coin.

Nowadays, we accept that bank sight or, as they term them in the USA, demand deposits, chequable bank deposits, should be included in virtually any definition of money; indeed the narrower definition of money, M1, includes outstanding currency holdings and just such chequable deposits held by the private sector with the banking system. But how about other deposits held with the banks which may not be immediately chequable, but which may be transferable automatically into demand deposits when the latter have been run down, or which may be capable of being used either immediately or at very short notice for a range, perhaps a restricted range, of transactions? Next, with the passage of the Building Societies Act in 1986, which should allow the building societies to provide a wider range of transactions and payment services for their customers, should we now include building society deposits or some subset of such deposits in our definition of money? Thus, in practice, besides the narrow definition of money, there are a whole series of alternative wider definitions, including in this country £M3 and PSL2, which for those not used to the acronyms adopted in this area stands for Private Sector Liquidity, definition 2. That may lead some to ask what private sector liquidity definition 1 might have been. This was an aggregate specially constructed and designed to avoid the distortion to the monetary statistics caused by the effect of that direct control, colloquially termed the corset, which was in operation between 1973 and 1980. With the abolition of the corset, the PSL1 definition lost its *raison d'être* and has since quietly fallen into disuse.

Perhaps even more important, it is perfectly possible to see forthcoming technological developments which will alter further still the structure and nature of payment systems. The further development of the ubiquitous plastic card, which will shortly come with its own built-in computer intelligence, together with the increasing ability to link homes, offices and shops with a single central computer to allow and facilitate electronic funds transfers, may shortly change the whole face and form of the payments mechanisms. We have progressed from a system of commodity money, in which money was overwhelmingly represented by coins of precious metals, to a system in which money primarily took the form of specific claims on financial intermediaries, with these latter claims in turn convertible on a one-for-one basis into legal tender. In response to the pressures on government finances in the last century, notably during periods of major war, such legal tender has come to take the form of government fiat money, that is to say, claims on the government which are not backed by the right to convert that claim into some other good or commodity; and we may now be moving towards a system wherein our generalised wealth holdings, whether claims over assets, or our command over our own human capital in the form of the present value of our expected future earnings, may all be capable of being utilised, up to a point, for immediate payment purposes. This latter raises the question whether the ability to exercise command over a much wider range of assets than in the past, for the purpose of making payments, might lead the definition of money to become so wide and general in future that it does indeed almost coincide with total wealth, and therefore ceases to have the more precise, narrower definition that has been employed so far. We shall return to this question in a later section.

THE USE OF MONEY

First, we embark on another associated historical tack, to ask why people need to engage in monetary exchanges at all. Now this may seem an odd, indeed unnecessary, question. If, however, one thinks of certain closely knit, small communities, such as the family or certain small tribal groupings, then there are frequently no monetary exchanges within that small closely knit group at all. Yet there are exchanges of goods and services within such small communities that are certainly economic in form, for example, my wife cooks for me; I ask my son to wash up for her. What happens within such small

social groupings is that our knowledge of each other, and the social pressures and sanctions that we can bring to bear when any member of this small group fails to play his or her allotted role, are strong sufficiently strong so that the provision of a good or a service at any time from one member of this group to another need *not* be exchanged for an object of comparable value, since the giver can be reasonably sure that he or she will subsequently obtain sufficient offsetting benefits by their membership of the group, or be faced with the ultimate sanction of being kicked out of the group. When the scale of our economic society increases, so that we have less information on, and fewer social sanctions over, the behaviour of those with whom we are dealing, then we will be generally unwilling to undertake a transaction in which we give up a valuable good or service, unless we receive what we may perceive as an asset of equal value in return. Initially, this led to barter, whereby the transactors would exchange goods which each felt was equal or greater in value in receipt than the good he himself was supplying. But the restrictions imposed on the feasible set of transactions by the need to find a double coincidence of wants, in which, for example, a hairdresser with a desire for apples might need to meet up with an apple grower who wanted a hair cut at that moment, were so considerable that from a very early stage our societies developed commodities, which, in addition to their uses for other purposes, could also serve as money.

The history of these early commodity monies, including such curious features as red feathers, wampum, nails, cowrie shells, and cattle, from which the term pecuniary has been taken, and leading up to the generalised development of gold and silver coin as commodity monies, forms a lively and exotic part of mankind's history. However much fun that history is, and I certainly find it enjoyable, one still has to realise that the development of such commodity monies was simply one among a range of possible alternative ways of dealing with the limitations on information, and on trust, and the restricted sanctions that we can impose on other people's behaviour, in order to be sure that in future they will provide offsetting value in exchange for the goods and services that we have given to them initially. Money should essentially be perceived as an instrument that allows an increasingly widespread and anonymous economic society to deal with the inevitable resulting shortcomings in information and trust of each of the members on the others.

This latter thought again raises the issue whether information

technology in general, and computer based electronic funds transfer in particular, could not be based on a more centralised intelligence system, allowing payments to be made on the basis of much wider generalised claims over assets, including future earning power, than has been the case in the past. To put it another way, monetary economists are increasingly beginning to wonder whether there will be any necessary role for a narrow and precisely defined monetary asset to exist within the future economy at all. There has, indeed, been some interest in the somewhat abstract theoretical issue of whether there would be a determinate price level, for example, in a world without money.

In my view, these latter issues remain, for the present, only of academic theoretical interest. It is as yet premature to envisage the arrival of a system without money. One reason for this is that the extent of information about each member of that society, necessary to run such a money-less economy, would be much greater than most of us would be comfortable to provide and have hanging over our own head. Even now, a surprisingly large proportion of outstanding currency – in particular the large denomination bank notes outstanding – is probably held for purposes such as the black economy and activities which we would rather keep anonymous, including criminal activities, to a much larger extent than is commonly recognised. Payment for prostitutes or drugs is rarely by cheque or credit card. For example, the Federal Reserve Board in the USA held a survey to try to ascertain where outstanding dollar bills were being held, and when they grossed up the results of their survey sample they found that they could only account for a small fraction of outstanding notes. In the UK, the value of notes and coins in circulation with the public was over 14 thousand million pounds at the end of 1987. If that is divided by the 50 million or so inhabitants of this country, the result is that *on average* every man, woman and child in the country must be holding about £275 in their pockets. Even when full account is taken of the accumulation of currency in shop tills, and so on, the size of the outstanding volume of currency remains really quite remarkable. In particular, it suggests that the general public will be quite widely unwilling to move to a centralised information, credit based payment system, and to drop entirely the present, anonymous fiat currency system.

Furthermore, the authorities obtain a small, but still worthwhile financial advantage in the form of seignorage from being able to print their own money. Moreover, the receipts from money creation,

effectively an inflation tax, can be increased substantially, if needs be, at times of crisis. Together with the power to control domestic monetary policy, an important instrument, which the ability to vary the monetary base gives to the authorities, I cannot see them willingly giving up their monopolistic command over the currency base of the system. While there have been certain recent suggestions, for example, US Secretary of the Treasury Baker's kite-flying exercise in 1987, in which he raised the possibility that the authorities might revert from a fiat currency system to one based again on a currency convertible into a wider basket of commodities, I think that it is politically naive and unlikely to expect the authorities to renounce lightly or easily their monopoly control over the supply of currency. Thus, for the foreseeable future, it is likely that the monetary system will consist of three main components. The first will be the fiat currency base, monopolistically supplied by the authorities, the central bank within a particular area. On top of this will be the traditional banking system, providing deposits with a range of payment services and an assortment of increasingly versatile plastic credit and debit cards. But beyond this will be a widening penumbra of other financial intermediaries offering limited payment services, together with a variety of plastic cards and other financial functions.

III THE DEMAND FOR MONEY

This account of the nature and role of money and how the actual set of monetary assets has changed over time, and is still changing, should already have provided a fairly clear insight into why we need to hold money balances. Given the limitations on information and on trust, there are various situations in which one can only complete a transaction if one actually holds currency. There is the famous pub sign which says that 'the bank has agreed not to serve beer, and we have agreed not to accept cheques'. There is a much wider range of transactions in which the promise to pay *at some future date* against a potential ability to raise funds by the sale of assets, or from earning power, would not be accepted as good enough. The seller may require an immediate payment which may either be guaranteed by a bank, for example via a cheque card, or by some other financial intermediary. To use the phrase adopted by many economists, in order to undertake a wide range of transactions, there is what is known as 'a cash-in-advance constraint'. Having access to future income is not

good enough. We need monetary balances now in order to be able to consummate current expenditures. This same analysis allows us to explain quite simply the factors that determine the demand for such monetary balances. The higher the level of expected future transactions, and the greater the uncertainty about the amount of such transactions, the more monetary balances people will want to hold.

Next, if two assets were offered, each providing the same expected interest rate yield, but one of them provided more payment services than the other, one would hold all one's wealth in the asset with the same interest rate, but providing more payments services. On the other hand, a bank providing deposits *with* payment services is going to have to hold more low yielding reserve assets, in the form of currency in its tills and balances with the Bank of England, in order to be able to honour the convertibility promise which it has implicitly made, to transform deposits that you hold with it into legal tender cash at any time. Accordingly, from both the supply and demand side, the yield or rate of return available on assets with significant additional liquidity and payment services is likely to be lower than on those assets with fewer monetary characteristics. So, the demand for money assets is likely to be inversely related to the margin between their own rate of return and the rate of return available on non-monetary assets. Indeed, in previous decades, we used to think of money as generally having a zero interest rate. But with the advent of high interest yielding deposits becoming available at commercial banks in many countries, including the UK, this is no longer strictly the case. It is only the authorities' legal tender currency which continues to have a zero nominal interest rate, and there have been those of us who have wondered whether this is necessarily so, or whether there would be ways that could be found of providing interest on currency also.

IV THE TRANSMISSION MECHANISM OF MONETARY POLICY

Thus we have established a demand for money to hold, depending ultimately on the cash-in-advance constraint, in which such demand is a positive function of the volume and uncertainty of the pattern of future expected transactions, and is inversely related to the margin between the lower rate of interest available on monetary assets and

the higher rate of return available on less liquid, non-monetary assets. Much of the broad thrust of this latter analysis of the determinants of the demand for money would be common ground to monetary theorists throughout the ages. Where there have been, and continue to be, many more disputes between monetary theorists concerns the results that might be expected to follow when the money *supply* is changed, perhaps by greater government expenditure or by the banking system granting more credit, leading to an initial *imbalance*, or initial disequilibrium, between the demand and supply of money. We should distinguish in this matter between various major schools of thought. The first of these is normally termed the Classical School. In its analysis, an increase in the supply of money, leading people to have more money initially in their pockets or in their bank balances than they would want to have at existing price and interest rate levels, would lead to a *general* expansion of expenditures as people use their excess money balances to buy a wide range of goods and services. But the Classicists saw the level of real output, and the production of all goods, as being brought back swiftly to an equilibrium level through changes in relative market prices: such latter shifts in relative prices include, importantly, variations in interest rates, which were seen as equilibrating the level of savings and investment; savings were seen as the demand for goods in the future rather than today, while investment was seen as the means of producing goods in future rather than today. The rate of interest was seen as the intertemporal price of goods in future, as compared with goods at present. So, for example, when savings, the demand for goods in future, were fairly high relative to the ability to produce goods in future, the rate of interest was expected to fall to equilibrate savings and investment. Thus, with the mechanism of prices shifting *relative* to each other, maintaining full equilibrium within the economic system, an expansion of monetary balances would simply lead in fairly short order to a generalised expansion in the overall level of prices. In a Classical world, monetary expansion leads directly and swiftly to a concomitant change in the general price level.

This was widely seen to be an idealised picture of a world in which, during the course of the nineteenth century, there were quite marked and often, it appeared, relatively similar cycles around a normal equilibrium level of output. This latter was explained by a group of economists, comprising such famous names as Wicksell, Hawtrey and Dennis Robertson, as reflecting certain imperfections within the economic system, especially and notably within the banking system,

that prevented the economy adjusting as perfectly as indicated within the simplified Classical model. In particular, bankers, like everyone else, would find it hard to distinguish between those occasions when an expansion would quickly exhaust itself in an inflationary impulse, or might actually represent a real economic improvement, say as a result of technological innovations and supply side improvements to the productivity of capital and labour. Thus a purely monetary expansion, increasing bank reserves, would make bankers keen to expand lending, and the fall in interest rates would make businessmen willing to invest and expand more generally. This would cause a period of high profit and prosperity, and general optimism, leading to further balance sheet expansion. Assuming that the expansion was *not* caused by improved real conditions, say in the form of technological productive advances, then the expansion would ultimately lead to inflation, a drain of cash reserves, and a falling away of profit margins as wages and input prices rose relative to final prices. This might then lead to a financial crisis or panic, which would lead bankers in turn to become much more risk averse, conservative, and unwilling to increase the size of their balance sheet, leading then to a down turn. Even so, these cyclical fluctuations took place around a full employment equilibrium, to which these economists saw the economy reverting in due course, even if the various imperfections of information within the system prevented such adjustment being rapid.

The main challenge of the Keynesian revolution involved the attempt to claim that no such reversion to a full employment equilibrium position was necessarily likely. Keynes reached this conclusion by arguing that savings and investment would be brought into equilibrium by changes in the level of income, through the multiplier process, *rather* than by variations in the rate of interest. Thus at any time, the level of income to which the economy would settle in the Keynesian analysis depended on that level of income which equilibrated the demand for investment with the demand for savings. Since under Classical and Wicksellian analysis such equilibration had been achieved by variations in the rate of interest, Keynes then had to explain what determined the rate of interest if it was not the relative volume of demand for funds to borrow for investment, as compared with the volume of flow of new savings into financial markets. Keynes then argued that the rate of interest, rather than equilibrating the demand for savings with the demand for investment, was instead the relative price that equilibrated the

demand and supply of money. As we noted earlier, the demand for money was a function of foreseen expenditures on the one hand, and the relative margin between the rate of interest on monetary assets and on non-monetary assets.

The Classicists, and economists like Dennis Robertson thereafter, argued that an excess of money supply would lead money holders to go out and spend more generally on all assets, increasing the general level of expenditures and thereby equilibrating the demand and supply of money. Instead, Keynes argued that an increase in the money supply, relative to money demand, would not lead people generally to go out and spend on *all* assets, but would make them switch their excess money balances into a subset of substitute, alternative, near-money financial assets. Thus, within the Keynesian model, the transmission route of monetary impulses, say an expansion, ran entirely through interest rates. A monetary expansion would be expected to reduce interest rates on liquid financial instruments, leading on to a reduction in interest rates on longer term bonds as people switched out of lower yielding liquid assets into higher yielding longer bonds, which would then lead to an increase in equity prices, after which the reduction in yields on financial assets generally would lead to some expansion in interest sensitive expenditures, such as investment and, in particular, house building.

Thus a distinction between the original Classical and the Keynesian models related to the prospective behaviour of agents in the economy when faced with monetary balances which at a particular time were higher than they felt they needed. The Classicists saw agents using such excess money balances quite generally for the purchases of a wide range of goods and assets; the Keynesians saw agents with excess money balances placing these in alternative, substitute financial assets.

In some large part, the monetarist revival, led by Milton Friedman, represented no more than a reversion to the earlier Classical view rather than the much more limited Keynesian transmission route. In the main, this dispute between the various schools of thought about the transmission mechanism of monetary policy relates back to a more fundamental issue about the comparative efficiency with which markets work, and the speed with which they restore equilibrium within the economic system. Although Keynes himself always claimed the opposite, that his theory represented a 'General Theory' which was true under all circumstances, it has been increasingly accepted that Keynesian analysis depends heavily on the assumption that there

are major imperfections within the economy which prevent the system of price adjustment in free markets from re-establishing a full equilibrium very quickly or easily. Certainly, the more perfectly markets work and adjust, the more nearly the Classical, and indeed the recent neo-classical schools of Lucas, Sargent and Wallace appear to be justified, and by the same token the less valid would be the Keynesian vision. But if you take the analysis of perfectly clearing markets to the extreme, which would appear to involve and imply virtually perfect information within the economic system, then it is difficult to see what role there would be for money at all within such a hypothetical economy, since the latter depends crucially on limitations in information and on trust. Indeed, in the latest perfect clearing market, full information models of real business cycles being developed by the current wave of neo-classical economists in the United States, there are many who have appreciated that there is no fundamental role for money and monetary phenomena within such hypothetical economies at all. But if the information available to agents in the economy is limited and people do not trust each other, and markets work imperfectly to eliminate surpluses and deficiencies, then not only will monetary assets be demanded but also the effect of monetary expansion is more likely to show itself, initially at any rate, in prices in those markets which do adjust relatively rapidly – namely, in financial markets such as the market for government bonds and the foreign exchange market, and also through temporary adjustments in real output – rather than working through directly into *generalised* price and inflationary changes.

These latter arguments continue to rage. If markets adjust relatively perfectly, bringing the economy back rapidly to an equilibrium full-employment level, then monetary fluctuations can *only* cause fluctuations in the rate of inflation. Under those circumstances, it would indeed be sensible to commit the authorities to a fixed rule, or target, to try to bring about a steady rate of increase in the money stock, and by the same token, a steady and predictable, and hopefully low, rate of inflation. If, on the other hand, markets worked very imperfectly, so that the economy frequently remains for some long period in conditions of severe disequilibrium, e.g. with periods of excessively low unemployment and inflationary pressure, then it is arguable that the authorities might be able to use their command over monetary policy to offset many of the remaining extensive imperfections within the economy that have allowed those conditions to appear.

Both these latter arguments, for rules and targets on the one hand, and for activist, fine-tuning monetary policy on the other to offset imperfections within the economic system, depend, at least in part, for their effective conduct on a knowledge of what the underlying demand for monetary balances may be. Only if one knows what the demand for money is likely to be can one set the target rate of growth of the supply of money, or change monetary growth through discretionary policy measures, in order to have the desired effect on price inflation, and/or on the level of economic activity. If the demand for money should become unstable, then one does not know how to set the money supply in order to achieve desired objectives.

What has happened in recent years, however, is that the various structural changes that were referred to in the earlier part of this chapter, including the payment of interest on demand deposits in banks and the provision of payment services on a wider range of assets outside the traditional bank deposits, has led to shifts in the demand for money, or in other words to shifts in the velocity of money which cannot be predicted at all well in advance. This latter has made the attempt to regulate monetary policy by reference to the perceived rate of growth of certain monetary aggregates a much more uncertain exercise. Instead, the authorities are having to revert to adjusting their monetary instruments, notably through administered changes in interest rates, in response to changes in inflationary pressures themselves, in response to changes in the rate of growth of nominal incomes, and notably of wage rates, and more recently and of growing importance in this country over the last couple of years, to the time-path of sterling in comparison with the Deutschemark. Thus we are moving back in these last couple of years effectively to a convertible currency system, in which the pound sterling takes its value from maintaining its convertibility with an external source of value – the Deutschemark.

3 In Defence of Monetarism

K. Alec Chrystal*

In some quarters an author who attempted to defend monetarism would be viewed in the same light as someone who supported paedophilia or incest! Monetarism represents the forces of darkness, ignorance and destruction which, it is alleged, inflicted untold damage on the British economy (see, for example, Gould, Mills and Stewart, 1981; and Kaldor, 1982). How could any intelligent person defend the indefensible?

The answer is that the most vehement critics of monetarism simply did not understand what it was they were criticising. The core of monetarism is a modest and sensible set of propositions, almost all of which are strongly supported by a mass of historical evidence. Indeed, most well-informed individuals would now accept all the main propositions of monetarism even though they were anathema to the Keynesian orthodoxy a mere twenty years ago. To this extent, it is safe to say that monetarism won. Its message has been accepted, and for this reason alone it seems to have outgrown its usefulness. The world at which it was originally aimed has changed considerably and so the policy implications of monetarism must change too. However, the principal debates are within the monetarist camp. There is no question of going back to the naive Keynesianism of the 1950s and 1960s. This will be true irrespective of the political complexion of future governments. Indeed, many Marxists (including Marx himself) have long subscribed to several of the major monetarist propositions. Hence it is not safe to associate monetarism with Conservative, Republican or right-wing politics in general. We shall return to this point in a while, however, it is helpful first to establish some general points of reference.

* The author wishes to thank David Llewellyn, Charles Goodhart, Gordon Pepper, Geoffrey Wood, Robert Sedgwick and Kevin Dowd for comments on an earlier version of this paper. None of these shares responsibility for the errors remaining or for the opinions expressed which are the sole responsibility of the author.

I WHAT IS MONETARISM?

In his famous essay 'The Structure of Monetarism', Thomas Maye set out twelve propositions which could be used to distinguis monetarists from (implicitly) Keynesians. Discussants of Mayer paper added others of their own, and no doubt virtually every schola would want to modify the definition or put different emphasis upo key elements. This does not show that monetarism was incoherent for the same range of differences would be found among the membe of any broadly based school of thought (Keynesians, Marxist Socialists or whatever). The unifying factor in the early days o monetarism was disagreement with the main strands of Keynesia macroeconomic policy strategy. These may be summarised as follow

(1) The level of aggregate demand can and should be 'fine tune by means of fiscal policy to ensure that the level of activit in the economy is maintained at a consistently high ('ful employment') level.
(2) Incomes policies should be used in order to control any infla tionary tendencies, which are inevitably of the 'cost–pus variety.
(3) Monetary policy should be largely subordinated to the needs o fiscal policy. In any event, 'money' exerts no independe influence upon activity or inflation, and there is no unambiguo aggregate money stock or stable demand function for money.

The unifying doctrine of the monetarists is based upon an especiall strong rejection of proposition (3). The rallying cry of the monetar is: 'Money matters!' Macroeconomic policy-makers ignore moneta policy at their peril. Excessive tightening of monetary conditio causes the onset of depression and excessively loose monetary poli causes inflation. The aim of monetary policy should be slow a steady expansion of monetary aggregates, not massively in excess the trend rate of growth of the real economy. What monetarism *do not entail* is the proposition that monetary growth should be tighten hard in periods of rapid inflation in order to squeeze inflation out the system rapidly. The traditional monetarist is a Fabian at hear believing that gradualism is always better than shock tactics. Howeve there are many modern advocates of monetarism who prefer t 'rational expectations' approach of rapid adjustment to new info

mation. Monetary policy is an inaccurate and blunt instrument whose impact is subject to 'long and variable lags'. Hence a steady long-term strategy is the only serious option available.

One source of confusion is worth eliminating at the outset. This is the belief among some Keynesians that a sufficient destruction of the monetarists is to show that the money stock is endogenous. In other words, part of the Keynesian justification for proposition (3) is the argument that money is *caused by* developments elsewhere in the economy, rather than being a cause of them. Monetarism, it is claimed, only makes sense if money is an exogenous variable. Fortunately, monetarists are not naive about monetary institutions. As Mayer (1978, p. 26) observes, 'the money stock is partly endogenous, being some distance removed from central bank action'. The money stock is not an instrument of monetary policy but this does not mean either that it should not be a target, or that there are no instruments which will enable the central bank to control the money stock.

There are common circumstances in which the money stock is demonstrably endogenous to the economy and yet monetary policy could be totally 'sensible' by monetarist criteria. The obvious case is of a small, open economy which follows the policy of fixing the exchange value of its currency to that of another (monetarist!) country. Fixing the exchange rate is a monetary policy. The authorities have to buy and sell domestic base money (in exchange for foreign currency) in order to supply just that quantity which will be willingly held. The domestic monetary aggregates are demand determined. Is this fatal to monetarism in some way? Of course not. Indeed, pegging to a neighbour who has a stable valued currency may be the best way to ensure an inflation free environment and stable monetary conditions which are the goals of all monetarist policy advisers.

Even in a relatively 'large' economy which has a floating exchange rate the authorities may behave in such a way as to make the monetary base endogenous. In practice, the broader aggregates are always likely to be partly endogenous – at least in conceivable institutional structures. For example, if they choose to peg interest rates rather than money, they will be obliged to supply the reserves which the banking system needs in order to avoid interest rate changes. The mere fact that they follow this policy does not prove that a focus on, say, the reserve base would not be preferable, or

that monetarist predictions about the costs of such policies would not be valid. In other words, if the authorities choose by their own behaviour to endogenise the base, this endogeneity cannot be taken as a demonstration that monetarist arguments about the link between monetary aggregates and inflation are invalid. Indeed, monetarists would argue that the authorities should change their behaviour and/or the institutional structures where necessary in order to ensure that monetary aggregates can be controlled.

An even more vexed question in the UK at present is that of which monetary aggregate to target. This is a technical question which requires a technical answer and a simple answer will not be attempted here. However, it is important to be clear that apparent instability in the velocity of previously favoured aggregates does not establish the negative proposition that no aggregate need be targeted. This is especially true when a large degree of the instability may be the direct or indirect result of institutional changes instigated by the authorities themselves. In the United States the abolition of Regulation Q, and in the UK the imposition and then removal of the 'corset' are obvious examples of the authorities causing relationships to shift. Many other important legal changes have had effects on the UK monetary system, such as the abolition of exchange controls and the derestriction of building societies. These problems do not prove that monetarists were wrong to rely upon stable velocity relationships but merely that the authorities' own behaviour makes it much more difficult to identify what the underlying behavioural relationship really are. Deregulation has turned broad monetary aggregates into 'savings' balances as opposed to pure 'transactions' balances. Rising interest rates on these savings balances make them more attractive to hold and as a result velocity falls.

A popular excuse used by officials in the Treasury and the Bank of England for having no explicit policy on major monetary aggregates is known as Goodhart's Law. This is interpreted as saying that any stable relationship will break down as soon as it is used for policy purposes. The validity of this proposition depends only upon the fact that the authorities in the past have attempted to use behavioural relationships in an absurd way. Targeting £M3 and hitting the target by suitable manipulation of the monetary base is one thing. Hitting the targets by imposing quantitative ceilings on banks' savings deposits is quite different. In the latter case, the intermediation which is being arbitrarily restricted will find another channel and the old demand for £M3 function will appear unstable. This does not prove that

money no longer matters, merely that the appropriate measure of money will have to change as the institutional structures change.

The implication that any monetarist would draw is that the monetary authorities have to have *some* target for monetary growth. Just because their own behaviour makes the identification of that target more difficult does not give them an excuse for abdicating responsibility. Needless to say, the suitable aggregate to target need not be any of the standard simple sum measures Mn ($n = 0 \ldots ?$) but, rather, may be some kind of weighted divisia style index. What is to be avoided is the 'seat of the pants' attitude of 'we look at everything' which means that anything can be justified by the gut feeling of those in power and not by objective economic criteria. In practice this means that the monetary aggregates are largely ignored because they all behave differently, and a nominal interest rate target is adopted instead. The reasons why this is the wrong thing to do have been explained *ad nauseam* by monetarists over the past twenty years and will not be repeated here (see, for example, Poole, 1970).

If the monetary aggregates really are impossible to interpret, there is a straightforward, alternative way of guaranteeing monetary discipline and this is to peg the exchange rate to a major inflation-free currency, the obvious candidate for the UK being West Germany. Indeed, it would probably be even better if we actually had the same money. Such a solution is totally consistent with monetarism though it has naturally not received much attention from monetarists in the USA. Indeed, this policy was followed in the 1950s and 1960s when the peg was to the US dollar – a period of low inflation and stable monetary conditions. Stability in that period is often claimed to be the result of Keynesian fiscal policies. In fact it was due to the combined effects of the monetary discipline imposed by fixed exchange rates and the stability of the external world in which other major countries, notably the USA, were also pursuing conservative policies, at least until the mid-1960s. Indeed, the system broke down in the early 1970s because of the perception that the USA was starting to export inflation.

Assertions as to the relevance of monetarism are of some use in reassuring the believers but will probably do little to convert the infidels. At least assertions can be seen for what they are. We are now about to resort to the somewhat devious tactic of presenting some empirical evidence. This evidence carries the usual health warning. It may be indicative but it is certainly not conclusive. While it does not prove that the monetarists are right, it certainly does not

show that they are wrong. We will attempt to provide support for the following three propositions, each of which would be widely accepted by monetarists and rejected by Keynesians.

(1) Velocity will ultimately turn out to be a stable (though not constant) relationship.
(2) Money causes inflation.
(3) The main effects of fiscal policy are on the public sector component of output. (This is consistent with the monetarist view that fiscal policy can affect the composition but not the level of output, except in the very short run. Money growth, on the other hand, can stimulate output in the short run, but with a zero cumulative effect.)

II VELOCITY PRECOCITY

One of the key empirical propositions popularised by monetarists is that of a predictable long-term relationship between money and money income. The velocity of circulation of money is defined as the ratio of nominal GNP (or GDP) to the money stock. Velocity has not exactly been constant throughout recorded history, but monetarists believe that it is predictable. In other words, the kind of things that influence velocity over time can be explained. Stable velocity implies that an increase in the money stock will be matched by a rise in nominal income. A further rule of thumb, derived from historical experience, suggests that such rises in money will first affect real output but will eventually come to be entirely reflected in higher prices.

In several major countries, most notably the USA and UK, the apparently stable trend in velocity suddenly changed course at the beginning of the present decade. In the United States velocity of M1, which had grown at an annual rate of about 3.4 per cent since the Second World War, suddenly started to decline sharply. Its decline continued until 1987 (see Stone and Thornton, 1987, for a discussion of the recent developments in the USA). Similarly, in the UK, the velocities of both M1 and M3, which had been rising steadily since the war, both started to decline.

These events caused jubilation in the anti-monetarist camps. Even the Thatcher government, allegedly the most committed to monetarism, abandoned targets on M1 and £M3, though even the

had really been adopting a pragmatic approach for most of the time, involving unstated nominal interest rate and exchange rate targets. What they failed to appreciate, of course, was that the changed behaviour of the monetary aggregates was almost entirely a result of the institutional changes that the government itself had initiated. The danger is that the authorities will draw the incorrect conclusion that money no longer matters and can safely be ignored. This is certainly not the case. Monetarism was not properly tested in the UK because, in practice, the authorities adopted multiple targets (including interest rates and the exchange rate as well as many different monetary aggregates). Excessive monetary growth always has and always will be a cause of inflation. Any government which forgets this does so at its peril.

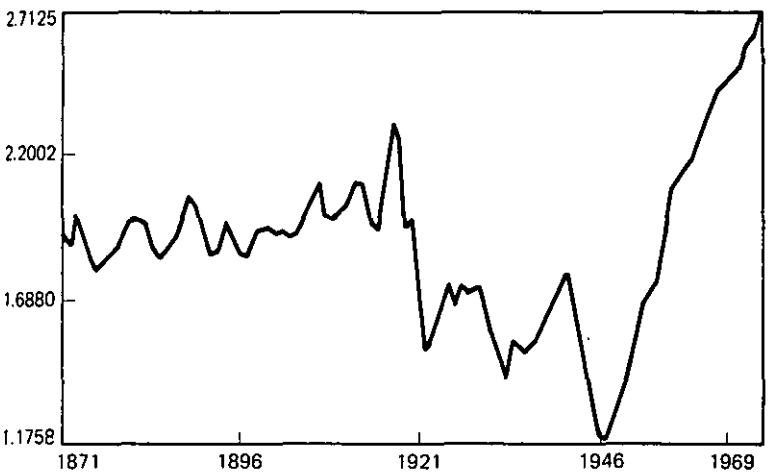

Figure 3.1 Velocity of M3, 1871–1969

Figure 3.1 shows the behaviour of M3 velocity between 1871 and 1969. From 1871 up to the First World War there was a cycle about a gradually rising trend. Both wars were associated with a sharp fall in velocity, as was the great depression. Since 1946, however, velocity (of both M1 and M3, see Figure 3.2) rose steadily over time. Figure 3.3 (on quarterly data and hence with a different scale) shows that the general trends in the velocities of both M1 and M3 continued to rise until about 1980, after which time they started to decline.

The sharp falls in velocity associated with wartime are easy to explain. Black market activity increases the demand for cash relative to

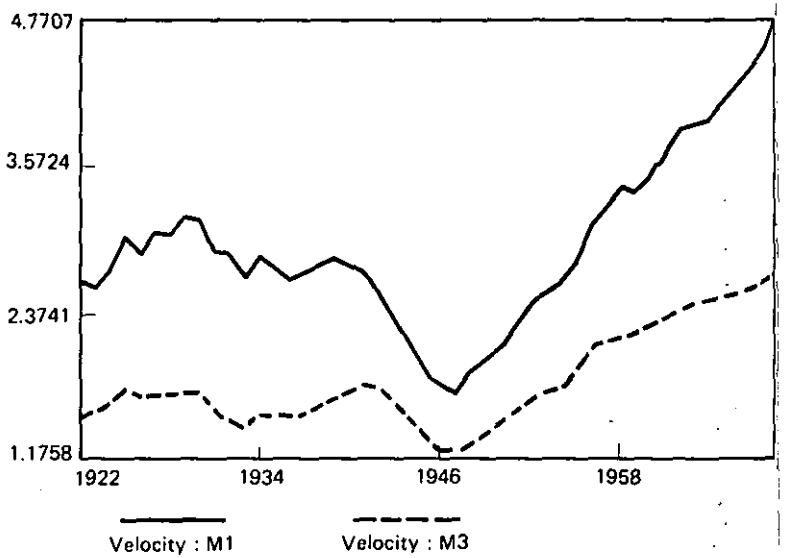

Figure 3.2 Velocity of M1 and M3, 1922–69

other assets and price controls restrain the growth of nominal income. What needs some explanation is why velocity grew so fast from 194 to 1980 as compared to the growth rate in the 1871 to 1914 period. Part of the explanation is usually said to be because of the acceleration in innovation in the payments technology. Such things as credit cards, cheque cards, cash dispensers, standing orders, wire transfers, etc. have increased the efficiency of payments technology so that individuals and firms require lower money balances relative to their income or turnover as time goes by. Equally, the growth of competition from institutions which provide a service which is a close substitute for that provided by banks – such as building societies – has made people hold less of their savings in the form of bank deposits than might otherwise have been expected.

There are probably elements of truth in all these arguments. However, it is something of a puzzle that the period in which velocity has been declining is the period in which both innovation has been at its fastest in the payments technology and the building societies have been most free to compete with banks. A plausible answer lies in the changing regulatory framework.

From the Second World War until 1971, banks were severely restricted in their ability to compete for deposits. Throughout this

period they were either subject to quantitative ceilings on their lending growth or to the potential threat of such ceilings. This meant that there was no incentive to compete for deposits. Under these somewhat artificial circumstances it is hardly surprising that the ratio of nominal income to money should rise. This rise was halted in 1971 with the introduction of the *Competitive and Credit Control* reforms. Quantitative ceilings on bank lending were removed. For a time M3 grew very rapidly and its velocity declined. The velocity of M1 did not decline as there was a great deal of substitution out of current accounts and into savings accounts. A new restriction on the banks, known as the corset, was introduced in late 1973. This once again restricted the banks' ability to compete for savings deposits; it had the original intention of protecting building societies from competition and mortgage owners from higher interest rates, but also of putting a brake on the ability of banks to finance an expansion of lending by attracting wholesale deposits at competitive interest rates. During a brief break in the corset (1978) M1 velocity declined and M3 velocity stopped rising. Once the corset was removed in June of 1980 velocity of M3 and M1 declined almost continuously to date.

Figure 3.3 also shows the velocity of the *stock* of bank lending. This is of interest for two reasons. First, it clearly shows the effect of the 1971 reforms in lowering the velocity of the outstanding stock of banks loans. The corset obviously worked in the opposite direction. Second, since the abandonment of the target range for £M3 growth, commentators have been looking at bank lending figures as a potential guide to monetary conditions. The data show that bank lending has a considerably less stable velocity than £M3 and hence is a less reliable guide to monetary conditions.

Monetary history suggests that the behaviour of the velocity of major monetary aggregates (and perhaps we will need to look at new ones like M4) will eventually settle down in the sense of returning to earlier trends. When the present wave of innovation has worked itself out, velocity will return to a modest upward trend. This could be sooner rather than later, and it is to be hoped that the authorities will be alert to this possibility as the tangible form it will take will be a rise in the UK inflation rate. However, it is possible that the broader monetary aggregates will come to be considered not as 'money' but rather as part of 'credit' intermediation. Attention will then revert to the narrower aggregates which are a purer measure of the stock of the transactions medium in the economy.

There is one aggregate whose velocity has remained remarkably

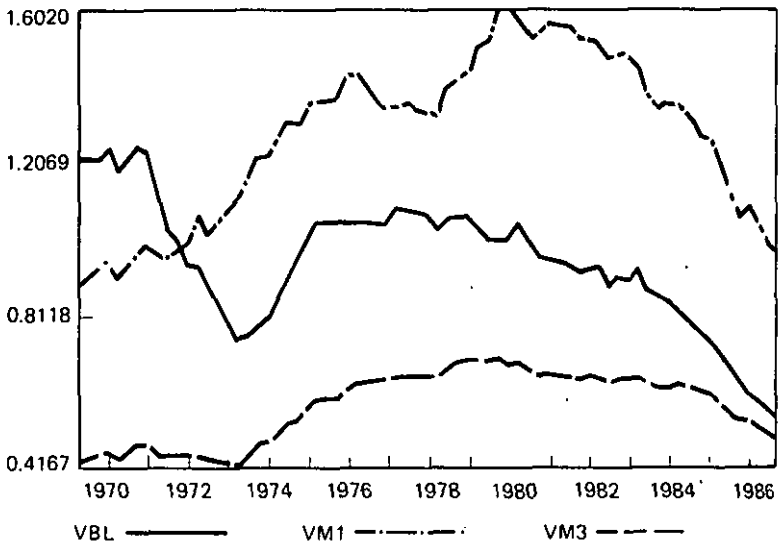

Figure 3.3 Velocity of bank lending, M1 and £M3 (1970i–87iii)

close to a constant time trend over the past two decades. This is M0. Figure 3.4 plots the log of the velocity of M0 and its regression line against time. It has grown at a fraction over 1 per cent per quarter since 1970. At present, this is the only aggregate for which there is an explicit growth target. Although many monetarists believe that base targeting is the appropriate way to implement a non-inflationary monetary regime, it does not appear to be the centrepiece of UK monetary policy at present. Certainly there is a target range for M0 (currently at 1 to 5 per cent) but it seems clear that monetary policy is actually conducted by influencing interest rates and that the target range for M0 would be abandoned like the rest if it proved inconvenient. None the less, the very stability (of the trend) of the velocity of the monetary base may be important support for the monetarist approach. We shall return to a discussion of the 'which aggregate?' question below.

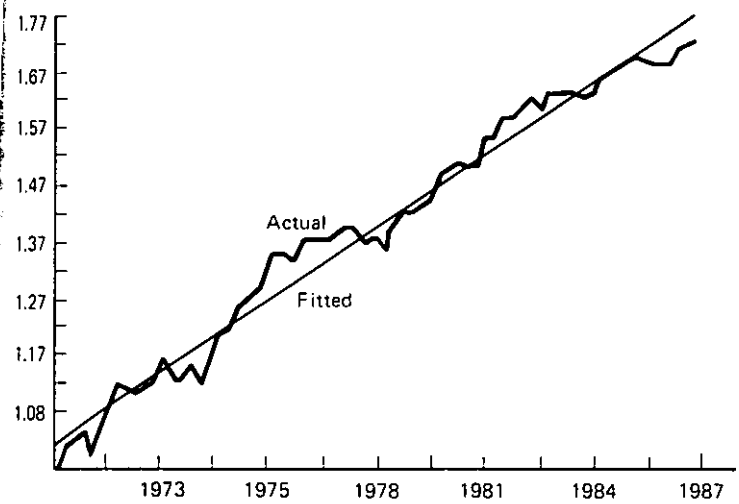

Figure 3.4 Regression of log velocity of M0 against time

III MONEY CAUSES INFLATION

It is tautologically true that money and inflation are closely connected. Not without good reason did Milton Friedman utter his famous statement: 'Inflation is always and everywhere a monetary phenomenon.' Inflation *is* a decline in the value of money, hence it would be ridiculous to suppose that money was not involved in the question in some fairly direct way. It has been known for centuries that a significant rise in the money stock will become associated with a more or less proportional increase in the price level. Equally, a rise in the price level will not be sustained without a parallel rise in the money stock. The direction of causation can go either way, depending upon the institutional structure, as we have seen above. This in no way invalidates the monetarist proposition that controlling the money stock will ultimately control the inflation rate. Reverse causation is always likely to be evident if the authorities are pegging interest rates or exchange rates, and, even if they are not, only the monetary base will be truly exogenous. However, policy determined or exogenous changes in the money stock do take place and these will have effects on prices in a unicausal sense.

We now look at the evidence of the link between money and both prices and activity. Two different data sets will be examined, annual

data produced by Capie and Webber (1871–1969 for M3 and 1922–69 for M1), and quarterly data (1970i–87iii for M0, M1, M3 and the stock of bank lending). Granger causality tests were used to determine whether there is any *prima facie* evidence for the proposition that money causes inflation. Tests were also conducted to see if there was any evidence of reverse causality. Of course, these tests do not prove causation in any absolute sense. All they do is indicate if, say, money can contribute to the statistical explanation of prices in addition to the past behaviour of prices themselves.

On the annual data set the basic autoregression contains five lags of the dependent variable. We then add the money stock lagged one up to five periods and test for the joint significance of these variables. A significant F statistic at the .05 probability level is taken as a rejection of the hypothesis of no causality, in the Granger sense. The results are shown in Table 3.1.

The results indicate that there is very strong causality running from M3 to both the CPI and the GDP price deflator. There is also causality running from M3 to real GNP, but the reverse causality from real GNP to M3 is even stronger. The results for M1 appear less satisfactory from a monetarist perspective. M1 causes real GNP but is itself caused by the CPI and GDP price deflator. However, notice that the data period is very different. The period 1922–69 is dominated by either years of fixed exchange rates or years of wartime price controls. It is highly plausible that reverse causation would dominate in this period.

The quarterly results reported in Table 3.2 use the same technique. The lag length used was six quarters and, in addition to M1 and M3, tests are reported for M0 and the stock of bank lending. What is discovered is that there is strong causation running from M0 to nominal GDP, the price level, the inflation rate and real GDP. There is reverse causation running from nominal GDP and prices to bank lending, and from prices and real GDP to M1. The surprising nature of these results is the very clear suggestion that M0 has causal significance, not only in prices and inflation but also for GDP, nominal and real. It is also clear that there is no causation running from prices or output to M0. So there is evidence here suggestive of the conclusion that M0 is genuinely exogenous in this period and that it has causal influence on prices and output. Monetarists should not be particularly surprised, but non-monetarists may be.

Another interesting result relates to the irrelevance of bank lending as an aggregate with any informational content. While bank lending

Table 3.1 Granger causality tests: money income and prices

1871–1969			
M3 causes:	GNP	F(5,83)	1.05
	CPI		4.57**
	PGDP		3.92**
	RGNP		2.70*
M3 caused by:	GNP		1.05
	CPI		1.67
	PGDP		1.13
	RGNP		5.06**
1922–69			
M1 causes:	GNP	F(5,32)	1.49
	CPI		1.10
	PGDP		0.74
	RGNP		2.56*
M1 caused by:	GNP		2.09
	CPI		3.34*
	PGDP		3.86*
	RGNP		1.38

Notes: GNP is nominal, CPI is the consumer price index, PGDP is the GDP price deflator, RGNP is real GNP. A* indicates significance at the .05 level. A** indicates significance at the .01 level.

is caused by GDP and PGDP, it has no causal influence upon anything else. This confirms our earlier suggestion that bank lending is not a useful monetary indicator. It is true that M1 and M3 do not do well in this sample either, but we have seen that the authorities could be as much to blame for this as anyone. Certainly, there is no result here which suggests that anything can be learned from looking at bank lending that could not be learned from M3. The latter, if anything has a more stable relationship with money GDP.

IV FISCAL POLICY AND THE REAL ECONOMY

An integral part of the monetarist controversy has been the question of the relative leverage of monetary and fiscal policy over the economy. The main debate began after the publication of Friedman

Table 3.2 Causality between money prices and output (Quarterly 1970i–87iii)

GDP caused by:	BL	F(7,51)	1.66
	M1		0.69
	M3		1.08
	M0		3.03*
GDP causes:	BL		2.47*
	M1		0.87
	M3		1.24
	M0		0.92
PGDP caused by:	BL		1.00
	M1		1.67
	M3		1.60
	M0		4.05**
PGDP causes:	BL		2.27*
	M1		4.09**
	M3		1.25
	M0		1.21
INF caused by:	GBL	F(6,51)	0.87
	GM1		1.39
	GM3		1.72
	GM0		4.07**
INF causes:	GBL		0.95
	GM1		1.24
	GM3		0.50
	GM0		0.41
RGDP caused by:	GBL		1.32
	GM1		1.78
	GM3		0.50
	GM0		2.28*

Notes: GDP is nominal GDP, RGDP is real, PGDP is the GDP price deflator, INF is the rate of change of PGDP, BL is the stock of bank lending, GBL is the growth rate of the stock of bank lending, GM1 is the growth rate of M1, etc.

and Meiselman (1963), but it reached its peak in connection with the results produced by Andersen and Jordan (1968). The question addressed by this earlier literature was the relative importance of monetary variables as opposed to autonomous expenditures (especially government expenditure) in influencing activity (nominal or real). We look at this issue from a slightly different perspective.

From a Keynesian view, an expansion of the net demand by the government should have a positive multiplier effect on output. Monetarists, however, tend to believe that an expansion of government demand 'crowds out' demand from elsewhere. However, this does not mean that an expansion of government spending has no effect. The government is not a pure consumer of private output, it constitutes a 'production' sector in its own right. Civil servants, etc. are assumed to produce an output equal to their resource cost (mainly wages) and this output is measured as a component of GDP. Hence an increase in government spending which involved hiring people would automatically increase national output. In a fully employed economy, the resources purchased would be lost elsewhere and the expansion of government output would lead to a reduction of output elsewhere. This negative effect would be fairly diverse and may be hard to pick up. However, there is a clear implication that the output effect of an expansionary fiscal policy would appear principally as an increase in the imputed output of the government sector itself. We test this by means of Granger causality tests and show that, indeed, the fiscal policy measure (change in the cycle-adjusted deficit as a percentage of GDP) does cause output changes in the two components of the public sector but not anywhere in the rest of the economy. For other evidence in the same spirit see Chrystal and Dowd (1987).

Once it is accepted that there are advantages to examining the impact of fiscal policy at a disaggregated level, it will be seen that there are similar advantages from doing the same thing for monetary policy. One of the major doubts surrounding reduced form tests of the impact of money on GDP was the question of exogeneity. At a sectoral level it is much more plausible to argue that the aggregate money stock is exogenous to individual sectors. A change in aggregate money growth is far more likely to be causal for output growth in a small sector than is the reverse to be true (see Chrystal and Chatterji, 1987). The monetarist rule of thumb is that a monetary expansion affects output after about a year and prices after about two years. Output subsequently declines so that the long-run effect is zero, or even negative, but money should show up as having a significant effect on output over a horizon of a few years.

Accordingly, we tested for the impact of money growth on the output of all the major sectors of the UK economy. Tests reported are only of the hypothesis of no causation in the Granger sense. We do not test for the zero cumulative effect. Table 3.3 shows the results

of Granger causality tests using M1 growth, M3 growth and CBD. The latter is the change in the cycle-adjusted budget deficit as a per cent of GDP. The dependent variables are the growth rates of real output of major sectors of the economy – real GDP and industrial production are also included for comparison. The only sector omitted is the imputed income from the housing stock, so the ten sectors add up nearly to GDP. Tests were on annual data (1952–85) and three lags were used, though contemporaneous values of the exogenous variables were included.

The results are remarkably clear cut. The fiscal policy variable causes ADMIN and EDU. These are the two main public sector components of GDP. The former is public administration and defence while the latter is health and education. Of the rest, five sectors (manufacturing, construction, transport and communications, distribution, and services) are caused by M1; in four of those very strongly so. This also shows up in the aggregates GDP and industrial production. M3 growth is causal in two sectors (construction, and energy and water supply). Not only is the fiscal variable not significant outside the public sector but it is nowhere near to significance in any sector except agriculture.

These results cannot be taken as anything other than *prima facie* support for the proposition that money matters. Also it appears that fiscal policy does much for public sector output.

V FURTHER ISSUES

Before drawing some conclusions from our arguments there are several issues which require discussion. Some of these are central to our theme, others are merely a source of external confusion.

Political Monetarism

We have discussed monetarism as if it were a doctrine about money and the conduct of monetary policy. In some contexts the term is taken to have much wider meaning than this. In the UK, it has sometimes been used to describe the entire policy set of Mrs Thatcher. This would include policies on cutting back public spending, privatising industries, deregulating markets, etc. This 'new right' agenda may well be a good thing and it may also have the support of

Table 3.3 Money and fiscal policy and sectoral output

GMANUF caused by:	GM1	F(4,26)	8.05**
	GM3		1.41
	CBD		0.16
GADMIN caused by:	GM1		1.38
	GM3		1.29
	CBD		3.95*
GEDU caused by:	GM1		0.54
	GM3		2.11
	CBD		2.97*
GCON caused by:	GM1		5.58**
	GM3		2.81*
	CBD		0.28
GTC caused by:	GM1		4.03*
	GM3		0.67
	CBD		0.21
GDIST caused by:	GM1		6.51**
	GM3		2.21
	CBD		0.17
GIBF caused by:	GM1		0.79
	GM3		0.25
	CBD		0.49
GSERV caused by:	GM1		4.33**
	GM3		0.89
	CBD		0.16
GEWS caused by:	GM1		1.74
	GM3		2.98*
	CBD		1.35
GAGRI caused by:	GM1		0.33
	GM3		1.98
	CBD		1.85
GIND caused by:	GM1		7.14**
	GM3		2.89*
	CBD		0.19
GGDP caused by:	GM1		6.60**
	GM3		1.64
	CBD		0.03

Notes: Dependent variables are all in growth rates. Sectoral output measures are: manufacturing; public administration and defence; health and education; construction; transport and communications; distribution; insurance, banking and finance services, energy and water supply; agriculture, forestry and fisheries; industrial production; and GDP. All are at constant prices. GM1 is the growth rate of M1; CBD is the change in the cycle adjusted budget deficit as a per cent of GDP. Data are annual 1952–85.

the majority of monetarist fellow travellers. However, this is well beyond the scope of the set of policies that is under discussion here. We have referred elsewhere to this broader agenda (Alt and Chrystal, 1983) as Political Monetarism. This is all part of a much bigger debate between capitalism and socialism, the state and the market, planning versus competition. This is really a separate question from the one of interest here. We are concerned only with macroeconomics, and especially monetary policy. Questions of industrial organisation and microeconomic structure are quite different.

The monetarism that is under discussion in this chapter is not a question of ideology, political or otherwise. It is a set of arguments and methods for determining the nature of technical and empirical relationships and for drawing policy conclusions from objective analysis. It is quite possible for a socialist, who believes in public ownership of major industries and central planning, to be monetarist, in the sense of believing that control of the money stock should be used to control inflation. Indeed, many would count both Keynes and Marx among the ranks of monetarists on this issue.

Post RE Monetarism

Since the mid-1970s, macroeconomic policy has been subjected to a new critique at the hands of the, so-called, New Classical School. This group has been associated with the introduction of the assumption of Rational Expectations (RE) into policy analysis. The new approach led first to the argument that only surprise increases in the money stock would have real effects. More important, perhaps, was the implication that government and the economy were not independent. Actors anticipate government actions just as government anticipates private behaviour. Optimal policy in this context is much more difficult to determine and, on some interpretations, systematic stabilisation policy is impossible. A further implication is that the economy will react fairly quickly to 'news'. This is in contrast to the traditional monetarist view of long and variable lags. It is beyond our scope to go in great detail into the issues which arise. However, it is safe to report that the policy implications of the new analysis are much more supportive of a monetarist approach than a Keynesian one. Monetarists have always supported 'credibility' and caution plus an avoidance of the illusion of fine tuning. For Keynesians, however,

fine tuning is their *raison d'être*. Without it Keynesian economics really is dead.

Particularly damaging to Keynesian demand management has been the so-called 'Lucas Critique' (Lucas, 1976). Lucas argued that the traditional strategy of setting fiscal policy after first evaluating the likely effects by means of simulations on an estimated forecasting model was invalid. The problem is that the parameters of the models are estimated in periods when a particular policy regime has been in place. Changing the policy regime will actually make private behaviour change as well (because of the interdependence mentioned above). This means that the outcomes of the policy changes will not necessarily be anything like what was expected. If this argument is accepted it has two important implications. First, fiscal policy has to be much more cautious than was hitherto thought necessary, and, second, the Keynesian strategy of building forecasting models based upon estimates of aggregate expenditure functions is never going to be an appropriate tool for the determination of 'optimal' fiscal policies. By default, the monetarist strategy of setting longer term rules and establishing a stable and credible policy environment appears to win the day.

Which Aggregate?

Because monetarists are not sure what is happening to major monetary aggregates and cannot say with confidence which aggregate should be targeted, critics have claimed that monetarism is dead. The obituaries are premature. Certainly there are problems, as we have indicated above. However, this does not prove that the monetarist strategy is inappropriate. The relationship between money and inflation has been evident for centuries and it is certainly not going to disappear now. Problems arise in that there is a confusion between 'money' and 'credit'. The latter is not directly related to inflation but the former is. In practice they are horribly entangled. Evolution of the financial system has meant that drawing the line between money and credit is very difficult and far from unambiguous. Money is the stock of the transactions medium. Credit is the volume of borrowing and lending. Deposits of various kinds fit both. Current accounts have long been thought of as money, but what about chequable savings deposits?

The position has been complicated recently by changes in regu-

lations which have enabled bank deposits (chequable) to pay interest and, in the UK case, formerly pure savings media (like building society deposits) to be drawn into the payments system. These developments make it even more difficult to decide what assets should be included in 'money'. However, it makes it much more likely that the suitable monetary aggregate for control purposes will be a very narrow one. There is no well-established explanation of the rate at which the volume of borrowing and lending should rise in an economy. Neither is there any strong reason why the monetary authorities should wish to limit this rate of growth. However, there are good reasons why they should want to limit the rate of growth of money and they are all to do with controlling inflation. Our evidence would suggest that a rigid control of the growth rate of M0 might suffice in this regard.

We do not wish to leave the impression that the answer to the question of which monetary aggregate to control is simple and clear cut. However, we are clear that the mere existence of these problems is not a refutation of monetarism. Monetarism has to change with its environment and answers will certainly be found. Those answers will involve a clear strategy to control the growth of 'money'. How that is best defined remains for another day, and indeed, the appropriate definition may well change over time – especially if the legal framework continues to change. However, the necessity of a stable and credible monetary policy has now gained universal acceptance, and it is on that basis that we can safely say that the monetarists have won.

VI CONCLUSION

Monetarism grew out of the study of history and as a reaction to the neglect of monetary policy in the post-war Keynesian strategy for managing the economy. In the UK, the Keynesian theory was not matched by practice because the pegged exchange rate regime imposed a tough monetary discipline. Only when the disastrous consequences of uncontrolled monetary expansion showed up after the advent of floating was it necessary to worry about domestic monetary control as a deliberate act of policy. Rapid monetary growth in the 1971–3 period produced equally rapid inflation. This lesson was quickly learnt by the 1974–9 Labour government. They introduced monetary targets, and in many ways were more successful

in hitting their targets than their Conservative successors. None the less, the authorities in the UK have not yet resolved the key issue of how to guarantee stable monetary conditions without resort to either quantitative ceilings or excessive swings in nominal interest rates. Monetarists cannot relax yet, there is work to be done, and, given the tendency of politicians to let judgement dominate principle, there always will be work for monetarists to do.

References

Alt, James and K. Alec Chrystal (1983) *Political Economics* (Brighton: Wheatsheaf).

Andersen, L. C. and J. L. Jordan (1968) 'Monetary and Fiscal Actions: A Test of Their Relative Importance in Economic Stabilization', Federal Reserve Bank of St Louis *Review*, vol. 50, pp. 11–24.

Capie, F. H. (1983) *A Monetary History of the UK: Volume 1 Data, Sources and Methods* (London: Allen & Unwin).

Chrystal, K. Alec and Monojit Chatterji (1987) 'Money and Disaggregate Supply in the United States, 1950–1983', *European Economic Review*, vol. 31, pp. 1211–28.

Chrystal, K. Alec and Kevin Dowd (1987) 'Would a Higher Fiscal Deficit Stimulate the Economy?' *Fiscal Studies* (February).

Friedman, Milton and David Meiselman (1963) 'The Relative Stability of Money Velocity and the Investment Multiplier in the United States, 1897–1958', in Commission on Money and Credit, *Stabilization Policies* (Englewood Cliffs, NJ: Prentice-Hall).

Gould, Bryan, John Mills and Shaun Stewart (1981) *Monetarism or Prosperity?* (London: Macmillan).

Kaldor, Nicholas (1982) *The Scourge of Monetarism* (Oxford: Oxford University Press).

Lucas, R. E. Jr (1976) 'Econometric Policy Evaluation: A Critique', in (K. Brunner and A. H. Meltzer (eds) *The Phillips Curve and Labour Markets*, Carnegie–Rochester Conference Series on Public Policy, vol. 1 (Amsterdam: North Holland) pp. 19–46.

Mayer, Thomas (1978) *The Structure of Monetarism* (New York: W. W. Norton).

Poole, W. (1970) 'Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model', *Quarterly Journal of Economics*, vol. 84, pp. 197–216.

Stone, Courtenay C. and Daniel L. Thornton (1987) 'Solving the 1980s' Velocity Puzzle: A Progress Report', Federal Reserve Bank of St Louis *Review* (August–September) pp. 5–23.

4 Credit, Broad Money and the Economy

Tim Congdon

Academic monetary economists often squabble with bankers and business economists about the precise meaning of credit and money, and about their implications for the economy. The aim of the present chapter is to clarify and resolve the key issues in these debates. It has two main themes. The first is that, in modern circumstances, the growth of money is driven by the growth of credit. Money and credit are nevertheless distinct and separate categories, and should not be confused. The second is that, in any economy, the amount of money has a strong and definite link with the amount of spending. As a result, when the amount of money changes sharply, there are profound short-run effects on the way people and companies behave, and so on the level of economic activity. In the long run, however, money cannot alter the economy's ability to produce real output and changes in the quantity of money mainly affect the price level.[1]

Professor Goodhart mentioned in a previous chapter that historically money has taken a great variety of exotic forms, including such objects as red feathers and cowrie shells. The evolution of money is a fascinating and important subject, and one of its key lessons needs to be strongly emphasised. This is that in the past societies have used such a diverse range of things as 'money' that grand generalisations in monetary economics should be treated with suspicion. In this chapter, the discussion will be confined to the circumstances of a modern economy with banks and a central bank.[2] The aim will be to provide an account (a 'special theory') of credit and money that is valid in contemporary market-based industrial economies. The same story could not be told in a pre-modern economy without banks or central banks; nor would it be altogether convincing today in a poor developing country or in a command economy like the Soviet Union's; and it might be totally misleading as a description of the operation of high-tech economies in the future.

I MONEY IS A LIABILITY OF THE FINANCIAL SYSTEM

The first point to highlight in a definition of money is that money has to be recognised as such by large numbers of people. Esoteric objects such as Chinese porcelain vases or Byzantine icons may be 'worth a lot of money', but they are not money as such. They could not be used to buy groceries from a corner shop or timber from a builders' merchant: Instead money comprises a fairly limited range of assets which can be used to pay for goods and services everywhere within a particular monetary area.

There is another key dimension to the definition of money. In his chapter Professor Goodhart argued that money consists of 'those assets that represent a means of payment'. The remark might seem straightforward enough, but he added a subtle – and vital – amplification. This was to say that one characteristic of such assets was that their transfer 'completes a transaction'. By so doing, Professor Goodhart excluded credit facilities which allow a transaction to take place but still leave a debt to be settled.

It is obvious that notes and coin are money under this definition. If payments are made with notes and coin, purchases are completed when they are handed over to the vendor. The purchaser has the goods, the vendor has the money and nothing remains to be done. Again, if payments are made by cheque against bank deposits, the purchaser has the goods, the vendor has the cheque, the purchaser's bank deposit is debited by a particular amount, the vendor is credited by the same amount and nothing remains to be done. But, if payments are made by cheque against a loan facility, the purchaser has the goods, the vendor has the cheque, the purchaser's loan account is debited by a particular amount, the vendor is credited by the same amount *and the purchaser has to repay the bank at some future date.* In this final example, the transaction is not completed even when the cheque has been cleared. It follows that notes, coin and bank deposits are money, but loan facilities are not. Similarly, proofs of creditworthiness (such as credit cards) may greatly reduce the inconvenience of buying and selling, but they are not money. We have here a very sharp distinction between credit facilities and money assets. There is no need for confusion.

Indeed, it is sufficient for most purposes to think of money as constituted by notes, coin and deposits. The issue can be complicated by devising different definitions of money, each of which includes a

specific range of monetary assets. Thus, we can think of an aggregate which consists of only notes and coin, and call it M0. Or we can think of another which includes notes, coin and deposits (so-called 'sight deposits') which can be spent without giving advance notice to a bank, and call it M1. In fact, in the UK today there are six M's, ranging from M0 to M5. The higher is the number attached to an M, the greater is the range of money assets included and the larger is the money supply concept under consideration. M0 and M1 are commonly called the 'narrow' definitions; M2 is an intermediate measure, usually described as consisting of transactions balances; and M3, M4 and M5 are measures of 'broad money'. (The precise definitions are given in Table 4.1.) But the basic idea – that money consists of notes, coin and deposits, and the money supply may be defined as some mix of these ingredients – is straightforward.

It is clear that notes, coin and deposits share the characteristic that they can be used to pay for goods, services and assets. But, in a modern economy, they also have something else in common. This is that they are liabilities of financial institutions, particularly the banks. Thus, notes are issued by, and are a liability of, the Bank of England. Similarly, if money is held in a bank deposit, the bank owes money to the depositor and must follow instructions with regard to payments. The bank deposits are evidently the banks' liabilities. Finally, since it is increasingly possible nowadays to write cheques against building society deposits, they are beginning to resemble bank deposits and can properly be regarded as money. But they are also liabilities, this time of building societies.

It may seem unnecessary to labour the point that nowadays all forms of money are liabilities of financial organisations. But there is an important reason for emphasising it. By so doing, we are alerted to the uniqueness of the monetary system in a modern economy. In earlier times (such as the eras of red feathers and cowrie shells), money was not a liability of financial systems, but a commodity. In other words, money had value not because a particular bank recognised an obligation to its depositors or holders of its notes, but because the commodity had intrinsic worth. The realisation that money could perform its functions without being a specific commodity was one of the key institutional innovations which made possible the emergence of advanced industrial economies.

Despite the benefits of modern monetary arrangements, nostalgia for commodity money is widespread and deeply rooted. It takes two particularly notable forms. First, sceptics of governments' ability to

Table 4.1 Relationships among monetary aggregates and their components

	Non-interest bearing component of M1		
plus	Private sector interest bearing sterling sight bank deposits	*plus*	Private sector interest bearing retail sterling bank deposits
equals	**M1**	*plus*	Private sector holdings of retail building society shares and deposits and national savings bank ordinary accounts
plus	Private sector sterling time bank deposits	*equals*	**M2**
plus	Private sector holdings of sterling bank certificates of deposit		
equals	**M3**		
plus	Private sector holdings of building society shares and deposits and sterling certificates of deposit	*plus*	Private sector holdings of foreign currency bank deposits
less	Building society holdings of bank deposits and bank certificates of deposit, and notes and coin	*equals*	**M3c**
equals	**M4**		
plus	Holdings by the private sector (excluding building societies) of money market instruments (bank bills, Treasury bills, local authority deposits) certificates of tax deposit and national savings instruments (excluding certificates, SAYE and other long-term deposits)		
equals	**M5**		

Source: *Bank of England Quarterly Bulletin*, May 1988.

manage 'paper money' yearn for the financial stability commonly, although perhaps mistakenly, attributed to the gold standard. Secondly, some economists (including such well-known monetarists as Milton Friedman and Karl Brunner) continue to theorise about economies with commodity money, apparently unaware that this approach is not fully applicable to economies with paper money. There is not enough space here to explain the difficulties to which this confusion gives rise. It is sufficient to say that many of the most heated debates in monetary economics stem from a lack of clarity about whether propositions relate to commodity-money or paper-money economies.[3] The discussion in the rest of this chapter relates to a modern economy in which money is explicitly a liability of financial institutions.

II A KEY DISTINCTION

Before we discuss the creation of money, one more idea needs to be developed. Although notes, coin and bank deposits are *all* money, a sharp distinction should be drawn between two forms that they take. Certain kinds of money are legal tender and must be accepted in law as a means of payment. In the UK today, these are represented by coins (a liability of the Royal Mint) and notes (a liability of the Bank of England). But there are other kinds of money which are not legal tender and it is not an offence to refuse payment in them.

Thus, I am fully within my rights to turn down someone's cheque. The writer of the cheque has no legal redress against me or against his bank. In effect, when I refuse a cheque I am indicating two things. First, I am not convinced that the writer of the cheque has enough legal tender in his bank account to honour the cheque and, secondly, if he does not in fact have enough legal tender, I am not prepared to hold a claim on the bank concerned. In some circumstances – for example, when a cheque is drawn on a bogus bank without capital or assets – I would be a fool to accept a cheque instead of legal tender. In the UK today we can, for virtually all practical purposes, regard notes and coin as legal tender, while other forms of money (bank deposits, building society deposits) are not.[4]

The last two paragraphs have a critical implication for the behaviour of interest rates. When I write a cheque, I am giving someone a mere scrap of paper. Why does this piece of paper have any value? The answer is that it is an instruction to my bank to pay the person or

company named a sum *in legal tender*. An obvious corollary is that the bank could not conduct its business unless it held legal tender among its assets. It is true that nowadays the practice of modern banking is so sophisticated that most cheques are cleared by the cancellation of debits and credits between the banks themselves. Banks do not need to make large and cumbersome payments in notes and coin either to each other or to their customers. Nevertheless, they must have the ultimate ability to make payments in legal tender.

The imperative need for banks to meet demands on them in notes and coin is the origin of the Bank of England's power to determine interest rates. The Bank is the monopoly issuer of legal-tender notes. It can therefore fix the interest rate at which these notes are borrowed and lent.[5] Since bank deposits are expressed in terms of legal tender and should be fully substitutable with them, the Bank of England's interest rate (variously described as 'Bank rate', 'Minimum Lending Rate', 'seven-day dealing rate' and so on over the years) is the key interest rate in the monetary system. Since there is no other issuer of legal tender, there is no other institution which can dispute the Bank's sway over interest rates.

This conclusion is of great significance. The operation of monetary policy has been a constant topic of debate in the UK in recent years, with uncertainty about how interest rates are set being a leading source of contention. There is no need for this uncertainty. Although there are a number of details to fill in, the essential message of our argument is plain and should be uncontroversial. In a modern economy interest rates are decided by the central bank. The power to determine interest rates is derived from the central bank's position as the monopoly supplier of legal tender. Its influence over interest rates is not based on convention and it does not survive because of the commercial banks' inertia.[6] Moreover, in principle, the central bank does not have to pay the slightest attention to 'market views'. It is true that, in the real world, central bankers are not known for intellectual iconoclasm and therefore try to respect the market consensus about where interest rates should be. But it is also true that there is nothing logically inevitable about this interplay of ideas between the markets and the authorities.[7]

III MONEY IS CREATED BY CREDIT

The nature of money in a modern economy – that it is a liability of financial organisations – has an important consequence. The liability side of any balance sheet can expand only if the assets side also expands. Banks and building societies increase their assets by making loans to their customers. It follows that money is created as a result of this extension of credit, while the rate of monetary growth is governed by the rate of credit expansion. In a pre-modern economy more money could come into being only if more of the monetary commodity was actually produced. Credit expansion, on the other hand, requires merely the simultaneous registration of debts (i.e. deposit liabilities) and assets (i.e. bank loans, mostly). The ability to create money by a stroke of a pen is strikingly efficient in cutting down on the quantity of resources needed to operate a system of payments. It constitutes a major advance in a society's productivity.

Unhappily, the negligible cost of producing money in a modern economy has the drawback that the issuers of money may be tempted to create an excessive amount. The result may be an inflationary process, with money losing value relative to other things and a consequent loss of confidence in the currency. This risk exists with privately owned banks, but it is subject to a tight constraint. Because their deposit liabilities are not legal tender, they must not allow their deposits to increase too much in relation to their holdings of legal tender. The quantity of bank deposits therefore cannot expand without limit if the quantity of legal tender is fixed or rising only gently over time. In fact, decades of monetary experience have shown that bank deposits tend to be a fairly stable multiple of the amount of legal tender money.

However, central banks are not subject to the same discipline as privately owned banks. If they (or their political masters) decide to issue legal tender money in reckless and inflationary profusion, they are not breaking the law and neither do they (or their political masters) have to worry about going out of business, although they may have to worry about the outcome of the next general election. The dangers of an inflationary overissue of credit-based money have to be balanced against the benefits to society from the trifling cost of creating it. This dilemma, which is at the heart of the controversies over monetary policy in a modern economy, is neatly captured in the title of a pamphlet – 'Proposals for an Economical and Secure Currency' – written by the famous British economist, David Ricardo,

in 1816. Credit-based money is economical in terms of the resources required to make it. But it is potentially insecure in value if too much of it is made. The responsibility for prudent monetary management ultimately falls on the central bank, since – as we have seen – the quantity of bank deposits cannot run out of control if the quantity of legal tender is limited.

The key points of the discussion so far may now be summarised. In a modern economy money is a liability of the financial system, particularly of the banks. Because of this property the growth of money is governed by – indeed, for most practical purposes, can be equated with – the growth of bank credit. The central bank, notably the Bank of England in the UK, can try to control the quantity of money by varying the rate of interest. It has the power to determine interest rates because it is the monopoly supplier of legal tender. Privately owned commercial banks, whose deposits are not legal tender, must kowtow to the Bank of England's interest rate decisions as they dare not risk being unable to convert their liabilities into legal tender.

We must emphasise, before we proceed to consider the impact of money on economic activity, that there is no muddle about the relationship between credit and money in our theory. To say that 'money is created by credit' is *not* equivalent to saying that 'money is credit'.

IV MONEY, ECONOMIC ACTIVITY AND MONETARY EQUILIBRIUM

Once money has been brought into being by credit expansion, what is the relationship between money and economic activity? Before answering this question, it is best to digress briefly to consider the relationship between any set of objects in the economy. For example, the economy produces each year a certain number of apples and pears. Market forces – the laws of supply and demand – establish a price ratio between the two fruits which keeps their producers profitable and their consumers happy. We can call this ratio, which satisfies buyers and sellers so fully that they have no wish to change the situation, an equilibrium ratio. If the quantity of apples rises or falls dramatically (because of the discovery of a new seed, a crop disease or whatever), but the quantity of pears stays the same, we would expect the relative price of apples and pears to change sharply.

There will be another equilibrium price associated with the new supply conditions. But the passage from one equilibrium price to another may involve disturbance and uncertainty, and we would not expect the new equilibrium to be attained instantaneously.

We could tell the same story about the relative price of bricks and mortar, or coal and electricity, or any other combination of goods and services we care to think of. Associated with each equilibrium price are also particular quantities of each good. If the quantities change, it is likely that the relative price must also change. The essential point is that there is an equilibrium relationship, in terms of both price and quantity, between any good and all other goods. When this equilibrium holds, there is no tendency for people or companies to try to upset it. The same set of prices and quantities continues from one period to another. The economy is at rest. Only if there is an unexpected change (in demand or supply conditions) is the equilibrium broken.

It does not take much imagination to think of money as just another 'good'. Indeed, it is particularly easy to think of it in this way since the prices of all goods are expressed in terms of money. If market forces establish the relative price of apples and pears (i.e. the number of apples required to buy one pear, say, 1½) they also establish the relative price of apples and money (say, 6p) and the relative price of pears and money (4p). The idea can be extended and generalised. If there is an equilibrium relationship between money and any particular good, there must also be an equilibrium relationship between money and national output as a whole. When this equilibrium holds, there is a particular level of national output (expressed in terms of £b, to put the idea in a UK context) and a particular amount of money (also in £b). Associated with the equilibrium is a price level of all goods and services taken together. In monetary equilibrium the demand for money (i.e. the quantity of notes, coin and bank deposits people want to hold) is equal to the money supply (i.e. the quantity of notes, coin and bank deposits actually in existence).

The concept of monetary equilibrium is not universally respected in the economics profession. Some of its critics think that it leads on too readily to the ambitious – and politically controversial – claim that the money supply and money national income tend to move together over time. In fact, any careful statement of the meaning of monetary equilibrium recognises that there are many influences other than income on the amount of money people want to hold.

Three deserve to be separately identified. The first comes under

the general heading of 'payments technology'. The more efficiently payments can be completed, the less money is needed in relation to income. For example, a society in which credit cards are widely used is unlikely to need as much ready cash (in proportion to national income) as one where they are unknown. Also important in this context are such institutional characteristics of the economy as the frequency with which people receive wages and salaries, and the preparedness of companies to defer payments to each other (e.g. by extending trade credit).

Secondly, the rate of interest people and companies receive on money affects how much of it they wish to hold. Interest is not paid at all on notes and coin, and there are still some bank accounts (e.g. the traditional current account) which do not pay interest. But nowadays the majority of bank deposits, and practically all building society deposits, pay interest. When we are considering people's desire to hold money relative to other assets, the key consideration is the rate of interest received on money *relative to the rate of return on these other assets*. When the general level of interest rates rises, people will want to cut down on their holdings of notes and coin because the relative attractiveness of these non-interest bearing assets has declined. *But it is possible, indeed quite likely, that the return on interest bearing bank deposits will have improved relative to the return on other assets and that people will want to hold a higher ratio of interest bearing money to income.* (We will return to this point – which has an important bearing on the interest rate sensitivity of the economy – later.)

Thirdly, it is clear that the expected rate of inflation affects attitudes towards holding money, since every increase in the price level reduces the real value of money balances. A high rate of expected inflation makes it worthwhile to keep wealth in the form of goods and tangible assets rather than money.

In fact, there are so many potential influences that we cannot hope to be comprehensive in a short discussion. But we can give an adequate summary by saying that the desired ratio of money holdings to national income depends on three main considerations – transactions technology, the rate of interest (or, better, the interest rate differential between money and non-money assets) and inflation expectations. *If these influences are stable, it is reasonable to expect the desired ratio of money to income to be constant.*

This is not a particularly bold or ideological statement. It is plain common sense to say that the number of apples people wish to

consume depends on how tasty they are, how expensive they are compared to pears and oranges, and how quickly they rot if they are not stored properly. Our remarks on money run on very similar lines. We can analyse the demand for money in much the same way as we analyse the demand for other things.

V SOME IMPLICATIONS OF MONETARY EQUILIBRIUM

Once we accept that, with certain conditions satisfied, the desired ratio of money to income is constant, some vital consequences follow. The most important is that an increase of x per cent in the money supply must be followed by an increase of x per cent in money incomes, and so in the nominal value of expenditure and output, if people are again to be happy with their money holdings. If national income does not rise by x per cent immediately, monetary equilibrium has been violated and people will change their behaviour until national income does rise by x per cent. We can think of an increase in national income as having two parts, an increase in output and an increase in the price level. If output is fixed, it is only the price level that can respond to the monetary injection. Indeed, monetary equilibrium requires that the x per cent increase in the money supply must be matched by an x per cent increase in the price level.

This does sound like a bold and ideological statement. It is undoubtedly very 'monetarist' in flavour. But our argument does not imply that, in any examination of actual data over a period of years, there will be a precise link between the money supply on the one hand and national income and the price level on the other. First, it has been emphasised that a precise link would be found *only if influences such as transactions technology, the rate of interest and inflation expectations were stable.* In practice, the character and strength of these influences are always changing, and their variations greatly complicate the relationship between money and prices. Secondly, the statement about money and prices is valid *only if monetary equilibrium has been established.* We have explained that people are always trying to move towards equilibrium. But in the real world the economy may not be in equilibrium. Just as it takes a period of microeconomic disturbance before the relative price of apples and pears adjusts to the discovery of a new seed or a crop disease, so there may be a period of macroeconomic disturbance

before national income and the price level adjust to an increase in the money supply. During this interval of monetary disequilibrium, the connection between money and prices may be difficult to identify.

We will discuss monetary disequilibrium in the next section. But before doing so, some consequences of the argument in the last paragraph need to be emphasised. It is possible both to believe that inflation is always and everywhere essentially 'a monetary phenomenon' (in Friedman's words) and to expect to observe, in the real world, considerable fluctuations in the ratio of money to national income. In policy debates the behaviour of the ratio of money to national income – and of its inverse, the velocity of circulation of money – attracts considerable attention. Many critics of a monetary approach to inflation claim that changes in velocity demonstrate the irrelevance of the money supply. But we can see that these claims are exaggerated and misleading. Indeed, the relevance of the money supply stems, at root, from a belief that the demand for money – like the demand for fruit, building materials or energy – can be analysed with the standard tools of microeconomic theory. All the interesting conclusions about money and prices are derived from the concept of monetary equilibrium. To deny the validity of this concept is also to deny the premise of rationality which is basic to all economic analysis.

VI THE CONCEPT OF MONETARY DISEQUILIBRIUM

The notion of monetary disequilibrium is best understood in relation to that of monetary equilibrium. We have said that when an economy is in equilibrium all prices and quantities set in one period are repeated in the following and subsequent periods. In monetary equilibrium, the demand for money is equal to the money supply and the ratio between money and income is stable over time.

Monetary disequilibrium arises when the demand for money is not equal to the money supply and people are changing their behaviour in order to restore equilibrium. In more familiar language, the amount of money people are willing to hold differs from the amount of money actually in the economy. If people have excess money balances they will seek to reduce them by, for instance, buying goods and services or financial and real assets. Decisions about spending and saving are adjusted until a more settled position, with desired money holdings equal to actual money holdings, is restored. This may sound rather

strange. In all economies at all times there is a particular quantity of notes, coin and bank deposits in existence and this quantity is held by people, companies and financial institutions. Surely, if the money is held at all, it is held willingly. There cannot be a mismatch between the demand for money and the money supply. It seems that the idea of monetary disequilibrium is incoherent and an intellectual *cul-de-sac*.

To dismiss monetary disequilibrium so abruptly is too superficial. A modern economy is extremely complex, with millions of prices being fixed every day only to be changed tomorrow, the day after tomorrow and so on into the indefinite future. At any given moment, the price level – and, indeed, many other characteristics of the economy (including, perhaps, transactions technology, the interest rate and the inflation rate) – may differ from the expectations prevailing when people last took action to adjust their money holdings. Moreover, very few economic agents know precisely how large their money holdings are at every instant in time. It is clear that actual money holdings can differ from the desired level. Monetary disequilibrium is a viable concept.[8]

With this idea accepted as part of our analytical tool-kit, we are almost ready to shift the discussion away from the abstract plane to a practical, real world level. But there is one further argument to develop. Our interest is in how decisions motivated by the behaviour of credit and money impact on output, employment and prices. We are not particularly interested in the behaviour of credit and money for its own sake. A transfer of money from one bank account to another, or from notes to bank deposits, is tangential to our main concern, since these transactions are purely financial and do not affect the 'real economy'. It follows that we need to identify and monitor a measure of the money supply which can make people reconsider their patterns of expenditure and saving. There is no point tracking a measure of money which is irrelevant to expenditure decisions.

VII NARROW MONEY VERSUS BROAD MONEY

The notion of monetary disequilibrium gives us the clue to making the right selection. In equilibrium the demand for money is equal to the money supply, monetary variables are neutral in their impact on the economy and it does not make much difference which particular

monetary variable (notes, coin or deposits; M0, M1, M3 or whatever) is the focus of attention. It is only in disequilibrium that money can disturb behaviour. Our question therefore becomes, 'For what measure (or measures) of money is there a possibility that the holdings people want to have differ significantly from the holdings that they actually do have?' This question could be rephrased more briefly as: 'What measures of money can behave in ways which surprise people and make them reassess their decisions to consume and invest?'.

Notes and coin are the small change of the economy. If people find that their holdings of notes and coin are too small for their requirements (to buy goods and services, mostly), they go to their banks and convert part of their deposits into notes and coin. (If, on the other hand, notes and coin are too large, they leave them on deposit with their banks.) The adjustment occurs through purely financial transactions, which we have already said are incidental to our main concerns. It is also obvious that no person or business organisation allows holdings of notes and coin to affect any major decision about the purchase or sale of large assets (shares, factories, buildings). In an advanced industrial economy, with its massive accumulation of capital assets, these decisions about asset disposition are critical to the economy's behaviour.

We have said enough to reject notes and coin (M0) from consideration. M0 cannot surprise people and make them review their decisions to consume and invest. This narrow aggregate has one further characteristic which needs to be emphasised. We have said that when individuals find that their holdings of notes and coin are out of line with their requirements, they restore equilibrium by transfers into and out of bank deposits. That could leave the banks with too much or too little cash, which creates another problem of adjustment. The banks respond by approaching the Bank of England in order to persuade it either to absorb the excess cash or to eliminate the deficiency. The Bank, which of course issued the notes in the first place, accommodates the banks' requirements as a matter of routine. *A large number of individual decisions to increase (reduce) holdings of notes and coin do lead to an increase (reduction) in the aggregate amount of notes and coin in the whole economy.* M0 adjusts to events in the economy; events in the economy do not adjust to M0.

Nowadays, the contacts between the banking system and the Bank of England are so harmonious, and the Bank's operations are so finely tuned, that the amount of M0 in the economy rarely differs

from the amount people want to hold. M0 is virtually always in or near to equilibrium. One consequence is that econometric work typically identifies a good, close-fitting statistical relationship between M0 and money national income.[9] *But this does not mean that M0 has a strong influence on decisions to spend or on the level of money national income. The direction of causation is rather from money national income to M0.*

Similar remarks apply to other measures of narrow money. M1 is larger than M0 because it includes bank accounts which can be spent without giving notice (sight deposits). But, again, if such bank accounts are too large or small, the natural response is to shift a sum of money to or from accounts which require notice (term deposits). An example is when an individual transfers funds from a current account at a clearing bank to a deposit account. This is clearly a financial transaction without implications for the real economy. Moreover, a host of such individual transfers will change the aggregate amount of M1. If M1 is too high or too low in relation to money national income, it is M1 which adjusts, not money national income.

We can summarise the last three paragraphs by saying that the various measures of narrow money are rarely in major disequilibrium, and even when they are, people and companies bring them back into equilibrium by purely financial transactions. The narrow-money aggregates – such as M0 and M1 – are therefore not the money supply concepts that we are seeking. Instead we need to look at broad money, notably M3 and M4.

VIII BROAD MONEY AND EXPENDITURE DECISIONS

We have seen that when people and companies have too much or too little narrow money, a more appropriate holding is restored – at the level of the whole economy – by switching between different categories of deposit or between deposits and notes or coin. The position is quite different with broad money. Broad money (on the M3 definition) includes all bank deposits in the economy. If the nominal quantity of such bank deposits is fixed by a separate and independent influence (such as the level of bank credit), a host of individual decisions to switch to and fro between different agents' bank deposits or between one type of bank deposit and another cannot change that nominal quantity. It follows that if the nominal quantity of broad money is too high or too low in relation to income,

interest rates or other macroeconomic variables, equilibrium can be re-established only by changes in these variables. This property explains why we must concentrate on broad money, not narrow money, if we wish to understand the link between money and economy activity.

The point may need a little elaboration. Suppose I discover, when I check my bank statement, that my holding of bank deposits is higher than I expected and require. Then I will attempt to shift the excess holding somewhere else. It will not solve the problem to transfer money from a deposit account to a current account (or vice versa) since that would leave the total of my deposits unaffected. The only way I can eliminate my excess money is to spend it on goods and services, or acquire an asset. Both these transactions will add to someone else's deposit, *but they will not reduce the aggregate amount of bank deposits in the economy*. Moreover, although I may eliminate my own excess money holding, the sudden addition of money to someone else's deposit may result in his having excess money holdings. Any one person may think that he can control the amount in his bank account, but

> For all individuals combined . . . the appearance that they can control their money balances is an optical illusion. One individual can reduce or increase his money balance only because another or several others are induced to increase or reduce theirs; that is, they do the opposite of what he does. If individuals as a whole were to try to reduce the number of dollars [or pounds] they held, they could not all do so, they would simply be playing a game of musical chairs.[10]

This game of musical chairs is the economy's attempt to move from monetary disequilibrium to equilibrium. It is not entirely futile. If everyone considers their broad money holdings excessive, they will all, more or less simultaneously, try to disembarrass themselves of the excess by increasing their spending on goods and services, or by purchasing more assets. These efforts will lead to higher aggregate expenditure and, in due course, probably raise the price level. At the new, higher price level, it may well be that the nominal quantity of bank deposits is again appropriate. Indeed, expenditure decisions will keep on being revised until the right balance between money and incomes is restored. While individuals may be

> frustrated in their attempt to reduce the number of dollars [or pounds] they hold, they succeed in achieving an equivalent change in their position, for the rise in money income and in prices reduce the ratio of these balances to their income and also the real value of these balances. The process will continue until this ratio and this real value are in accord with their desires.[11]

We may summarise the message of this section. *A large number of individual decisions to increase (reduce) nominal holdings of broad money does not lead to an increase (reduction) in the nominal aggregate amount of broad money, but instead causes changes in expenditure on both current and capital items. The behaviour of the economy therefore adjusts to broad money, rather than broad money to the behaviour of the economy.*

IX INTEREST RATES AND PRICES

There has now been enough analytical preparation for a rough and ready account of how interest rates, credit and money affect economic activity and the price level. It can be related, if rather casually, to the position of the UK economy over the last twenty or thirty years. Let us suppose that the economy is in approximate monetary equilibrium. Interest rates are set at a level where both the growth of credit and the associated growth rate of broad money are such as to keep output expanding at about its trend rate (say, 3 per cent a year) and inflation is at its average value in recent years (say, 5 per cent). Let us also suppose that – perhaps under political pressure to promote faster growth – the Bank of England cuts interest rates substantially. How would we expect the economy to respond?

First, the growth of credit is stimulated. The explanation is simply that with lower interest rates the attractions of borrowing are increased. If interest rates are cut, there will be a wider range of assets where the return exceeds interest costs and there will also be higher borrowing. Experience in the UK suggests that two kinds of credit – mortgage borrowing for residential property and borrowing by property companies to invest in offices and other kinds of commercial property – are particularly susceptible to interest rate changes. Increased expenditure on these assets often represents the economy's earliest 'real' response to lower interest rates.

Secondly, the faster growth of credit leads to faster growth of

broad money. If broad money growth was previously appropriate to maintain a steady rate of increase in money national income of about 8 per cent a year (i.e. 3 per cent increase in output, 5 per cent increase in prices), it must now be too high. Economic agents discover – because of the quicker increase in the nominal amount of bank deposits – that their money holdings are excessive. For this reason they must think about how their money holdings can be brought into a better relation to their expenditure and income.

But there is yet another reason for adjusting behaviour. As mentioned earlier, in the UK today most deposits are interest bearing. When interest rates are cut, the desired ratio of interest bearing deposits to income is lowered. This effect would stimulate expenditure even if the nominal amount of broad money were constant. Since there is actually more rapid growth of nominal money due to the extra buoyancy of bank credit, the urge to move out of money assets into either current expenditure or non-money assets is doubly strong.

We have explained – in the last section – why the excess holdings of broad money cannot be eliminated except by changes in incomes, interest rates or other macroeconomic variables. In practice, the economy's efforts to restore monetary equilibrium are very complicated and work initially via asset markets (the stock market, the property market) rather than goods markets (i.e. through immediate changes in consumption and investment). For example, when they have 'too much' money in the bank, private individuals switch much of the excess balances to building societies (where they finance the purchase of houses), to financial institutions such as unit trusts and insurance companies (where they become available to buy shares and government bonds) and to companies. Companies can then use the money either to finance stockbuilding and investment, or to purchase more assets (the shares of other companies, or commercial and industrial property such as offices, warehouses and factories). Typically, in the early stages of an upturn, when there is only nascent optimism about future output growth, companies are more eager to buy existing assets than commit themselves to increased expenditure on new capital equipment and buildings.

In other words, a cut in interest rates is often followed in the first instance more by a surge in asset values than by an upturn in output growth. But just as there is an equilibrium relationship between money and national income, so there are an assortment of equilibrium relationships between the market values of capital assets and their

replacement values. If the market value of capital assets is driven far in excess of replacement value by a boom in credit and money, more new investment becomes worthwhile. To talk in terms of 'market value' and 'replacement value' may sound technical, but the underlying economic logic is obvious. After all, if house prices soar above the cost of building new ones, it is only common sense that there should be a surge in housebuilding. In due course, the jump in asset values stimulates higher investment.

The length of the lag between the interest rate cut and the revival in most forms of capital expenditure may confuse economists into thinking that investment – and therefore the economy as a whole – is not sensitive to interest-rate changes. Indeed, it needs to be recognised that a standard feature in the early stages of a boom is that only one kind of investment, in private residential houses, is notably strong. Consumers' expenditure, which is often regarded by economists as little affected by interest rates, may show a more definite response than investment. Closer examination is nevertheless likely to demonstrate that the increase in consumption is concentrated in long-lived items like cars and durables (e.g. furniture, carpets, washing machines). These items are effectively investment by the personal sector and the increased demand for them may be motivated, in large part, by the cut in interest rates.

Once the boom has started it becomes difficult to stop. Indeed, the rise in asset values which reflects attempts to redispose wealth holdings more effectively may give further impetus to credit demand. Some businessmen may be tempted to project a rate of asset price appreciation persistently above the rate of interest and they borrow even more heavily to capture the expected capital gains. Unless interest rates are raised, speculative excitement becomes self-feeding. Credit growth – and therefore the growth of broad money – accelerate further.

Eventually the economy reaches a condition of boom. The rate of real demand growth may be between 5 and 7 per cent a year, far in excess of the 3 per cent trend growth rate. Output may grow at an above trend rate of 5 or even 6 per cent for a time, but in due course signs of strain emerge. In the UK, which has a medium-sized economy highly exposed to international influences, a classic symptom of excess demand is balance of payments deterioration. But other indicators, such as a sharp fall in unemployment and a rise in the proportion of companies reporting capacity shortages, usually tell the same story. The lack of spare capacity in factories now leads to the rapid growth

in manufacturing investment which was missing at an earlier stage in the cycle.

Companies and individuals are, throughout the upswing and the boom, attempting to bring their money holdings into line with their incomes. But with credit growth strengthening because of the emergence of speculative activities in the property and other asset markets, they may find that every time they adjust their behaviour, a new and unexpected addition to their bank deposits throws them out of equilibrium again. The ratio of broad money to their incomes may rise to levels far above the long-run figure they regard as sensible. To put the same point in more technical terms, the velocity of circulation of M3 and M4 may fall substantially beneath its equilibrium value. Strangely, a repetitive pattern in UK cycles at this stage – indeed, virtually a recurrent cyclical phenomenon in its own right – is that economic commentators point to the drop in velocity as evidence of the poor relationship between the money supply and economic activity.

Sooner or later inflation spreads from asset markets to the prices of goods leaving factories and appearing in the shops. The excess demand for all types of products causes shortages which can only be eased by price increases; the decline in unemployment leads to tightness in the labour market which provokes higher wage increases and aggravates the spiral in industrial costs; and the worsening external payments position undermines the pound on the foreign exchanges which increases the price of imported goods, including the costs of many of the raw materials and inputs used in UK factories. At this point the growth rates of M0 and M1 – which were probably unaffected by the asset price surges in the early stages of the boom, but are highly responsive to the higher money value of transactions consequent on rising inflation – may accelerate markedly.

Now, with inflation as well as real output growth moving above its previous trend figure, the government becomes alarmed. It mandates the Bank of England to raise interest rates to restrain the pace of expansion. The higher level of interest rates causes falls in asset prices and deters the more speculative forms of credit. But broad money growth remains high for several quarters, as companies complete the expansion programmes initiated during the boom and take up banking facilities already arranged. Beneath-trend output growth of under 3 per cent is needed for a time to compensate for the excesses of the boom. If the government is lucky, credit expansion, money growth and inflation return – without too much fuss or delay –

to the values associated with the previous condition of approximate monetary equilibrium. However, the price level is x per cent higher than it would have been if interest rates had not been cut in the first place. The value of x is likely to be very close to the excess of broad money growth over the figure that would have occurred if interest rates had been kept constant throughout. The episode of excessive credit and monetary expansion has achieved nothing positive in real terms. But it has imposed on society, even if only temporarily, all the awkwardness and inconvenience of coping with higher inflation.

X CONCLUSION

The sequence of events described in the last section may sound familiar. It is, in the form of a simplified idealisation, the story of the UK economy between mid-1986 and mid-1988. The early 1980s had been a rather tranquil period for the UK economy, as output grew at about the trend rate of 2¾ per cent a year and inflation was steady at about 5 per cent. But a marked upturn in demand and output growth in the second half of 1986 followed a reduction in interest rates from the rather high levels of 1985 (when clearing banks' base rate averaged 12.25 per cent). It gathered dangerous momentum in early 1988 after base rates had dropped to 8½ per cent and below. Share prices soared in the initial phase of above-trend output growth, while property values rose sharply throughout the boom. Serious financial problems eventually emerged, with inflation on the rise and the current account of the balance of payments lurching heavily into deficit. Between June and August 1988 base rates were raised eight times from 7½ per cent to 12 per cent, as the Bank of England tried to compensate for previous monetary looseness.

The behaviour of both real and financial variables during this period is inexplicable except in terms of interest rates, credit and broad money. Some economists have suggested other causes for the rapid expansion of demand and output, but these are all implausible. The world economy was not notably vigorous over these years and, in any case, such strength as it had cannot account for the UK growing faster than the rest of the industrial world. Fiscal policy was somewhat contractionary in effect, even when adjustment is made for the impact of cyclically strong tax revenues in forging a large budget surplus. The claim that the oil price fall of 1986 caused a

significant sterling depreciation, which then stimulated exports, is valid up to a point.[12] But over the two years to mid-1988 imports rose much faster than exports and the change in the balance of payments actually withdrew demand from the economy. The non-monetary explanations of the 1986–8 boom (which may be fairly called the 'Lawson boom' after the Chancellor of the Exchequer who presided over it) are random and miscellaneous; the monetary explanation – which focuses on official interest rate decisions, the upturn in credit expansion in late 1985 and 1986, and the subsequent acceleration in broad money growth – fits the essential facts.

Indeed, the Lawson boom has several incontestable similarities to the Barber boom of 1971–3 and what might be termed the 'Healey boomlet' of 1977–9. At some point in all three of these episodes base rates dipped beneath 8 per cent and gave a clear stimulus to credit and monetary expansion. Apart from these instances, base rates were never at 8 per cent or less in the seventeen years from 1971. The year 1971 is an important landmark since it saw the abolition of artificial restrictions on bank balance sheet growth. When the low level of interest rates had been established, share and property prices rose quickly, demand and output moved forward at above normal rates, and financial difficulties developed. Interest rates then had to be raised to cool the economy down.

If the broad outline of our analysis is accepted, it is evident that the Bank of England has enormous power over the economy. Interest rates are under its absolute control, while interest rate changes cause fluctuations in the growth of credit and broad money, and these in turn cause fluctuations in the growth of demand and output. The Bank of England may abuse its power, perhaps under pressure from overoptimistic Chancellors of the Exchequer. But there should be no doubt about the extent of its ability to determine macroeconomic outcomes. It would be of great benefit to society if the Bank of England's power were exercised more responsibly in future than it has been in recent years.

Notes

1. Strictly, changes in the quantity of money are matched by changes in output and the price level. The effect on prices dominates only in an inflationary economy, where the rate of increase in prices is two, three or more times the rate of increase in output. See pp. 116–20 of Sir Alan Walters's *Britain's Economic Renaissance* (Oxford: Oxford University

Press, 1986) for an example of the claim that money and credit are frequently confused.

2. 'Central bank' is a generic term for the bankers' bank. Nowadays it is invariably banker to the government and its note liabilities are legal tender. But there is nothing preordained about these arrangements which have evolved over centuries. See Tim Congdon, 'Is the Provision of a Sound Currency a Necessary Function of the State?', pp. 2–21 in *National Westminster Bank Quarterly Review* (August 1981), for an outline of the historical development of the existing system.
3. Monetarist economists are known for emphasising that control of the money supply is necessary and sufficient for the control of inflation. Associated with this essentially technical proposition are a number of political beliefs, including a particularly hostile attitude towards state intervention in the economy.

   The author registered his own protest against the failure to differentiate between commodity- and paper-money economies in his 'Has Friedman Got It Wrong?', pp. 117–25 in *The Banker* (July 1983). The same theme appears in Kaldor's 1980 evidence to the House of Commons Treasury and Civil Service Committee, reprinted in N. Kaldor, *The Scourge of Monetarism* (Oxford: Oxford University Press, 1986).
4. There is a trivial exception. The Scottish banks issue notes which, although they are perfectly acceptable for most payments throughout the UK, are not legal tender.
5. In practice, the Bank of England expresses its wishes on interest rates more by setting the price at which it buys and sells seven-day bills (seven-day dealing rate) than by announcing the rate of interest at which it will lend money. The detailed institutional arrangements for money market operations are extremely complicated, but it would not change the basic argument if they were described here. The two key articles are: 'The Management of Money Day by Day', in *Bank of England Quarterly Bulletin* (March 1963) and 'The Role of the Bank of England in the Money Market', in *Bank of England Quarterly Bulletin* (March 1982). They are reprinted in the Bank of England's *The Development and Operation of Monetary Policy* (Oxford: Oxford University Press, 1984).
6. This statement is intended as a direct contradiction of the general argument in chapters 3 and 4 of J. C. R. Dow and I. D. Saville *A Critique of Monetary Policy* (Oxford: Oxford University Press, 1988) and of the particular statement on p. 61 that 'bank base rates are determined by conventions that are largely historically determined, and thus subject to considerable inertia'.
7. The view that short-term interest rates are strongly influenced by market sentiment, and are not therefore under full Bank of England control, has been argued by Professors David Llewellyn and Brian Tew in 'The Sterling Money Market and the Determination of Interest Rates', in *National Westminster Bank Quarterly Review* (May 1988).
8. The idea of disequilibrium money is associated in the UK at present particularly with Professor Charles Goodhart of the London School of Economics and Professor David Laidler of the University of Western Ontario. See, for example, chapter 10 of C. A. E. Goodhart *Monetary Theory and Practice* (London: Macmillan, 1984). But it can be traced

back a long way. Arguably, it is implicit in the distinction between long-run and short-run monetary equilibria in D. Patinkin *Money, Interest and Prices*, 2nd ed. (New York: Harper & Row, 1965), particularly on pp. 50–9, and perhaps can be found in Keynes (notably, according to Richard Coghlan, in two articles Keynes wrote in 1937). (See R. T. Coghlan *Money, Credit and the Economy* (London: Allen & Unwin, 1978, p. 27).

9. See, as regards M0, R. B. Johnston *The Demand for Non-Interest Bearing Money in the UK* (London: Government Economic Service Working Paper, No. 66, HM Treasury, 1984) and, for M1, R. T. Coghlan, 'A Transactions Demand for Money', *Bank of England Quarterly Bulletin* (March 1978).
10. See M. Friedman, 'Statement on Monetary Theory and Policy' given in Congressional hearings in 1959, reprinted on pp. 136–45 of R. J. Ball and Peter Doyle (eds) *Inflation* (Harmondsworth: Penguin Books, 1969). The quotation is from p. 141.
11. Again, the quotation is from p. 141 of Friedman 'Statement on Monetary Theory and Policy'.
12. As argued by Mr Philip Stephens, the economics correspondent of *The Financial Times*, in an article in *The Financial Times* of 6 August 1988.

5 Monetarism and Stagflation

Brian Reading

The task of Congdon and Chrystal was to clarify the issues which monetarism raises. Mine, it may be concluded, is to confuse. There is something in this. But a medical analogy is more apposite. They have been describing the anatomy of monetarism, what happens to a healthy system when all goes right. My objective is to look at the pathology of monetarism, what happens when the system misbehaves. If the result causes confusion, no apology is made for the real world generally does behave in a confused and pathological manner.

It is convenient to introduce the arguments within the framework set by that basic identity of monetary analysis, which must be familiar to all economists:

$$MV = PT$$

where:
M is the stock of money;
V is the velocity of circulation, the average number of times each unit of the money stock changes hands during a given period of time;
P is the average price level; and
T is the volume of transactions during that period.

M is ambiguous, as is well known. There is now a veritable motorway system of alternative definitions from which to choose. And rather like a motorway system, some people appear to switch from one M to another the better to reach some predetermined destination. Britain has frequently changed its target aggregates, from £M3 to M3 to M0; Germany has gone the other way and dropped its long-standing target aggregate, the narrow central bank money stock, after it misbehaved in 1987, and has targeted M3 for 1988 instead. The argument between rival definitions is not our concern here. The

one used in the statistics produced here is close to M2. It is what the IMF statisticians define as 'money plus quasi-money' and that simply means the nearest thing they can get to M2 for each country. It is best for the purpose at hand because it is the broadest measure readily available for international comparisons. Broad measures are probably preferable to narrow ones because switching between different forms of money is more likely to take place within them.

I SUPPLY-PUSH MONEY STOCK GROWTH

M in the identity of exchange is also ambiguous in another sense. It is often loosely called the money *supply*. It is also, of course, the money *demand* as *ex post* demand and supply must be equal. A less misleading term is the money *stock*. *Ex ante*, demand and supply do not need to be equal. Charles Goodhart stressed that most of the determinants of money demand have been common ground to monetary theorists throughout the ages. The problem is to know what follows from a change in supply. It is, in effect, the difference between *demand-pull* and *supply-push* money stock growth.

Demand-pull money stock growth occurs where borrowers actively approach financial intermediaries for additional credit. Intermediaries will oblige provided extra lending can be accommodated within prudential or legal reserve ratio requirements. New loans create deposits, increasing the money stock. Demand-pull growth normally occurs where some new investment opportunity presents itself, such as North Sea Oil, or when people want to buy new products, such as video recorders. A rise in the demand for credit, which pulls out extra supply, may push up interest rates in the process. But as the new money is mostly spent on consumption or investment, increasing the output or the price of new products and services, nominal GDP also rises and the velocity of circulation remains stable or increases.

Supply-led money stock growth occurs when financial intermediaries take the initiative to increase their lending and are able to do so without breaking prudential or legal reserve ratio requirements. Since most intermediaries want to expand their business as much as it is legally and prudentially safe to do so (and many are under competitive pressure to go even further than is prudent), they react to any opportunity to lend more. They get such opportunities when increased government borrowing supplies them with additional reserve assets. Moreover, governments do not always need to run big budget deficits

in order to borrow more from the banking system. They may also borrow to finance large unsterilised exchange market intervention. Where money stock growth is supply-led, the intermediaries actively go out looking for borrowers. They may advertise nationally on television, or may offer credit through selective mail shots.

People encouraged to borrow may choose to consume or invest more. But more likely they will be tempted to purchase financial claims instead. In so doing, they will alter the size and composition of their personal balance sheets, increasing both their financial assets and financial liabilities relative to their incomes. Mostly they will incur financial liabilities fixed in nominal terms, in exchange for financial assets whose value will (hopefully) rise according to market developments. Indeed the change in the size and composition of personal balance sheets has been one of the most striking features in recent British experience (see Llewellyn's chapter to this volume).

When more money is lent to people who buy claims over existing real or financial assets, rather than buying newly produced goods and services, the result is asset price inflation, rising land and property prices, price hikes for collectibles and old masters (see Figure 5.1). Asset price inflation does not immediately cause real or nominal GDP to increase. Feeling richer, people may subsequently borrow more against their enhanced wealth, in order to consume or invest more. But on the whole, supply-led money stock growth can be expected to do little directly to raise GDP and so will lower the velocity of circulation.

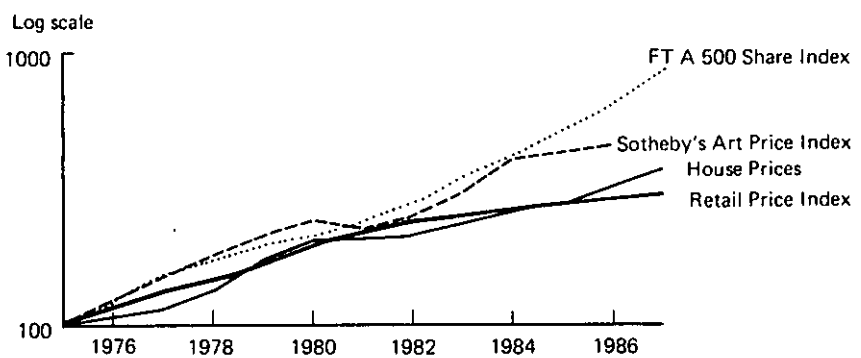

Figure 5.1 Asset price inflation (1975 = 100)

Asset price inflation introduces another complication. People are most eager to borrow to buy real and financial assets when they

expect capital gains. Enhanced money supply triggers asset price increases. The initial rise in price creates the expectation that prices will continue to rise, rather than fall. This then increases the demand for credit, even to the extent that interest rates are pushed up instead of down. Higher interest rates reduce credit demand for normal consumption and investment, depressing the growth in nominal GDP. Where this happens in an open economy, foreign funds attracted by the higher interest rates will produce additional supply-push money stock growth keeping the asset price pot boiling. In such circumstances, higher interest rates have a perverse effect as they accelerate money stock growth.

II A DIGRESSION ON RECENT BRITISH EXPERIENCE

The relevance of this to the present policy dilemma in Britain should now be apparent. But it is worthwhile to sidetrack slightly to describe it. In 1986, Britain had a remarkably healthy economy with a faster

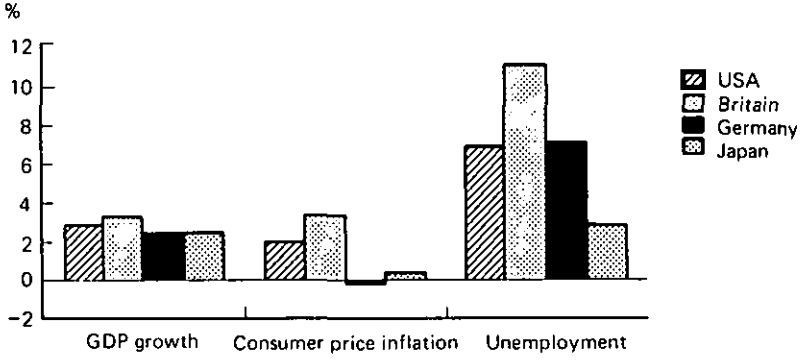

Figure 5.2 Britain's relative performance in 1986

rate of growth than in other industrial countries (Figure 5.2). Savings and investment were broadly in balance at home without the need for any substantial Budget deficit to keep demand growing satisfactorily. The external payments were in near equilibrium. Britain's growth rate was one of the fastest in the OECD and inflation was at a twenty-year low. Interest rates were, however, on the high side. Meanwhile, other countries were suffering acute external and internal imbalance

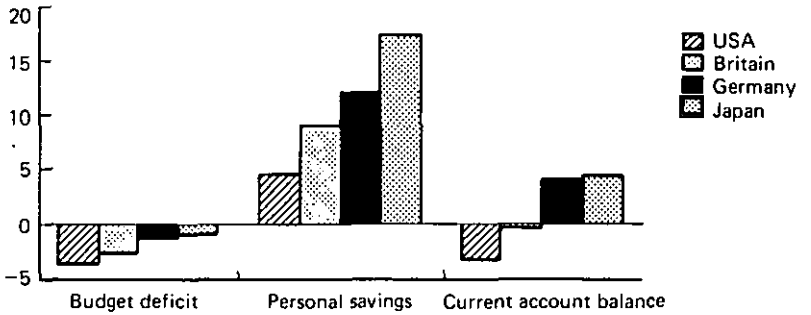

Figure 5.3 Britain's balanced economy in 1986 (percentage of GDP)

(Figure 5.3). The Germans and Japanese saved too much; their economies ran excessive payments surpluses. They were in danger of sliding into deep recession. Americans spent too much; their economy was in excessive payments deficit and in danger of overheating. Obviously, the disequilibrium between surplus and deficit countries had to be corrected, and obviously this would involve some rise in the Deutschemark and yen and some fall in the dollar. It is clear that we should have stood apart from these adjustments. If sterling went down all the way with the dollar, that would upset our equilibrium causing us to run trade surpluses and risk higher inflation. But equally if the pound went up with the Deutschemark and yen we would be pushed into trade deficit.

To keep the pound in the middle, it had to fall against the Deutschemark and rise against the dollar. To achieve this, interest rates should have been cut. But because the pound had been unwisely pegged to the Deutschemark, they were not cut. Yet so healthy and attractive was the British economy at these interest rates that foreign capital poured into this country. In today's world, major countries' exchange rates are dominated by capital flows. Trade balances adjust to these capital movements and not the other way round. A big capital inflow causes the trade balance to move into large deficit, instead of a large deficit causing capital to flee the country. So when foreign money poured into Britain in 1987, the government's main choice was how would it like its trade deficit, with or without inflation?

The government could let the exchange rate rise to deter the capital inflow. But that would have priced British exports out of world markets. The trade balance would move into deficit, causing the

economy to stagnate. Alternatively, the government could instruct the Bank of England to intervene in the foreign exchange markets to stop the pound rising, which is what it did. The Bank of England spent £14.5bn in 1987 to stop the pound going higher, buying a $20bn hoard of foreign currencies. The pounds the Bank spent on the government's behalf had to be borrowed, by the sale of an additional £14.5bn of government debt.

The effect of this on the money supply could not be sterilised by an 'exact funding' policy, without destroying the gilt-edged market. Instead the borrowing had to be partly from the banking system. This increased the banks' ability to expand credit. The authorities consequently lost control of the growth of the money stock in Britain. Here was supply-push money stock growth with a vengeance. It produced asset price inflation, particularly in the housing market, as expected. There was a sharp fall in the velocity of circulation. But the consumer then began to borrow and spend against his new-found wealth. A spending spree followed and a house building boom started. These booms caused economic growth to accelerate to 5 per cent, leading to capacity and labour shortages. The consumer boom sucked in imports, causing Britain's trade balance to deteriorate. In early 1988 we were running a deficit at a rate of £9bn a year. So keeping interest rates too high, while preventing the exchange rate from rising, flooded the economy with money, caused it to overheat at home and to plunge into payments deficit abroad. Yet despite the profoundly damaging and perverse effects from holding interest rates too high, conventional wisdom demanded that they go higher.

Velocity

So far the focus has been on money stock growth, but the concept of the velocity of circulation has already been touched upon. It is now time to finish what has to be said about velocity. It is not measured separately but defined as money *use* divided by money *stock*, PT/M. This definition ensures that MV always equals PT. This would be fine if V were either stable, or changed in a predictable way. That's what the classical economists supposed that it did. But it isn't and doesn't. Figure 5.4 shows how velocity has varied in industrial countries over the past thirty years. It rose during the 1950s, then bounced down to the mid-1970s. It has since gone up and down again. In all fairness to the monetarist, it is admitted that velocity

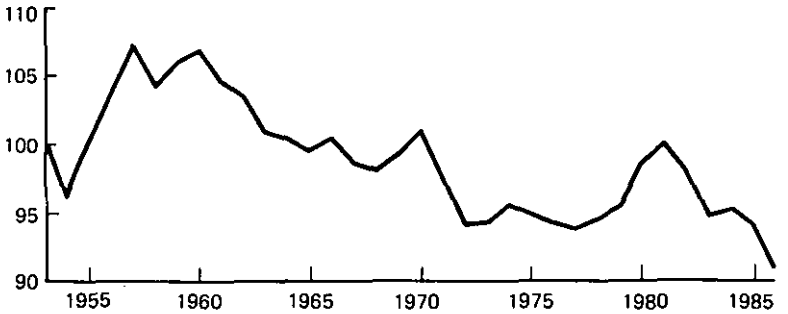

Figure 5.4 Money velocity in industrial countries. Nominal GDP divided by money plus quasi-money stock – Index 1953 = 100

has been most volatile in Britain and America, the countries with the freest and most innovative financial systems.

The volatility of velocity underlies the monetarists' rather unhelpful suggestion that rapid money growth affects prices with a long and variable lag. The same could be said of almost any cyclical time-series with regard to almost any other. But undoubtedly having more money to spend normally encourages people to spend more, even if

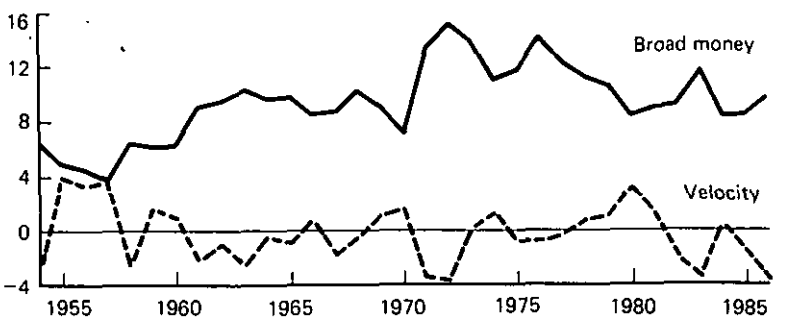

Figure 5.5 Industrial countries' money stock growth and velocity. Percentage change over previous year

the link is not immediate or pound for pound. Indeed as Figure 5.5 clearly shows, there has been a noticeable tendency in recent years for strong movements in the money stock to be smoothed by opposite, though lesser, movements in the velocity of circulation. This strongly suggests that rapid money stock growth is more normally supply-led than demand-led, and it is argued later that this is indeed the case.

Rarely do velocity changes overpower money stock movements to cause MV to move in the opposite direction to M. But it does sometimes happen, particularly when speculation in financial assets is rife. This was the case, for example, in America shortly after the War of Independence. A contemporary wrote that everywhere the scene was one of 'stock gambling, agriculture, commerce and even the fair sex relinquished to make way for unremitted exertions in this favourite pursuit' (Henry Lee, 1791, quoted in Channing, *History of the United States*). Sometimes 'money' can be the more exciting of the two topics which Charles Goodhart observed 'stay in the forefront and in the deeper recesses of our mind'.

Inflation and growth

So far I have dealt with MV, the monetarist side of the equation. The difference between monetarists and Keynesians, as Goodhart has implied, lies in what happens to velocity when the money supply rises. The monetarists reckon it stays more or less constant. Keynes says it can fall. The fall in V as M rises is clear evidence that one may be pushing on a piece of string. But generally a rise in M is associated with a rise in MV. This obviously means that a rise in M is generally associated with a rise in PT. So we now turn to the other side of the equation, to describe the odd things that have been going on with regard to prices and output in recent years.

T is the volume of transactions, the current production and sale of goods and services. It is equivalent to real GDP. It is therefore convenient to use the GDP deflator as the measure of average prices. PT is nominal GDP and T is real GDP. We know nominal GDP normally rises when M rises. The problem is to know how much of the rise in nominal GDP is due to a rise in prices and how much due to a rise in real output. The Classical economists reckoned that most of the rise would be in prices since the economy in equilibrium would be working close to full capacity. Later, as it was seen that the pre-First World War economy followed a cyclical path about its full employment level, the rise in nominal GDP was reckoned to be shared between P and T. When the economy was below full employment, most of the rise in nominal GDP was due to real growth. When full employment was reached, most of the rise was in prices. This concept is still accepted today, but the idea of what constitutes full employment has changed. It is no longer some low

level of unemployment. Instead a circular concept has been invented, NAIRU (the non-accelerating inflation rate of unemployment). NAIRU marks the divide between real growth and inflation, which can be at quite high levels of unemployment. But the important point is that inflation is positively correlated with growth and negatively correlated with unemployment. Faster growth and lower unemployment cause inflation, slower growth and higher unemployment cure it.

Keynes's world does not differ from the Classical world in this respect. But in Keynesian theory the economy can reach equilibrium at a high level of unemployment with prices stable or falling. Keynes's theory is a special not general theory, as the monetarists rightly claim. It requires there to be a liquidity trap which, by preventing real interest rates from falling, stops interest rates from equating savings and investment. When prices are stable or falling, nominal interest rates cannot go below zero – for then nobody would lend – and hence real interest rates cannot be reduced and may rise. That is what happened between the wars. But neither Keynes nor the monetarists cover the possibility that inflation may cause unemployment as well as unemployment curing inflation. Neither envisaged our modern disease, *stagflation*, in which inflation is rampant while economies stagnate. But a look at recent experience shows that this has been the pathological condition of industrial countries' economies in recent years and it is a phenomenon which must be explained.

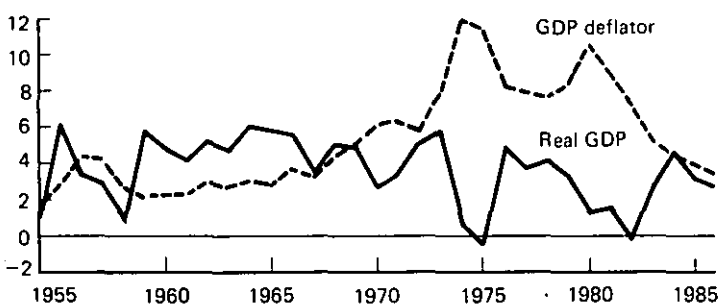

Figure 5.6 Industrial countries' growth and inflation. Percentage change over previous year

Figure 5.6 shows inflation and growth in industrial countries over the last thirty years. From the late 1960s, inflation accelerated as growth slowed down. There was even slumpflation in 1974–5 and

1981–2, when falling output was associated with record inflation

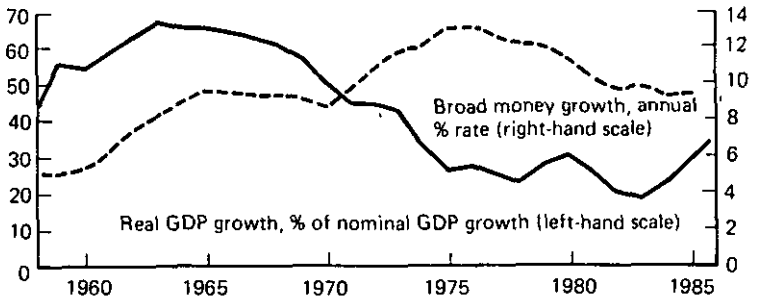

Figure 5.7 Industrial countries' real and money growth. Five-year moving averages

following the two oil price explosions. Figures 5.7 and 5.8 show the share of real growth in nominal GDP growth over consecutive five-year periods. Whereas in the early 1960s real growth accounted for 2 per cent out of every 3 per cent rise in nominal GDP, by the early 1980s inflation accounted for 4 per cent out of every 5 per cent rise in nominal GDP. This worsening of the trade-off between growth and inflation was not associated with economies working flat out with

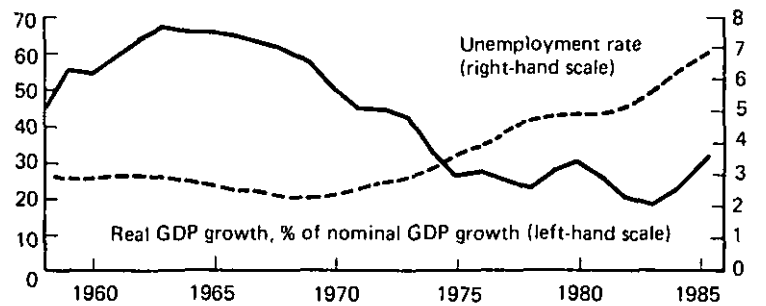

Figure 5.8 Industrial countries' real growth and unemployment. Five-year moving averages

full employment, but rather the reverse. As Figure 5.8 shows, faster inflation was associated with rising unemployment. Figure 5.9 shows that it was also associated with economic growth from the mid-1970s falling well below its 1954–74 trend level. Faster inflation was increasingly linked to weaker not stronger economic activity.

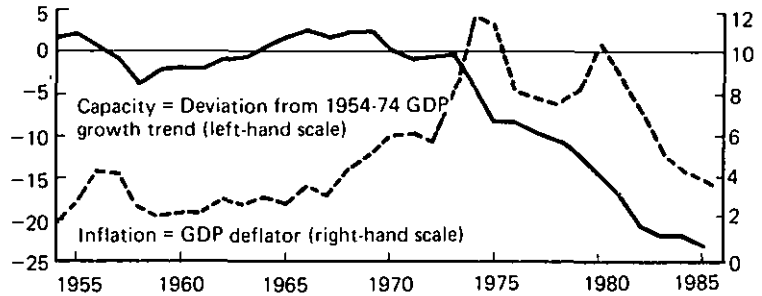

Figure 5.9 Industrial countries' capacity and inflation

The analogy that can be used here is of a car with a slipping clutch. The faster the engine races, the slower the car moves forward and the hotter its engine becomes. The more money growth accelerated from the mid-1960s, as Figure 5.7 shows, the less it drove the real economy forward and the more it pumped up the rate of inflation. In this situation, inflation can still be cured by stamping on the monetary brakes. But this causes the car to stall completely. Monetary stringency then merely exchanges the evil of unemployment for the evil of inflation. It is not even certain whether high unemployment, like an incomes policy, merely defers rather than defeats inflation. It is beginning to look a bit that way in Britain today.

It is ironic to look back at the past and see that in the 1950s and early 1960s, when monetary policy might have worked in the Classical way, it was not actively employed. Everybody was then a Keynesian. Full employment was the first priority. Economies were fine tuned by fiscal policy to achieve this objective, and monetary policy was largely accommodating. While in recent years, when so many economists and governments have become monetarist, the economy simply has not behaved as it should. What then explains the great near-stable price prosperity of the 1950s and 1960s and the lapse into slumpflation in the 1970s and early 1980s?

III THE ANATOMY OF COST-PUSH INFLATION

The answer is that at some time in this period the familiar demand-pull inflation was replaced by cost-push inflation. Demand-pull

inflation is the monetarists' norm and has been superbly encapsulated in the dictum, 'too much money chasing too few goods'. But recently we have been suffering from its opposite, 'too much money chasing too many goods', which equally well encapsulates the phenomenon of stagflation. Cost-push inflation occurs where economic agents collectively demand greater shares in the national income than can be satisfied at existing prices. Prices rise to squeeze out the weakest claimants. The economic agents are:

— *workers and companies* who together produce the national income;
— *savers and lenders* who provide the capital with which it is produced;
— *the government* which takes and spends for us or redistributes income; and
— *foreigners*, with whom we can run surpluses and deficits.

If any of these groups tries to absorb a larger share of the GDP, others must take less or retaliate.

Workers are well placed to maintain or increase their share, provided they don't mind some of them losing their jobs in the process. The trade unions in Britain reached their apotheosis in the post-war years, particularly in the large, publicly owned monopolies. Governments are also well placed. They can always cover public spending by taxing or borrowing more. Savers are strong, provided they do not suffer from money illusion. Once capital markets fear and fully anticipate inflation, realised real interest rates remain positive whatever actually happens to inflation and savers can no longer be cheated out of their wealth.

Foreigners are losers of first resort from rising inflation. When higher prices make an economy uncompetitive, it soon runs into external payments deficit. This means that foreigners initially lend it the money to go on spending beyond its means. But no country can run a payments deficit for ever. Sooner or later its growing debts destroy its credit and the lending stops. Exchange rate depreciation normally follows, which returns the pressure on resources to the domestic economy.

Companies are the weakest contenders in the struggle for resources. Tax costs, wage costs and interest rates can always be pushed up against them. They get caught in the middle of a tax–wage spiral. Workers demand inflationary wages increases, which at existing prices would cause losses at all levels of output. Governments try to stop

companies paying these increases by pursuing austere fiscal and monetary policies. To fight a strike means certain and immediate losses. To give in means later losses. Where unions are prepared to pursue their demands to the point at which companies are put out of business, the only thing management can do is to pay up, put prices up and hope. If the government keeps a tight grip on the money supply, companies then shut plant and fire workers. Inflationary pressures immediately cause higher unemployment. If the money supply is increased fast enough to finance higher prices, and people are willing to spend their enhanced money holdings, the first result is inflation. This is so even where a company has the ability to expand output in response to higher spending. If it is making losses on each unit of output it produces, higher output at existing prices merely increases the number of units on which losses are made. Prices are pushed up instead of and before production is raised.

Faster inflation rather than higher output results from greater money supply growth, even when the economy is depressed. The authorities are then left with a dilemma, whether to deflate because of inflation or reflate because of unemployment. When they choose the former, as ultimately they always do, the efforts to subdue inflation cause even greater unemployment. In a system where cost-push inflation causes unemployment, stagflation results. When tackled by conventional deflation, slumpflation follows. The root of post-war cost-push inflation in the major Western democracies was that collectively voters demanded that governments spent more than they were individually willing, as earners and tax payers, to release from the real output they produced. Stagflation and slumpflation have been the mechanisms which have persuaded governments and voters to change this behaviour.

IV THE GREAT POST-WAR PROSPERITY

Public spending in all OECD countries increased as a share of GDP from the mid-1950s to the early 1980s. At the same time, two decades of full employment and the public commitment to preserve it made unions more concerned to increase their members' real wages than to protect their jobs. These were the powerful factors which caused cost-push inflation. The great stable price prosperity of the 1950s and 1960s was an aberration. A series of favourable accidents prevented cost-inflationary pressures from having a serious effect. But these

accidents, of their nature, could neither persist nor be repeated.

(a) Following the end of the Korean War, in July 1953, defence expenditure fell, putting a brake on the rise in total public spending in the mid-1950s.
(b) At the start of the Korean War there was a commodity price explosion. This was quickly reversed and industrial countries enjoyed a prolonged period of stable or falling food, raw material

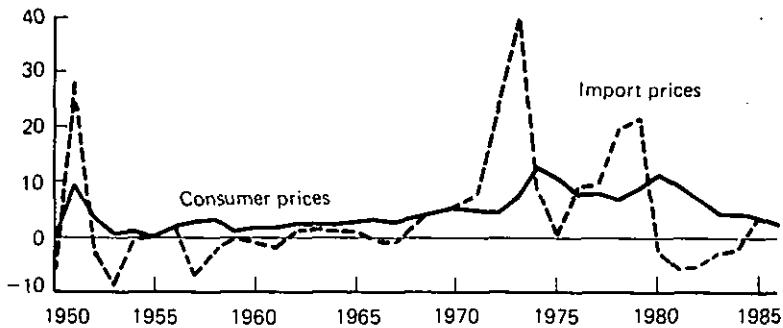

Figure 5.10 Industrial countries' import prices and inflation. Annual percentage growth

and energy prices (see Figure 5.10). Their terms of trade improved steadily against primary producers. In part, this was the result of technological improvement in production, and in part due to the rise in the volume of primary producers' sales. What developing countries lost from lower prices they recouped from increased output. Their economies were also supported by strong foreign investment.

(c) There was money illusion. Inflation was sufficiently slow for governments to operate unindexed tax systems. For years governments got the credit for tax rate cuts, even when the proportion of income taken in taxation sharply increased. Savers also accepted low and sometimes negative real interest rates. So long as the secular rise in inflation was small relative to year-to-year cyclical changes, savers did not anticipate it. Industry obtained cheap capital by borrowing in fixed nominal terms to finance appreciating physical assets. As long as inflation was low, companies made no effort to prepare inflation-adjusted accounts.
(d) The Bretton Woods fixed exchange rate system internally created two natural losers, the reserve currency countries America and

Britain. The Bretton Woods system was more the consequence than the cause of international prosperity. It lasted as long as the sacrifices demanded of the reserve currency countries were bearable. Adjustment to international trade imbalance was through changes in relative growth rates rather than through changes in relative prices. Rich economies can tolerate slow growth in a fast growth world. They will not tolerate falling output and rising unemployment in a slow growth world to preserve the currency system.

V THE GENESIS OF STAGFLATION

By the mid-1960s, falling defence expenditure was no longer making any contribution to the resources available for increased public and private spending on other things. Instead US defence spending began rising strongly as that country's involvement in the Vietnam War escalated. But the collapse of the Bretton Woods fixed exchange rate system in the late 1960s and early 1970s was the main factor leading directly to slumpflation. Through the 1950s and 1960s rival currencies became cumulatively under and overvalued. The reserve currency countries suffered vicious cycles of slow growth, low investment, slowly rising productivity, intensified inflationary pressures, growing payments imbalances and hence even slower growth to protect currency values. The pound and the dollar became increasingly overvalued and the cost in lost growth from protecting them continued to escalate.

Competitor countries, Japan, Germany, France and Italy, enjoyed virtuous cycles leading to their currencies becoming increasingly overvalued. The slow-down in activity needed to correct the imbalance for Britain, and then in America, ultimately became unacceptable. After a long and damaging delay, Harold Wilson's Labour government was finally forced to devalue the pound in November 1967. The Bretton Woods system then progressively collapsed. On 15 August 1971, President Nixon suspended the gold convertibility of officially held dollar reserves. An attempt was then made, in the Smithsonian Agreement of 18 December 1971, to re-peg currencies, with margins widened from 1 to 2.25 per cent either side of their par values. But Britain floated the pound again in mid-1972 and by early 1973 most currencies were floating against the dollar – albeit that

most European Community agreed upon a joint float through the EMS snake.

The Bretton Woods system did not collapse without a fight. In the late 1960s, problems of external imbalance were tackled by allowing deficit countries more time in which to achieve domestic structural readjustment. Their deficits consequently remained larger and lasted longer and thus required greater financing. The result was a dramatic increase in the supply of international liquidity. SDRs were invented and first issued on 1 January 1970. Gold prices rose and between 1969 and 1972 holdings of foreign currencies in national reserves more than trebled (see Figure 5.11).

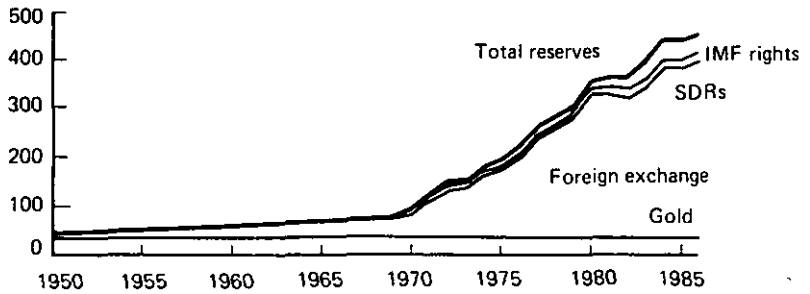

Figure 5.11 World reserves (liquidity in SDRs)

The end of Bretton Woods led to the synchrocycle. When adjustment to payments imbalance was through changes in relative growth rates, countries marched out of step. But floating was seen as allowing payments' adjustment to take place through relative price changes, without deficit countries having to deflate other than by the amount necessary to release resources for the external trade sector. Everybody could enjoy rapid growth. In consequence, a modest world recession in 1970–1 was followed by a powerful and synchronised upswing in industrial countries' activity over 1972–3 (see Figure 5.12). Although the boom was not unusually strong in individual countries (Britain excepted, thanks to Edward Heath's U-turn), the combined effect of countries marching in step was to produce an exceptionally strong boom for them all put together.

Stacks of liquidity, coupled with synchronised growth, led to commodity shortages and speculation. The result was an end to commodity price stability in the first price explosion since the Korean War. Industrial countries' terms of trade began to deteriorate. The

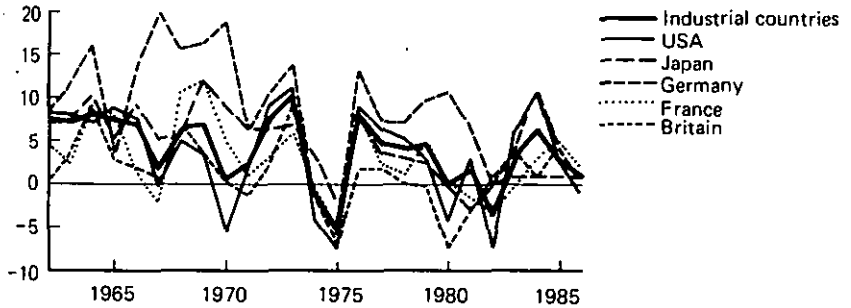

Figure 5.12 The synchrocycle (GDP percentage change on previous year)

deterioration became dramatic over the turn of 1973–4 when oil prices suddenly exploded. Special factors explain the severity of the first oil price explosion, but it was none the less part and parcel of the general drift into stagflation. Higher oil and commodity prices added directly to price-inflationary pressures in industrial countries, while the increased savings of Arab oil sheiks produced demand deflationary effects. Cost-push inflationary pressures were thus increased on both sides of the account.

VI THE RESPONSE TO STAGFLATION

Governments in industrial countries were slow to react to the first oil price explosion. The result was the inflationary slump of 1974–5. When they did react they tackled demand-deflation with fiscal and monetary stimulation and price-inflation with prices and incomes policies (see Figure 5.13). The former worked while the latter failed. As a result, the first oil price increases were validated by the generalised rise in world prices. In real terms, oil prices declined from 1974 to 1979, with the result that by the late 1970s demand once more recovered. In this period international credit was again sharply expanded, this time through petro-dollar recycling. OPEC surpluses were originally earned at the expense of the main OECD industrial countries. OPEC savings were largely deposited with industrial countries' banks. These savings were then lent to developing countries in Latin America and Africa and to communist Eastern European economies. These poorer and developing countries were thereby encouraged to continue to buy from the OECD even when,

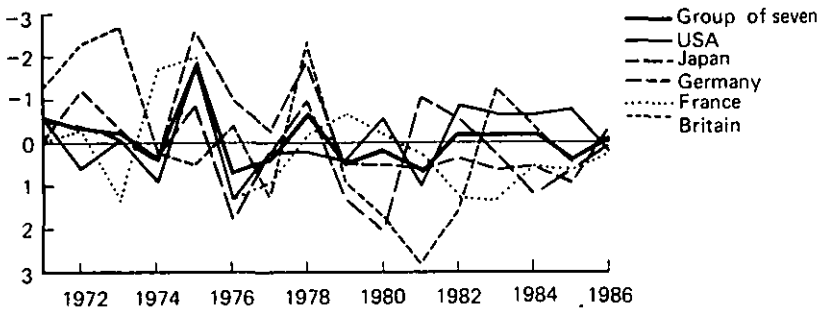

Figure 5.13 Discretionary fiscal policy. Policy changes to budget deficits, per cent of nominal GDP. Negative is reflationary

through slow growth, OECD imports of food and raw materials were falling. The counterpart to the OPEC surplus was thereby recycled from OECD economies to the LDCs. To make matters worse for these debtor countries, a part of the money borrowed was ploughed into expanding commodity production in response to the 1972–3 shortages and price hikes. Meanwhile, within the OECD, governments in Europe and Japan were persuaded during 1977–8 to run increased budget deficits in order to share the locomotive role for the world economy with the USA.

The response to the first oil price explosion generated the second. Over the turn of 1979–80 oil prices again exploded upwards. Once more the initial result was an inflationary slump. But this time the major industrial countries reacted in a different way. OECD overall fiscal policy remained neutral. The rise in the US budget deficit allowed Britain, Germany and Japan (later followed by France) relatively painlessly to reduce their public sector deficits. The deflationary effect on home demand was mitigated by increased exports to the USA as that country slid into deep trade deficit. Monetary policy became restrictive and interest rates rose. Outside the USA, slow growth and rising unemployment brought inflation rapidly under control. Internationally the rise in oil and commodity prices was negated instead of being validated. Supply outran demand and in 1986 oil prices imploded, falling from $30 a barrel to briefly under $10.

US policy in both periods played a distinctive and crucial role. As a major producer and minor importer in the early 1970s, the USA tried to insulate its economy from the effects of the first oil price

explosion. The domestic price of oil was controlled. In consequence US production continued to stagnate and US consumption steadily rose. This was one reason why world demand failed to respond fully to the first hike in oil prices. But following the second oil price explosion the USA moved domestic oil prices to the level of world oil prices. Unlike other economies, its inflation therefore peaked in the early 1980s rather than the mid-1970s. At the same time, the USA performed most powerfully the role of locomotive economy in the 1980s. Supply-side tax cuts led to both the big US budget and payments deficits, without which the world would probably have slid into a slump. The rising US trade deficit also saved the world from a more acute LDC debt crisis. Fortunately for debtor nations, when US banks stopped lending abroad in early 1982, precipitating the debt crisis, US consumers started spending abroad, making it possible for the crisis to be managed without any disaster for the world financial system. (But by no means has the problem been solved.)

Through the 1980s the forces which caused stagflation have been steadily reduced. High unemployment (outside the USA) has substantially reduced unions' aspirations for real income growth. In the USA rapid labour force growth has kept labour plentiful despite relatively lower unemployment levels. Wage-indexation systems have generally been scrapped, particularly in Italy and France. Job preservation has become the unions' major objective. Governments, meanwhile, have struggled with bludgeoning budget deficits. The trend towards increased public spending as a share of national output has been painfully halted and is now being reversed. In the meantime, money illusions have all been destroyed by the inflation explosion. Governments (particularly in Britain), were forced to index personal tax allowances and tax bands, meaning that public sector revenues no longer painlessly increase as a percentage of incomes. Savers are no longer cheated by negative real interest rates. Instead rates have been positive and historically high. Any suggestion that governments might be embarking on inflationary policies now results in immediate market pressure for higher nominal interest rates. In consequence, any tendency towards resurgent inflation is immediately countered by incipient recession. The cost-inflation of the 1970s has been burnt out by the demand deflation of the 1980s.

VII INTO A DEFLATIONARY DECADE

As cost-inflationary forces have moderated during the past decade, demand-deflationary forces have gained the upper hand. Only a further massive growth in credit has saved the world from sliding into a new depression. But the growth in this credit, and the movement of funds between countries which it has spawned, have created new problems of instability in foreign exchange and domestic capital markets. Imbalances have been generated within and between economies and their elimination is now the most pressing problem facing world fiscal and monetary authorities.

Deflationary pressures stem from technological and demographic change. On the one hand, the technological revolution has depressed the propensity to invest, while on the other, demographic change is increasing the propensity to save. The problem with the present technological revolution is that it has produced relatively few new products; the video recorder and the microwave oven are notable exceptions. Instead it has revolutionised the way we make the same old products. In consequence, our homes, high streets and countryside have been little changed. There is nothing akin to the transformation which railways, cars, jet planes successively brought in their wake, or to the consumer durable revolution which transformed our homes during the great post-war prosperity. We are merely able to produce mostly the same products massively more efficiently, using less labour, less materials, less energy and less capital in the process. Nor has the present technological revolution spun off increased investment demand in other areas. The lack of new products has meant that there is nothing now to compare with the highway programmes, the airports, the huge investments in new sources of raw materials and energy which fuelled demand during the 1950s and 1960s. Technical change has transformed one area of activity, financial services. But even here the effects have been mixed. The enhanced ability of individuals and intermediaries to move money cheaply, quickly and knowledgeably between capital markets and countries has created such exchange rate and interest rate instability that greater risk premia are now required before corporations are willing to invest in expanding capacity. Not only *ex ante*, but *ex post*, the share of investment in the output of industrial countries has declined in recent years (see Figure 5.14). One of the dominant characteristics of the long and slow recovery from the 1980–1 recession has been that it has remained largely consumption-led. Only recently have signs

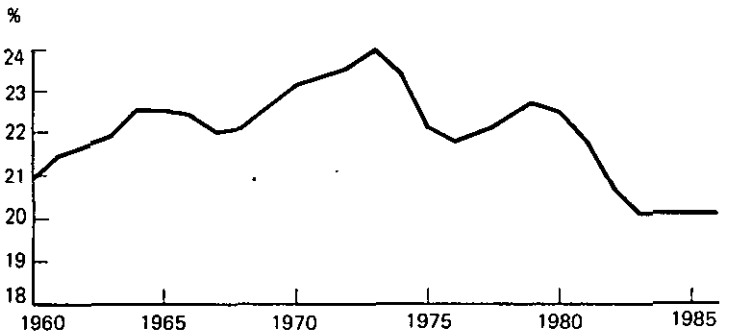

Figure 5.14 Share of investment in OECD GDP

appeared that the present recovery could mature with an acceleration in capital investment.

Increased uncertainty about the future, coupled with demographic change, has enhanced the propensity to save. The population in all advanced industrial countries is now ageing. The post-war baby boomers are reaching middle age without they themselves having so many sons and daughters. The timing of this demographic phenomenon varies between countries. Germany has been a lap ahead and its total population is already falling. The USA has been a lap behind. Its working population is still rising. But everywhere the effect of these demographic changes on savings is dramatic. There is essentially a lifetime pattern to saving and lending. When we are young we borrow and spend, buying and furnishing homes, raising children. When we are old and retired we draw down past savings and spend. In middle age, however, we save to repay past loans and to accumulate the capital on which to live when we are old. So as the labour force grows older, savings increase. Nowhere is this more apparent than in the finances of pension funds and social security systems. In Japan, for instance, the social security system has been generating annual surpluses equal to 3 per cent of GNP, which have financed the greater part of the deficit that the government has been running on its general accounts budget. By the middle of the 1990s these surpluses are set to disappear. Not only will more people be retiring but each will have a larger pension to look forward to, having paid into the scheme for more years since it was started. In the USA, a rising surplus on social security funds is about to transform the savings propensity of that economy. The result will be that the USA, by the mid-1990s, will be running a large and rising public sector surplus. The Federal

government budget deficit will not be a matter of concern for much longer, while the US payments deficit could soon become an embarrassing surplus.

When people want to save and lend on an increased scale, others must be persuaded to borrow and spend, or economic growth comes to a halt. This has been the story of recent years. Only the continued injection of additional credit into the international economy has prevented the world from sliding into a depression. But except where credit is put to productive use (and this has been on an inadequate scale due to the decline in the propensity to invest), debts rise relative to income. This ultimately undermines the credit-worthiness of any borrower. When borrowers can no longer borrow, they must stop spending more than they earn and start earning more than they spend. From being a stimulus to growth, they become a drag upon it. So ultimately the build-up of debt sets a limit to the further extension of credit. At this stage some debt shake-out must follow, or some other forces intervene to eradicate the excess propensity to save. Without such a development, a financial crash and consequent international depression become inevitable.

So far a succession of borrowers and spenders have saved the world from this catastrophe. These have included:

(1) Lesser Developed Countries. The manner in which petro-dollar recycling produced the LDC debt problem has already been described. In all probability this problem, caused when the USA stopped lending abroad in 1982, will be resolved by widespread debtor default in 1988–9 as the US stops spending abroad.

(2) Texas oilmen, Californian farmers, condominium buyers, smoke-stack industries. Excessive US bank lending was not confined to LDCs and all the above categories of borrowers and spenders went to the limit of their credit in recent years, and well beyond it. Many of these problems are already being resolved by debt default.

(3) Governments in Britain, Europe and Japan. All ran big budget deficits in the late 1970s and early 1980s, as already described, and all were left with bigger debts which now limits their room to reflate. The Japanese government is particularly worried by the build-up of its debts to its own social security system.

(4) The US government and consumer, the biggest and best borrowers and spenders ever. Their contribution to saving the world

from recession is enshrined in the twin US budget and balance of payments deficits.

(5) Stock market speculators, leveraged buyers-out, take-over merchants, stock buyers-in, real estate and property speculators, collectors of collectibles everywhere, and old master buyers. The ultimate stage of credit creation is where the money goes into asset price inflation generating continued demand only through wealth effects.

All the evidence is that the supply of credit-worthy borrowers and spenders has now almost been exhausted. By early 1987 foreign private investors were no longer willing, at the existing dollar exchange rate and interest rates, to finance the continuing US external payments deficit. Had nothing been done, the USA would have been forced by a foreign exchange market crisis to stop borrowing and spending so much. This would have precipitated the feared world recession. In the event foreign central banks, following the February 1987 Louvre Accord, stepped in to pick up the tab. Official foreign exchange intervention in 1987 totalled somewhere between $100 billion and $150 billion. The consequences for Britain from its $20bn intervention have already been described. The consequences for the world were similar.

Central bank intervention caused the governments of all major countries outside the USA to lose control of the growth of credit during 1987. In the United States, the time gained by central bank intervention was used to neglect rather than correct fundamental imbalances. In Britain, America and Japan, the result was speculative excesses. These produced the booming stock market which preceded the October crash. In their attempts to stabilise foreign exchange markets, governments successfully destabilised domestic money and bond markets. When they became alarmed, from mid-summer, with the speculative excesses which resulted, they moved to sterilise the effects of intervention. This involved raising interest rates causing bond prices to tumble, which in turn destabilised the stock markets leading to the crash.

The world economy is now balanced on a knife-edge. The massive expansion of credit, and the asset price inflation which it induced, did lead to a recovery in consumer spending and construction activity in Europe and Japan. Growth from the middle of 1987 was everywhere much faster than forecasters expected it to be. In the USA, the downward adjustment of the payments deficit now seems under way.

US output remains buoyant even if US consumer demand is flagging. Expansionary forces outside the USA may now be sufficiently strong to prevent the correction of the US twin deficits from causing a severe world recession.

A more pertinent issue is whether the deflationary decade is now drawing to its close. The depressed level of expenditure on new capacity during the long, slow-burn recovery has created the need for a sharp investment upswing, which is already occurring in the more dynamic economies, i.e. in the USA and Britain. Secondly, contagious tax reform is producing less penal and more efficient systems. Not only are they more neutral between savings and investment, but also they are generally more efficient in collecting a maximum of revenue with the minimum of rates, instead of the other way round. This gives governments higher revenues with less pain and greater scope to increase public investment in sadly neglected social infrastructure. Tax reform is producing greater income inequality. While this may tend to increase the propensity to save, it also changes the composition of expenditure between mass produced goods, crafted goods, skilled work and services. Spending by those with high incomes is skewed towards areas least affected by the technological revolution, where the labour content is greatest. Each pound or dollar of expenditure directly provides more jobs and more income for those less able today to find traditional industrial employment.

But possibly the most crucial factor is what now happens to interest rates. The one anomaly of the deflationary decade is that real interest rates have remained so high for so long. At first sight, this makes it hard to sustain the argument that there has been a surfeit of savers chasing a dearth of productive investors. But there are several explanations for persistently high real interest rates. Fears of resurgent inflation have irrationally persisted following the excesses of the 1970s and early 1980s. Increased volatility in financial markets has created uncertainty, against which higher risk premia are required. Problems of larger and more intractable domestic and international disequilibria have placed greater strains on capital markets. Higher interest rates are required to move funds on the necessary scale to where they are needed. Interest rates measure the return on specific financial instruments and not the credit status of those who borrow by way of those instruments. Bad credit risks have been able to borrow on a scale not formerly possible. The quality of credit has declined as the

quantity of credit has increased. Higher interest rates reflect this deterioration in quality.

Money has thus been plentiful but dear, another pathological development. High real interest rates have been associated with rapid money stock growth and a falling velocity of circulation – totally the

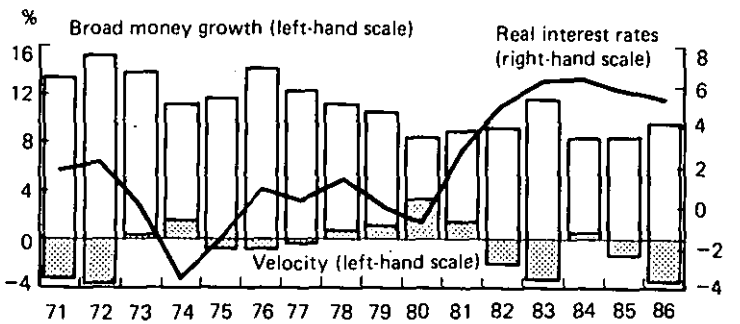

Figure 5.15 Industrial countries' interest rates and velocity

opposite of what some monetarists would expect (see Figure 5.15). With the stock market crash, the importance of risk relative to reward has been emphasised. In consequence, more attention will now be paid to the quality of loans and less to the interest they notionally yield. This will make money scarcer but cheaper. Scarcer credit will undermine asset price inflation (which is the last gasp of the old inflationary age, not the harbinger of a new one). Lower interest rates will mean greater demand to finance productive investment, which enhances non-inflationary real output growth. Plentiful and dear credit was associated with asset price inflation, product price deflation and decreased velocity of circulation, and a more adverse division between inflation and real output growth. In a nutshell, too much money chasing too many goods. Cheaper and scarcer credit will be associated with higher investment in productive assets, faster growth, less inflation, and a rising velocity of circulation. In other words, we could be about to return to a more stable and more prosperous age. Let us hope so.

6 Structural Change in the British Financial System

David T. Llewellyn

A different perspective is taken in this chapter. While the other contributions consider various aspects of 'money' and credit, the focus here is on two related aspects of the British financial system: the major changes in its structure that have been a feature of the 1980s, and the acceleration in the pace of financial innovation. Both are pertinent to the theme of this volume, and after considering their nature and implications, their relevance for money and credit are highlighted. The focus is upon financial innovation and the nature of structural change in the British financial system as a whole rather than a detailed discussion of particular subsectors such as building societies, the stock exchange, etc. In particular three aspects of the subject are considered: the precise nature of structural change and financial innovation, the forces producing these changes, and their implications.

1 FINANCIAL INNOVATION AND STRUCTURAL CHANGE

The basic themes may be summarised as follows:

(1) After decades when the structure was ordered and stable, the British financial system has, over the 1970s but most especially in the 1980s, undergone a degree of structural change greater than ever previously experienced. It is the combination of the extent of changes and the range of institutions affected that makes the recent phase of evolution unique in historical terms.

(2) There is a close parallel between the process and causes of this innovation and change: both have reduced rigidities and demarcations between institutions and instruments, and both

have been induced by the simultaneous pressures of competition, technology, changes in regulation, and the specifically international dimension to competition. Structural change and financial innovation are two parts of the same process.

(3) Structural change is not a random process, but can be analysed by identifying systematic pressures undermining traditional structures. These pressures are not fundamentally different from those that produce structural change in other industries, though the impact in finance has unique features.

(4) Competition has come to have an increasingly powerful international dimension. National financial systems are effectively in competition with each other as historic 'market' and 'imposed' advantages of domestic financial mechanisms have been steadily undermined, again by the simultaneous pressures of competition, technology, changes in regulation, and financial innovation.

(5) A major feature of all subsectors of the financial system has been the trend towards diversification. In the process the financial system has become less structured and more integrated as the historic divisions between subsectors have been eroded. This has changed the traditional structure of a system based upon specialist financial institutions. In particular, the historic separation (also common in many other countries) between commercial banking, merchant or investment banking, housing finance, insurance, securities trading and portfolio management has been eroded with the emergence of financial conglomerates.

(6) The structural changes have implications for the interpretation of the substantial growth of credit in the 1980s and the simultaneous rise in personal sector debt and financial assets.

(7) There is a strong presumption that financial innovation and structural change have increased *structural* and *allocative* efficiency though this is not to say that the benefits are unqualified.

Historical Structure

For most of this century the structure of the British financial system was remarkably stable; historically it has been a highly structured system based on specialist and differentiated financial institutions and markets with clear functional demarcations between them (Revell, 1973). Notably, and in common with many other financial systems (especially those in the USA, Canada, Japan and Australia) but in

contrast with many in Europe, there has been a clear functional distinction between the major business areas of commercial banking, investment banking, insurance, fund management and securities trading. However, in contrast to the experience in other such systems, this has not been predominantly a product of official regulation, but more a function of restrictive practices, anti-competitive mechanisms, and self-imposed constraints.

On the other hand, the regulatory authorities (and the Bank of England in particular) have been sympathetic to this structured approach and have supported limits on the range of business operations of different subsectors of the financial system. It was viewed as one way of dealing with potential conflicts of interest that can arise in institutions conducting a wide range of business. It also made regulation easier since if the objective was to regulate *functions* there was a clearly defined set of *institutions* performing each function. It also meant that to a large extent regulation could be 'sub-contracted' to the sectors themselves. Such self-regulation through 'clubs' has been a distinctive feature of financial regulation in the UK. The authorities also took the view (which can be questioned) that risks are reduced through limiting the range of business activities, and that systemic risk is lowered by separation – reducing the danger of contamination of one part of the business (say banking) by risks in other parts (e.g. insurance).

At the same time, historically there have been powerful pressures in the British financial system limiting competition. Regulation has often been anti-competitive in nature based upon the alleged dangers of 'excessive competition'. Official attitudes to competition have been different in finance than in other industries, with a strong strand in the history of regulation based upon the alleged dangers of excessive competition inducing both immoderate risk-taking by institutions and resultant systemic dangers (Llewellyn, 1986). Restrictions on the range of activities can frequently be anti-competitive (as protection is given against competition developing from outside the regulated group) and even perverse, as they may have the effect of increasing risk by denying the benefits of a diversified business structure. There is a trade-off between the objectives of regulation and the benefits to be derived from competition and the chosen mix varies over time (Bain, 1981; Llewellyn, 1987). It is also the case that for decades monetary policy was operated in a way (e.g. through direct credit controls) that had the unintended effect of limiting competition in banking.

However, in practice, regulation has not been the most powerful force limiting competition in finance. There have been many self-imposed restrictive practices and anti-competitive mechanisms which have fixed prices and limited the range of business activities of different subsectors. The structured nature of the British financial system, has been as much a product of these self-imposed restraints and restrictive practices as official regulation. Nevertheless, such mechanisms would not have survived had they not been condoned, if not positively encouraged, by the regulatory authorities. They were condoned partly because they were viewed as an alternative to explicit regulation and as performing the same role as regulation in limiting risk by limiting competition.

II STRUCTURAL CHANGE IN THE FINANCIAL SYSTEM

In this historical context, the major structural changes that have dominated the recent evolution of the British financial system relate to competition, the process of diversification, the undermining of the historically structured system, the form and extent of regulation, the increasing internationalisation of finance, and the acceleration in the pace of financial innovation. Overwhelming all other forces has been the role of competition as the competitive environment has intensified within and between subsectors of the system. In the process, the specialist basis of financial institutions has been powerfully eroded as diversification has produced a trend towards financial conglomerates. This has been a comparatively easy process as historical demarcations have been largely self-imposed rather than forced by legislation. At the same time, regulation has become more formalised and explicit, and competitive pressures have undermined most of the traditional features of self-imposed regulation based upon restrictive practices.

The dominant change, and one that has been both created by competitive pressures and in itself reinforcing them, is the process of diversification and the erosion of the historically structured basis of the financial system (Llewellyn, 1985). This has been occurring across the whole range of financial institutions and markets. The clearing banks have bought securities trading firms and merchant banks, and they now own unit trusts (thus merging four areas of finance that have traditionally been kept separate). One bank has also bought an

insurance company. At the same time, the process of diversification into a limited range of financial services includes non-finance companies such as the BAT company and Marks & Spencer (which holds a banking licence). Marks & Spencer launched its own unit trust in October 1988 which it sells through a directly owned subsidiary. Several national retail stores (with access to a large number of personal customers) offer credit card services and personal loans. Thus conglomeration has come to straddle financial and non-financial companies.

There have been major and far-reaching changes in the securities industry (Goodhart, 1987). Historically, this sector has been highly regulated through practices enforced by the London Stock Exchange. In particular: (i) a strict division was enforced between market-makers (jobbers) and brokers; (ii) brokers were required to deal exclusively with jobbers rather than matching customer deals between themselves; (iii) minimum commission charges were established so as to remove the incentive to bypass the jobbers which, it was alleged, would have undermined the market-making role to the eventual detriment of all users of the stock exchange; and (iv) the extent of external ownership (e.g. by banks) was substantially limited. This meant that brokers were undercapitalised for trading in securities rather than performing an exclusively brokerage role.

As a result of deregulation in the stock exchange and the securities industry, British and foreign banks have become major elements in the securities industry as part of integrated financial groups comprising banking, merchant banking, securities broking and market-making, together with fund management components. In October 1986, fixed commissions for securities trading in the London market were ended following the agreement between the government and the stock exchange. In March of the same year, stock exchange rules were changed to allow 100 per cent outside ownership of stock broking and jobbing firms. The abolition of minimum commissions would inevitably undermine single capacity, but the industry as previously constituted was undercapitalised for dual capacity. This required a change in the rules to allow a capital injection via banks buying securities firms. The resulting capital injection by a wide range of British and foreign banks has changed the UK market in two major respects: the hitherto separate and specialist role of market-maker and broker (enforced by internal regulation since 1908) has been abandoned, and all the major securities firms have become parts of integrated financial institutions.

Building societies are similarly diversifying. From being a set of highly specialised institutions, whose business was exclusively to collect retail savings deposits to finance mortgage loans, they are emerging as diversified retail banks encompassing consumer lending, insurance underwriting, stock exchange brokerage services, portfolio management and money transmission services which have hitherto been the monopoly of a small number of clearing banks.

Factors Producing Structural Change

The substantial changes in the British financial system evident in the 1980s are a product of several factors operating simultaneously; these are summarised in Table 6.1 and discussed in this section.

Table 6.1 Factors inducing structural change

(1)	Competition
(2)	Changes in regulation
(3)	Technology
(4)	Changes in the market environment
(5)	Changes in the business constraints of the suppliers of financial services
(6)	Changes in the relative efficiency of competing groups
(7)	The strategies of suppliers of financial services
(8)	The dynamics of financial innovation
(9)	International pressures

1 *Competition*

The competitive environment has intensified greatly in the 1980s, not least as subsectors have sought to diversify the range of business operations in several dimensions: within subsectors (e.g. building societies), between subsectors (banks and building societies are in competition with each other across a wider range of business areas as each has diversified into the others traditional territory), between banks and the capital market, and also internationally. (The role and impact of competition are discussed in more detail in Llewellyn, 1989.)

2 *Regulation*

Regulation has a major potential impact on the financial system, and has in its turn experienced far-reaching changes in the 1980s. 'Excessive' regulation imposes a tax on financial institutions which, in the final analysis, will be reflected in prices. To the extent that any of the various forms of regulation (legislative, self-imposed, etc.) prescribes the allowable range of business, this necessarily affects the structure of the financial system. There will also be significant structural effects if regulation is not competitively neutral as between competing sectors, since it imposes a competitive 'tax' or 'subsidy' on one set of institutions compared with another which has implications for relative efficiency and the supply price of financial intermediation.

The style of regulation in the UK has changed substantially since 1980. The financial system has become more complex and the range and number of institutions has increased substantially. There has also been a major influx of foreign institutions not steeped in the City's traditions and frequently they have come from countries where regulation is highly legalistic. In this environment the traditional, flexible approach has become less viable.

The regulatory environment has been changing substantially over the 1980s in several respects:

— competitive pressures have undermined some traditional features of self-imposed regulation;
— regulation has become more formalised and explicit;
— while institutions have broadened their range of financial services, the way each sector of business is conducted has become more explicity regulated;
— a wider range of financial services has been brought within the regulatory regime through a series of self-regulatory organisations (SROs) under the general authority of the Securities and Investment Board;
— the number of regulatory or supervisory bodies has increased;
— in the context where distinctions between types of financial institutions have become increasingly blurred, regulation and supervision have increasingly come to be based upon functional rather than institutional criteria;
— regulation in general has become more interventionist and detailed.

What has emerged is that two trends in regulation are operating in opposite directions: various forms of deregulation have enabled financial institutions to widen their range of services, while at the same time the authorities have imposed a more stringent regulatory environment relative to the conduct of each of their particular areas of activity. With respect to the former, the current regulatory environment is less constraining now than ever before and in this dimension the financial system is less confined than virtually any other in the world.

In order to place the impact of regulation in perspective, it is generally the case that in the evolution of the British financial system the general pressures of competition have been considerably more powerful than specific episodes of legal deregulation; the less formal system of regulation has in practice been broadly accommodating to changes in the competitive environment, although particular episodes of formal deregulation have been powerful and have acted as catalysts to underlying competitive pressures. Throughout the 1970s and 1980s, the process of deregulation has been largely a response to competitive pressures and financial innovation rather than dramatic policy changes designed to force the pace of change, competition and innovation. *Competition and Credit Control*, and the pressure on the clearing banks cartel, were a product partly of the competitive thrust introduced by the entry of foreign banks, and the abolition of direct controls on lending was induced by a steady process of circumvention. The abandoning of the building societies' interest rate cartel in 1983 followed increased competition in the hitherto virtual monopoly position of building societies in the mortgage market. The changes in the securities industry were motivated not so much by the threat of reference of the Stock Exchange rule book to the Restrictive Trade Practices Court, but by the force of international competition in the securities industry. The more recent diversification allowed by building societies developed largely because, in the new competitive environment, they were placed at a competitive disadvantage in being substantially restricted in the range of allowable business. Thus deregulation has been as much a product of competitive pressures as a factor creating a more competitive environment in the financial system (Bank of England, 1983). As a result the regulatory authorities have moved in the direction of supporting, even encouraging, the trend away from historical specialisation by different types of financial institution.

3 *Technology*

The development of technology has also had a major impact on the financial system as it has the potential to increase the efficiency in the provision of existing services; lower the entry costs to some services; enable new services to be provided; lower transactions costs in financial markets, and undermine the role of branches as a major delivery system.

In general, recent and prospective developments in technology reduce the cost of entry to the payments system. This implies that institutions other than banks (building societies are an obvious example) are able to challenge the banks' one major near monopoly – the payments system. To the extent that non-cash payments can be transacted and settled without the expensive paper-based cheque system, institutions can enter without being members of the traditional clearing system.

Secondly, new technology challenges the role of branches both of banks and other retail institutions. This in itself lowers the cost of entry into 'banking' services. Technology can displace many of the standard functions performed by branches. One of the implications of technology is to remove the necessity in some financial services for the supplier of the service to be near to the customer. In this sense it challenges the comparative advantage of institutions with branch networks. This also has an international dimension as it enables foreign suppliers of financial services to compete with domestic institutions. A particular example is the development of securities trading technology which enables investors to trade in securities via dealers and brokers located abroad. This was a major factor in the structural changes in the UK securities industry as the London Stock Exchange came to be increasingly bypassed as UK institutions could easily trade via New York.

4 *The Market Environment*

Changes in the market environment have a major impact on financial systems and the way they perform the basic functions because they affect both the suppliers and demanders of financial services. Obvious examples of this impact were the effect of the substantial rise in domestic and international imbalances in the 1970s, and the high and volatile rates of inflation and interest rates.

The high and volatile rate of inflation throughout the 1970s had a

marked impact on financial institutions' balance sheet positions, changed the structure of corporate financing, induced companies to shift from fixed interest (capital market) sources of funds towards floating rate (bank) sources, induced a higher level of personal sector savings, undermined some pension fund portfolios and necessitated 'topping up' operations; and increased the level and volatility of interest rates. Financial institutions and their clients respond to general changes in the economic and financial environment. There are legion examples of financial innovation in response to the increased volatility of interest rates, exchange rates and the rate of inflation, and the high level of interest rates.

5 *Constraints*

On the other hand, institutions and the suppliers of financial services may face constraints (perhaps due to regulation) in meeting demand. If a constraint is eased this may be reflected in the structure of the financial system. An obvious example is with respect to capital. During the 1970s the substantial growth of the banking sector was facilitated in part by a decline in capital–asset ratios. During the 1980s, partly under regulatory pressure, banks have raised capital–asset ratios. But the potentially constraining effect of this was offset by a regulatory decision of the Bank of England to allow banks to augment capital through issues of various forms of debt instruments. Banks have made substantial use of this and in the process capital was augmented by a total of £6.5 billion for the four clearing banks in the period 1980–7, and loan capital rose to 35 per cent of the capital base. In the early and mid-1980s, the trend towards 'securitisation' and off-balance sheet business (most especially in international finance) was associated in part with the emergence of capital and other balance sheet constraints, which created incentives to develop off-balance sheet business and generate fee income to raise the rate of return on assets and equity.

Changes in regulatory constraints (whether externally or internally induced) can have a major impact on the financial system as seen in, for instance, the building society sector which, since the 1986 Building Societies Act, has been freed of many of the restrictions previously imposed upon its business operations. Similarly, changes in the rule-book of the Stock Exchange has removed constraints over dual capacity and the external ownership of securities firms.

6 *Relative Efficiency*

Structural change can also occur as a result of changes in the relative 'efficiency' of different subsectors of the financial system. During the 1970s, for instance, bank lending became more flexible with the extension of Term Loans, new types of lending, and new pricing formulae. In general, banks became more innovative in funding strategies and developed new techniques of liability management to a high degree of sophistication.

A major example in the 1980s has been with respect to the structure of the mortgage market. Building societies have traditionally funded in the retail savings market where interest rates have historically been lower than in wholesale markets. In the early 1980s, the mortgage rate and the societies' maximum retail deposit rate moved above wholesale money market rates. This made mortgages an attractive asset to hold for a new set of institutions (such as the National Home Loans Corporation) created specifically in response to this change in the structure of interest rates, and funded exclusively in the wholesale markets. They had a competitive and efficiency advantage over building societies which, due to regulatory limits on the extent of their wholesale market borrowing, were not able to take advantage of lower funding costs in these markets. Thus a change in relative 'efficiency' had a significant impact in a short period on the structure of the mortgage market, as it created a competitive advantage in favour of banks (who have no regulatory limits on the structure of funding) and new institutions funded in wholesale markets. In 1983, the share of new mortgages of these new institutions was less than 2 per cent but this had risen to close on 15 per cent in the final quarter of 1987.

7 *Business Strategies*

Financial institutions are not passive but have their own portfolio objectives which change over time. Changes in the structure of the financial system can occur as a result of changes in financial institutions' own strategies. As an obvious example, throughout the 1970s the portfolio behaviour of banks changed towards asset-growth objectives, internationalisation of portfolios, active and innovative liability management and the objective of increasing the range of services. More recently, the emergence of capital and portfolio constraints has increased the incentive towards off-balance sheet

activity designed to generate fee income without increasing assets.

In the 1980s, changes in business strategies have continued to have an impact on the structure of the financial system. The overall strategy of diversification and the offering of a wide range of financial services by all financial institutions have eroded the traditional structure of the financial system comprised of essentially specialist institutions. The strategic decision of banks in the early 1980s to develop retail financial services on a larger scale has also had a significant impact. In the mid-1980s, and in the context of potential capital constraints, the strategic decision of building societies to focus more upon basic profitability rather than asset growth had a notable effect upon their business performance.

8 *Dynamics of Financial Innovation*

A major feature of the 1980s has been the sharp acceleration in the pace and range of financial innovation, and the emergence of several secondary markets in which new instruments are traded. During the 1980s new financial instruments, markets, facilities and techniques have emerged on a large scale most especially in the British and American financial systems, and in international financial markets. In most developed financial systems, there has been a substantial growth in the range and issue volume of short-term financial instruments, together with a much higher volume of secondary market transactions.

Financial innovation has two central features: (1) the creation of new financial instruments, techniques and markets, and (2) the unbundling of the separate characteristics and risks of individual instruments and their reassembling in different combinations. In the process three central features of financial innovation are important: (i) it increases the range, number and variety of financial instruments; (ii) it combines characteristics in a more varied way and widens the combination of characteristics thereby reducing the number and size of discontinuities in the spectrum of financial instruments; and (iii) it has the effect of eroding some of the differences between different forms of intermediation as, for example, where a Floating Rate Note is a capital market instrument similar, except for liquidity, to a floating rate bank loan.

The development of technology has had a major impact on financial innovation, not least because it enables traders in new instruments

(options being a good example) to immediately calculate arbitrage opportunities in complex situations. Technology also contributes to the design and pricing of new instruments, and facilitates the identification, measurement and monitoring of risks in portfolios containing complex instruments. It also reduces trading costs in international markets, and has the effect of widening the market for new instruments to an international dimension. In the process, the supply and demand for instruments in isolated markets can be more readily matched. It also enables international markets to be linked and therefore gives access to a wider range of borrowers; an obvious example is swaps.

There are several reasons why the pace of financial innovation accelerated in the 1980s. The dominant factor has been the role of competition. The competitive environment in all financial systems, and in international markets, intensified greatly in the 1980s. A second major factor has been the parallel development of information and trading technology. Changes in the economic and financial environment may have created a demand for new instruments, but the development of technology has enabled these demands to be met. A third factor has been the volatile financial environment especially with respect to interest rates and exchange rates. Many of the new instruments have been designed to offer some form of protection against this volatility. Fourthly, and as part of the more competitive market environment, a 'marketing ethos' has developed in the management of financial institutions. The creation of new instruments and facilities partly represents an attempt by the suppliers of financial services to capture market share and create demand.

A further characteristic feature of the 1980s has been a steady 'internationalisation' in finance. This is discussed in detail in Llewellyn (1988), and can be counted as a fifth major factor in the acceleration in the pace of financial innovation. Many of the new instruments and facilities that have been created were developed initially in international markets (partly because of the special risk characteristics of international financial intermediation and the more intensive competitive environment) and as a response to regulation in national markets. They have subsequently been adopted in national financial systems.

The analysis of financial innovation may be summarised by considering in general terms the benefits that accrue through innovation in financial systems. These may be briefly summarised as they are implicit in much of the previous discussion:

(1) The costs of financial intermediation are reduced in two major ways: (i) they give access to borrowers to a wider range of markets and facilities, and (ii) in some cases they allow different institutions to exploit their comparative advantages. In the process market imperfections are eroded.
(2) New instruments facilitate arbitrage between markets in different countries and instruments, and in principle this erodes pricing anomalies and may reduce market imperfections.
(3) Some instruments and techniques also allow individual borrowers to exploit their comparative advantage in different markets which again lowers the costs of financing; swaps are a good example.
(4) Several instruments (such as futures, options, etc.) widen the range of hedging possibilities and hence enable risks to be protected against.
(5) As part of the same process some instruments allow risks to be priced (one of the basic functions of a financial system) and to be shifted to those able and willing to absorb them.
(6) To the extent that innovations create secondary markets, they have the effect of increasing liquidity and facilitate the management and adjustment of portfolios. In the same process they also serve to price risks.
(7) Many instruments allow various risks to be unbundled, separately priced and 'sold'. If the risks are correctly priced this contributes to the financial system's role of allocating resources efficiently.
(8) By increasing the range of financial instruments the process of 'spectrum filling' offers a wider range of choice with the presumption that the requirements of users are more readily and efficiently met. In the process it reduces the number and extent of discontinuities in the range of facilities.

Financial innovation has become a major feature of financial systems. It has contributed, in some countries more than in others, to major structural change in national financial systems and to a general integration of such systems. The process of financial innovation contributes to four specific functions of a financial system: (i) it increases the variety of financial instruments for both borrowers and lenders by offering various combinations of standard characteristics; (ii) it increases the liquidity of the system to the extent that tradeable assets and markets are created; (iii) it enables particular

risks to be priced and transferred; and (iv) it may give borrowers access to particular markets from which they were hitherto excluded.

In the process financial innovation has had the effect of increasing the integration of markets (including between financial centres) as many instruments straddle different types of markets and those in different countries. As part of this the formal distinction between bank and capital market intermediation has tended to become less powerful. In this sense, financial innovation has contributed to a more general structural change implying a less structured, more integrated system with the specialist nature of different types of financial institutions becoming less pronounced. Just as some instruments have the effect of integrating and linking *markets*, so they may also have the effect of eroding specialist functions of financial institutions. Both are a reflection of the 'spectrum filling' role of innovation which is a natural product of a competitive financial system.

A general theme of this chapter has been that structural change in the financial system has involved the eroding of demarcations between different subsectors. There is a close parallel with the process of financial innovation in two respects: (i) it implies the same feature of mixing characteristics of different instruments and in the process reducing the distinctions between them, and (ii) it has been induced by the same pressures of competition, technology, changes in regulation and the specifically international dimension to competition. In essence, structural change and financial innovation can be viewed as two components of the same process.

9 *International Pressures*

A final force inducing structural change has been the emergence of powerful international pressures (Llewellyn, 1988; BIS, 1986). Although the bulk of world financial intermediation is conducted through national mechanisms, finance has become increasingly global. National financial systems are losing some of their traditional 'efficiency' (e.g. transactions costs, access to customers, information) and 'imposed' (e.g. exchange control, etc.) competitive advantages. In practice, this amounts to domestic institutions and markets being subject to competitive pressures from foreign financial systems.

In varying degrees the basic roles of a financial system can be performed either domestically (within a national financial system) or externally. Historically, domestic mechanisms have been dominant

though the factors accounting for this are being steadily undermined. The central functions of a financial system can increasingly be performed by external systems instead of exclusively within domestic systems. Which system or set of markets and institutions are used depends upon the relative efficiency of the alternatives, and the ability of users to arbitrage between them.

The international dimension of financial intermediation emerges at three levels: (i) neither surplus nor deficit units are restricted to mechanisms (institutions and markets) located within their own country; (ii) the suppliers of financial services are not restricted to business within their own country; and (iii) financial institutions have the option of locating outside their home country. In all three respects finance has become increasingly internationalised. The process of *globalisation* is emerging in several dimensions. The most obvious is in securities trading where shares and other national capital market instruments are now traded in several national markets. A second dimension is location. There has been a growing trend for financial institutions to enter foreign markets not only to conduct international business but also to develop local domestic business. Thirdly, cross-border business has become more important for financial institutions. Borrowers are able to arbitrage and increasingly have international financing options. This is reflected in the progressive use of international as opposed to domestic funding instruments. (This is discussed in detail in Clements's chapter.)

In the provision of financial services arbitrage would tend to eliminate differences in supply prices which can vary between countries for three general reasons: (i) differences in basic efficiency; (ii) different regulatory 'taxes'; and (iii) differences in the competitive environment sustaining a higher level of costs or profits in one system compared with another. In principle, a high price in one national system (as measured, for instance, by the margin between deposit and lending rates of financial intermediaries or issuing costs in securities markets) would be competed away either through the entry into the domestic system of foreign firms, or by the users of financial services using foreign mechanisms. In practice, though it varies considerably between countries and for different users, this process can be impeded by exchange control, location costs, regulation, and entry barriers together with information and transactions costs. But in practice these impediments are becoming less powerful with significant implications for the evolution of national financial systems. A specific European dimension will be added with the ambition to complete the EEC 'internal market' by 1992.

As finance becomes increasingly globalised so competitive pressures within national financial systems would be expected to intensify. A characteristic of the globalisation of finance is that, at least for some users (predominantly corporate), and for suppliers of financial services, competition has developed from outside the national financial system. Also the entry of foreign institutions has intensified the domestic competitive environment. International competition has developed through users having international options, through financial innovation being transferred, and through the entry of foreign institutions.

However, competitive pressures are not equal for all sectors of business. In practice, the competitive environment in wholesale and corporate sector banking has intensified more than in retail banking though the latter has not been immune. The net result is that market conditions in much wholesale and corporate banking business approaches the characteristics of perfect competition, while profit margins in general are wider in retail business. In the process, a two-tier structure of banking has emerged in which the corporate sector increasingly has global options while the financial intermediation of the retail–personal sector is still limited mainly to within national financial systems. This is partly because, despite developments in technology, servicing the retail sector frequently requires a branch network, and the size of individual transactions is small. It is largely through the corporate sector that global competition has developed between national financial systems.

Diversification and Risk

The traditional, prudential case for a financial system based on specialist financial institutions, or at least keeping separate certain types of financial business, was outlined earlier. However, the risk argument is open to challenge on the same grounds as an efficiently diversified portfolio of assets is less risky than a concentrated one. Enforced demarcations may increase risk. In a US study Heggestad (1975) finds that commercial banking is one of the most risky activities when risk is measured in terms of the variability of profits. He also found that the returns to some activities that might be conducted by banks (real estate agency, insurance broking, etc.) are negatively correlated with banking and hence that banks could reduce their risk

exposure by diversifying. Wall and Eisenbeis (1984) also conclude that had banking organisations been passive owners of some prohibited activities, their earnings might have been less volatile during the 1970s. In fact several activities that *prima facie* appear more risky than banking have returns that are negatively correlated with those of banking, suggesting that such activities and a more diversified business would be risk-reducing for banks. Carron (1983) also finds that the lack of portfolio diversification at thrift institutions in the USA has caused both lower and more volatile earnings compared with banks.

In general, diversification into new (albeit related) areas by financial institutions can reduce overall risk if either the new areas are in general less risky (earnings less volatile) or the earnings are negatively correlated with those in banking. This is an issue of importance when considering diversification.

Some analysts make reference to the experience in the USA in the early 1930s. In fact there is little evidence to support the proposition that the cause of banking failures was associated with the links between commercial and investment banking. The wave of banking failures was associated more with local adverse conditions in particular States combined with a totally inappropriate monetary policy response of the Federal Reserve, which operated in a direction opposite to the lender-of-the-last-resort role.

III CREDIT TRENDS

Other contributors to this volume have focused upon credit, money and monetary policy. Structural change and financial innovation in the financial system have a relevance to the interpretation of money supply and credit trends. Thus Chrystal, in his chapter, notes that it changes the relevant concept of 'money' with direct implications for velocity and the efficiency of particular aggregates for target purposes. It is also likely that a more competitive, innovative and dynamic financial system makes the conduct of monetary policy more difficult as the instruments of policy become less effective. It certainly implies that direct control mechanisms become increasingly ineffective, especially in the absence of exchange control as their effect can be circumvented via external intermediation such as occurred through the Euro sterling market in the year between the abolition of exchange control, in 1979, and the abandoning of the 'corset'.

Structural change in the financial system has direct implications for the exceptionally sharp growth in credit during the 1980s. The savings ratio has fallen sharply and the volume of borrowing has risen substantially. At the same time, the personal sector has built up holdings of liquid assets on a substantial scale. These simultaneous trends are associated in part with structural and competitive changes in the financial system. They are linked with a general process of deregulation combined with a more intense competitive environment and changes in the behaviour pattern of financial institutions.

The extension of credit (more especially consumer credit) accelerated sharply after 1980. The volume of outstanding debt of the personal sector rose, as a proportion of disposable income, from 45 per cent in 1980 to a record 85 per cent in 1987. At the same time, the savings ratio (savings as a proportion of personal disposable income) declined steadily from 15 per cent in 1980 to around 5 per cent in 1988 – its lowest level since 1959. Simultaneously, the personal sector acquired liquid assets on a substantial scale. As a proportion of personal disposable income holdings of liquid assets rose from 67 per cent in 1980 to 86 per cent at the end of 1987. As the discretionary savings ratio has been negative, and yet there has been a continuing build-up of liquid assets, this necessarily implies that part of the increase in holdings of liquid assets has effectively been financed by borrowing.

A sharp rise in personal sector credit of the magnitude experienced since 1980 may be due either to a rise in the *demand* for credit or an increased *supply* potential associated with the behaviour of financial institutions. A general review of the determinants of credit trends may identify four major components: *trend*, *cyclical*, *stock-adjustment* and *structural* determinants. *Stock-adjustment* factors are by their nature finite. They represent the process through which financial institutions and their customers adjust the structure and level of their portfolios to a new position following some identifiable external change. Such external shocks are frequently associated with policy and regulatory changes. While the adjustment is being made, the flow of credit can be substantially increased. The impact of such adjustments, although by their nature finite, may be substantial while they are taking place. Structural factors relate to the efficiency of the financial system and institutions, and changes in their behaviour. If financial institutions become more competitive and efficient, innovative and responsive to market demands, it is likely that both the supply and demand for their services will rise. This will affect the

demand for both financial assets and credit. If, for instance, institutions become more asset-driven, it is likely that the volume of credit will expand through a supply-side stimulus. In both the United Kingdom and the United States the trend towards financial deregulation, and more intensive competitive environment, have had a significant impact on the total supply and availability of credit.

This general framework can be applied to the UK experience of substantial credit expansion and debt accumulation, together with the rise in the holdings of financial assets, since 1980. Three particular changes in the period are likely to have contributed to a once-for-all stock-adjustment effect in the financial system, inducing a greater volume of credit during the period the adjustment was taking place. The elimination of remaining HP controls in July 1982 and the abolition of the corset would both have had such an effect. After the abolition of the corset a substantial stock-adjustment was made towards mortgages as, being freed from the previous balance sheet constraints, banks sought to increase the proportion of mortgages in their loan books (Llewellyn, 1983). The transfer of houses from the public sector will have had a similar effect. In the period 1981–4, around 630000 dwellings were transferred to the private sector sometimes at substantial price discounts. This created a once-for-all demand for mortgage finance that would not otherwise have developed.

A more permanent change, which has had the effect of inducing a more expansionary trend of credit, has been the change in the competitive environment of the financial system. Deregulation has had the effect of enabling institutions to develop new business and to expand existing business. Above all, it has had the effect of changing the climate towards asset growth and the cultivation of new business. Credit facilities have been more actively marketed. There has also been an internal process of deregulation in some sectors and notably with building societies. The abandonment of the interest rate cartel was not dictated by government, but by the forces of competition from outside the building society sector. It had the effect of reducing the queues for mortgages. Building societies have become more asset-driven and accommodating; queues are shorter, mortgages are being advertised, and the average advance as a percentage of purchase price has risen. The rise in the mortgage rate relative to wholesale money market interest rates, evident since 1980, has also brought into the mortgage market a host of new institutions funding exclusively in the wholesale markets.

. Overall, a more competitive environment where the full range of domestic and foreign financial institutions have targeted the personal sector in their strategic objectives is likely to generate a larger volume of credit. At the same time, partly because the effect of competition has eroded the profitability of lending to the corporate sector, banks made a strategic decision to target the personal sector for new lending business. To the extent that this general competitive environment has also made financial assets more attractive, and a narrower interest rate spread between deposit and loan rates, both the volume of personal sector debt and the acquisition of financial assets would be expected to rise. In practice, the financial system has become more competitive for both deposits and loans and the average spread between loan and deposit rates has tended to narrow. The narrower the differential the lower is the cost of maintaining liquidity on the basis of borrowed funds.

Thus the new competitive environment, induced in part by external and internal deregulation, has been a significant factor in motivating a strong simultaneous expansion in both credit and the holdings of liquid assets. Part of the rise in credit and debt of the personal sector reflects increased efficiency in the financial sector.

IV AN ASSESSMENT

Having reviewed the nature of the structural change in the financial system, and the forces producing it, two specific issues are considered, both related to the fundamental issues under discussion. The first relates to the trend towards financial conglomerates, and the second considers the issue of the basic efficiency of the financial system.

Financial Conglomerates

The wisdom of the trend towards financial conglomerates and substantial diversification might be questioned on grounds of efficiency. There is no certainty that financial institutions have the necessary management skills to handle a very diversified business where the technicalities can be complex. There may also be doubt in some cases about whether, for instance, the different' ethos and cultures of clearing banks, merchant banking and broking can be mixed; the 'deal-oriented' culture of securities trading does not easily

fit within the ethos of a clearing bank. Evidence indicates that alleged 'synergy' is difficult to identify and achieve when totally different cultures and practices are brought together. This appears to be most evident in the case of the American banks who, post Big Bang, bought UK securities trading firms.

Those who are sceptical of the financial conglomerate strategy believe that the costs of policing all the complex boundaries (Chinese Walls and compliance officers) designed to avoid conflicts of interest are high, and the risk remains of losing clients when breaches are published. These may outweigh the alleged benefits of synergy and economies of scale and scope even when they can be identified. It is known that banks have lost banking customers who faced hostile take-over bids from companies advised by their corporate finance departments. In some cases, size has been a problem with unwieldy structures and internal tensions together with information and control problems.

This in turn raises the question of the extent to which the different facets of the conglomerate are managed independently, and of the type of control systems within the administrative structure. The dilemma is that if the different areas are integrated so as to secure the advantage to the consumer of a 'department store', then questions arise about whether such a diverse but integrated business can be effectively managed. On the other hand, if different elements of the conglomerate are established as semi-autonomous companies, the question arises as to whether there is any advantage to the user of financial services. The issue is particularly apposite in the context of a trend (for at least large corporate customers) to buy financial services not from a single bank but from where each particular service is most effectively supplied. The corporate sector is tending to 'unbundle' its banking business, a process which undermines pricing strategies incorporating cross-subsidies.

There may, therefore, be a case for considering the experience of the industrial sector because industrial companies appear to be moving away from the notion that large-scale and highly diversified business is necessarily more efficient. In the 1960s, industrial conglomerates were in favour because they were thought to be able to spread risks and to enjoy economies of scale, but there is a certain disillusionment with the consequences of industrial conglomeration. That strategy is now questioned and many industrial companies subsequently divested themselves of some of the new areas they went into in the 1960s. In some cases, this was because they lost sight of

old strengths and did not fully understand the new business areas, or appreciate the management demands of a diversified conglomerate. In many instances the problems of managing a diversified business diverted management attention from the mainstream business. Overall, large-scale and extensive diversification does not invariably mean higher efficiency.

Within the financial sector a warning that diversification may not always be successful is provided by the 'lifeboat' organised by the Bank of England in 1974. Here the biggest institutions that had to be rescued were former successful specialist HP finance houses, such as UDT, Bowmaker and Mercantile Credit, which had recently diversified into lending to property developers. The wide-ranging diversification of business by the big UK banks has not always been conspicuously successful. Another example of such diversification by commercial banks is to be found in their move into medium-term sovereign lending. Some of the problems encountered may be attributable to many of the banks being new to this type of business. In general, the traditional wisdom and expertise of bankers engaged in international lending is located in the merchant banks, not the clearing banks.

Efficiency of the Financial System

This chapter has considered the parallel issues of financial innovation and structural change in the financial system. A major consideration is the extent to which they have contributed to the basic efficiency of the financial system (see Revell, 1973). The concept of efficiency is difficult both to define and measure in finance and a distinction is made between *structural efficiency* (the range of choice offered in the system and its adaptability to changing circumstances and preferences of users), and *allocative efficiency* (the ability of the system accurately to price risks and allocate funds to where the risk-adjusted rates of return are highest). A third concept of *resource efficiency* (the extent to which financial institutions and markets use efficiently the real resources absorbed in the supply of financial services) raises different issues and is not considered further. However, there is a presumption that competition (both domestic and international) imposes a powerful discipline on the use of resources, as in all industries.

The basic requirements of an efficient financial system are outlined

Table 6.2 Requirements of an efficient financial system

Structural efficiency	
(1)	Institutions and markets should offer a wide range of instruments and facilities with respect to maturity, risk, rate of return, marketability, etc., so that different portfolio preferences can be readily accommodated.
(2)	Particular financial services should be provided in a competitive environment by the most efficient suppliers as reflected in consumers' choice.
(3)	All users of the system should have access to the full range of financial facilities.
(4)	Facilities should exist for changes in the structure of portfolios without disturbing prices.
(5)	Institutions and markets should be innovative, responsive to changing market conditions, and able to accommodate changing customer requirements.
Allocative efficiency	
(6)	The system should facilitate the accurate identification, pricing, diversification and transfer of risk.
(7)	All interest rates and prices of financial instruments should be equally flexible so as to minimise financial flows due solely to price rigidities.
(8)	Pricing signals should contribute to an optimum allocation of real resources.
(9)	The system should provide for savings to be channelled to where the rate of return on capital is highest.
(10)	Regulation should be competitively neutral as between potentially competing institutions and markets, and directed exclusively to issues of consumer protection and systemic risk.
(11)	Allocation effects due to the taxation system should be based only upon market failure and market imperfection considerations, and should be competitively neutral as between potentially competing institutions and markets.

in Table 6.2. In general, and compared with those in most other countries, the British financial system meets these requirements well. This is due in part to four major characteristics: the minimum of regulatory limits on the business of financial institutions; the freedom of entry of foreign institutions; the ability of those using financial facilities to use foreign markets; and the generally competitive conditions in the markets for financial services.

The structural changes in the financial system that have been the focus of this chapter do seem, on the face of it, to have contributed to enhancing the *structural efficiency* of the system in terms of the required characteristics outlined in Table 6.2. There is now considerably more choice in the provision of financial services and the range of available instruments and markets. Both in the range of financial instruments and the type of services available, choice has been increased through the process of 'spectrum filling' described earlier. All financial institutions offer a wider range of products and services than a decade ago and a process of product differentiation has meant that there are fewer discontinuities in the range and types of services and instruments available. In many instances, the precise product can even be 'tailor-made'. One of the most basic of products (the bank account) has also been subject to the same process with banks now offering a much wider range of accounts from the basic current account (with or without the payment of interest) to accounts based upon wholesale money market rates of return, and accounts specifically targeted at particular types of customer.

A theme of this chapter has a direct bearing on the second aspect of structural efficiency: the competitive environment. This element of efficiency has clearly been enhanced as, through the process of diversification, all services are now provided by a wider range of suppliers rather than by specialist institutions. In the retail sector, the most obvious example is the mortgage market where for decades building societies had a virtual monopoly. Basic household mortgages are now supplied also by clearing banks, foreign banks, insurance companies and specialist mortgage institutions funded entirely in the wholesale markets. In many services, the corporate sector increasingly has international options. It is also the case that competition and financial innovation have had the effect of widening the range of customers who can have access to particular services, products and instruments. Although it remains the case that large customers have access to more instruments and facilities than small ones, the trend has been for the minimum size (both of customer and transaction) to decline. A particular example is with respect to risk capital to finance new businesses which, historically, has not always been readily available. A new group of Venture Capital companies (some established by the clearing banks) has had the effect of making risk capital more readily available to small firms.

In general, there has been a significant expansion in the number of secondary markets and the volume of transactions, and hence

efficiency in the facility to change portfolio structures has been increased. However, with respect to stock exchange investments it is only alpha and beta stocks for which there is a fully developed market-making system. Thus for gamma stocks registered market-makers are obliged to post only indicative two-way prices on SEAQ although most quote firm prices on enquiry. The requirements are even less for delta stocks.

Competitive pressures have unambiguously made institutions and markets more innovative and responsive both to changing market conditions and customer requirements, with the result that the fifth characteristic of *structural efficiency* identified in Table 6.2 has been enhanced.

Issues related to *allocative efficiency* are more difficult to identify and measure because their effects are not directly observable. Nevertheless, there is a presumption that structural change and financial innovation have contributed to *allocative efficiency*. The development of the Unlisted Securities Market since 1980 has created a secondary market in the equity of (small) companies who could not be listed on the regular stock exchange (see Buckland and Davis, 1984). Many of the financial innovations of the 1980s have had the effect of allowing risks to be priced and transferred. Interest rates in general have become more flexible and differentiated in several respects: (i) the abolition of the building society cartel has made building society interest rates more responsive to market trends, with the result that the 1980s have not experienced the sharp changes in competitiveness that were evident during the 1970s when building society rates were adjusted with a lag and within a narrower range than competitor interest rates; (ii) interest rates have also come to be more a reflection of relative risks and costs than in the past; thus building societies now frequently charge a lower interest rate for large mortgages compared with the historical norm of charging premium rates for large loans; (iii) building society interest rates are also now adjusted to market-clearing levels rather than set at levels that maintain an excess demand for mortgages; and (iv) banks also now apply a wider range of loan interest rates, partly to reflect relative risk and costs of loans to different types of customer. In general, these trends create a presumption that pricing signals now more accurately reflect risk and contribute to an efficient allocation of real resources in the economy.

In many respects, changes in regulation have been beneficial as measured in terms of the contribution to efficiency. Regulation can

be anti-competitive in nature and undermine efficiency in many ways: overregulation imposes a 'tax' on financial institutions which raises the supply price of services; frequently it limits the allowable range of activities, and it is sometimes not competitively neutral as between competing institutions. At the same time, entry regulations offer a degree of protection to domestic institutions from foreign competition. Explicit and self-imposed deregulation has removed limits on the allowable range of business activities of financial institutions, and the 1987 Building Societies Act has abolished many of the restraints on building societies which were clearly not competitively neutral. Similarly, competition and government prompting have undermined cartels and restrictive practices in finance. However, there is more doubt about the general impact of regulation which, as already noted, and most especially as a result of the 1987 Financial Services Act, has become more interventionist, detailed, encompassing a wider range of services, and involving a multiplicity of regulatory agencies. This may have the effect of imposing costs on the suppliers of financial services which outweigh the benefits derived in terms of consumer protection and potential systemic stability.

Historically, taxation has had a major impact in finance, not least with respect to determining the effective rate of return on certain assets and the encouragement given to certain types of saving and investment. There have been substantial tax incentives in favour of government borrowing, investment in houses, saving via life assurance and pension fund schemes, and leasing, etc. Although many such allocative distortions remain, there has been a trend during the 1980s to limit the range and impact of these tax effects. In particular, the tax benefit in favour of life assurance and leasing has been abolished, the limit on the size of mortgage that qualifies for tax relief on the interest payments has been held comparatively static while property prices have risen sharply, and the composite rate of tax paid by building societies on behalf of their depositors has been extended to banks.

Assessment on Efficiency

There must be a strong presumption that competition (both domestic and international), and the wider range of choice in financial services, has contributed positively to both *structural* and *allocative efficiency*. Competition has had the predictable effect of eliminating excess

demand (most especially in the mortgage market), and of undermining cross-subsidies, restrictive practices, and cartels (building societies and the Stock Exchange). The allocation of finance has also come to be more dominated by the price mechanism than by rationing. To the extent that many of the structural changes in the financial system that have emerged are associated with the abandoning of restrictive practices and the undermining of self-imposed anti-competitive mechanisms, then the presumption is that efficiency has increased, however difficult in some areas it is both to identify and measure. Technology has also contributed to efficiency in ways already noted.

On the negative side, it is probably the case that in some areas conflicts of interest are less satisfactorily dealt with compared with the past. It might also be argued that systemic risk will increase in three ways: (i) competition may have the effect of lowering the risk-threshold of lenders; (ii) competition may make profitability more volatile and on average lower thereby undermining the capital base of financial institutions; and (iii) because of the eroding of traditional demarcations and the wider range of business activities of financial institutions, though for reasons already noted, diversification may have the effect of reducing both micro and systemic risk. In the final analysis these are empirical rather than theoretical issues and will be tested in the course of time.

References

Bain, A. D. (1981) *Economics of the Financial System* (Oxford: Martin Robertson).

Bank for International Settlements (1986) *Recent Innovations in International Banking* (Basle).

Bank of England (1983) 'Competition, Innovation and Regulation in British Banking', *Quarterly Bulletin* (September).

Buckland, R. and E. Davis (1984) 'The Unlisted Securities Market', *Lloyds Bank Review* (October).

Carron, A. (1983) 'Political Economy of Financial Regulation', in R. E. Noll and B. M. Owen (eds) *The Political Economy of De-regulation* (Washington: American Enterprise Institute).

Goodhart, C. A. E. (1987) 'Structural Change in the British Capital Markets', in Goodhart *et al.*, *Operation and Regulation of Financial Markets* (London: Macmillan).

Heggestad, R. (1975) 'Riskiness of Investments in Nonbank Activity by Bank Holding Companies', *Journal of Economics and Business* (June).

Llewellyn, D. T. (1983) 'Banks in the Mortgage Market', *Building Societies Institute Journal* (November).

Llewellyn, D. T. (1985) *The Evolution of the British Financial System* (London: Institute of Bankers).

Llewellyn, D. T. (1986) *The Regulation and Supervision of Financial Institutions* (London: Institute of Bankers).

Llewellyn, D. T. (1987) 'Competition and the Regulatory Mix', *National Westminster Bank Review* (August).

Llewellyn, D. T. (1987) 'Financial Intermediation and Systems: Global Integration', in D. E. Fair and C. de Boissieu (eds) *International Monetary and Financial Integration: European Dimension* (The Hague: Martinus Nijhoff).

Llewellyn, D. T. (1989) 'Competition and Structural Change in the Financial System', in E. P. M. Gardener (ed.) *The Future of Financial Systems and Services* (London: Macmillan).

Revell, J. R. S. (1973) *The British Financial System* (London: Macmillan).

Wall, L. and R. Eisenbeis (1984) 'Risk Considerations in De-Regulation of Bank Activities', Federal Reserve Bank of Atlanta, *Economic Review* (autumn).

7 Money in International Trade

Alan W. Clements

Any essay on 'money' and 'international trade' must begin with some attempt at definitions. But the trouble with 'money' is that definitions of it not only are numerous, but also elusive. Perhaps the most interesting are the most general. For our purposes a general definition should suffice: money is whatever is generally accepted and used as a means of payment or in settlement of debts. This means that this chapter will be using figures of world money supply, as supplied by the World Bank and analysed in IMF International Financial Statistics, and at times the definition will be expanded to cover the whole array of international payments systems, including money markets, foreign exchange markets, loan and bond markets, and so on. Similarly, with the other part of my theme – 'international trade' – if a start is made with exports or cross-border trade flows, but I move on to include investment and the financing of investment, it will be because they are inextricably mixed and part of the same whole. Investment follows trade, and both call for finance, and no discussion of the relationships between 'international trade' and 'money' can avoid a detailed look at the developments in the global international financial scene since the 1950s.

However, it is wise to start with a relatively simple model of the relationship between money, or means of payment, and international trade. There is a rather obvious, but nevertheless important point to make at the outset, namely that international trade differs from domestic trade in that one party to a trading transaction must deal in a foreign currency. Thus, we will be much concerned with the influence of foreign exchange markets on international trade, and in particular with exchange rate movements over the longer term, and fluctuations in the shorter term. We will need to look also at the effect of the change from 'fixed' to 'floating' rates in the early 1970s. In other respects, international trade is much the same as domestic trade in its monetary aspects, in that it calls for significant amounts

of finance – finance to keep the exporter's business alive while he waits for payment by the customer, and finance to meet the costs implicit in the substantial stock levels needed in a variety of locations if one is attempting to trade on a global basis. In addition, and here there can be a difference when compared with domestic trade, investment in fixed assets – factories, plant and machinery, and so on – often follows trade, and calls for more permanent forms of finance than are normally employed in financing working capital. The level of demand for the company's goods in a foreign country may well grow to the point where it becomes desirable to have a manufacturing or production facility there, close to the foreign customer, and providing him with greater security of supply. This will call for medium-term and even for long-term finance. Thus, in addition to the foreign exchange markets, we will be concerned with the money markets, with the market in loans from banks, and with the bond and securities markets. We will need to look at their ability to finance the needs of trade, and its back-up investment, and we will have to see to what extent the evolution of those markets has moved in sympathy with developments in international trade, or not. And since these markets exist to serve many parties other than international traders – governments, domestic traders and manufacturers, investors, speculators, and so on – we will, from time to time, have to comment on the ways in which developments in the 'money markets' at large, caused by extraneous factors, have in turn affected international trade.

1 STATISTICAL BACKGROUND*

With this as a very general background to our theme, what do the statistics tend to suggest? If we use export volumes to measure world trade growth, we can use some statistics in the World Bank's 1987 Annual Report to compare the growth in world trade with the growth in GDP worldwide. They show that trade has, in fact, been growing faster than GDP, as seen in Table 7.1.

Our objective is not really to enquire into why this has been so – clearly, it has a lot to do with growing specialisation and increasing interdependence of economies. For our purpose, the significant point

* I am indebted for the material and background in this section to Richard Freeman, ICI's Chief Economist.

Table 7.1 Growth of world trade and world GDP, 1950–85

Annual % changes	*1950–73*	*1973–9*	*1979–85*
World real GDP	5.7	2.7	2.0
World export volumes	9.8	6.0	3.5

is that we would expect to find that the financing of world trade – the 'money aspect' of it – has been rising at a faster rate than the total financing of all economic transactions. Up to about 1980 prices in world GDP rose at much the same rate as world trade prices. Since then, however, world trade prices have fallen by some 10 per cent, whereas general world inflation has more than doubled – it has in fact increased by 120 per cent.

In nominal terms, world money supply grew at much the same rate as world exports, but from a slightly higher base, from 1950 to 1980. Since 1980, the two series have diverged sharply – world money supply has increased by 155 per cent, and the value of world trade by only 16 per cent. The two are plotted in Figure 7.1.

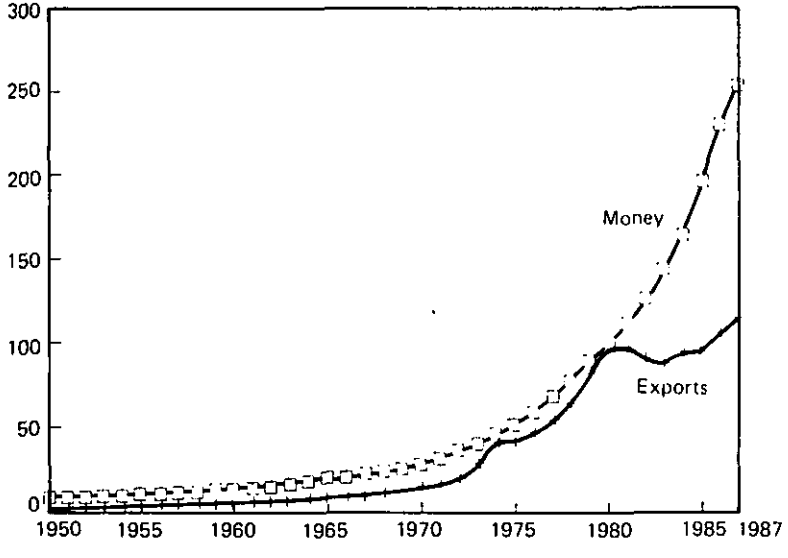

Figure 7.1 Growth of world money and world exports, 1950–87 (1980 = 100)

Obviously, this is not the whole story from our point of view – we need to determine what part of the growth in world money crossed borders, i.e. involved foreign exchange, and was related to the financing of trade and non-trade flows. There could be lengthy debate over which measure to use, but most economists would probably agree that a reasonable indicator of cross-border monetary flows is the published figure of 'cross-border interbank liabilities' in the IMF International Financial Statistics. Unfortunately, the series only goes back to 1971. However, in prefloating days and with fairly general exchange controls, their relationship to world trade was probably fairly constant. Again, after 1980, cross-border interbank liabilities moved sharply away from world trade values, indicating a massive growth in non-trade related monetary flows. The figures are plotted in Figure 7.2.

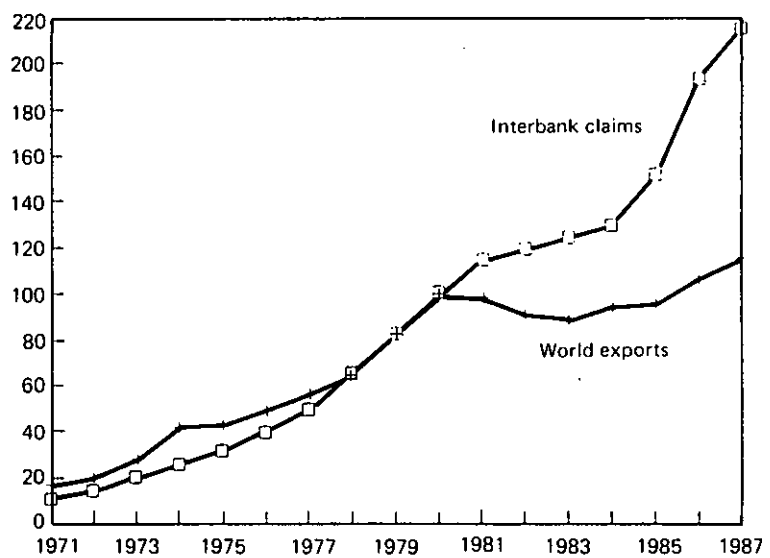

Figure 7.2 Growth of interbank claims and world exports, 1971–87 (1980 = 100)

Other statistics tell essentially the same story. For example, world money reserves have grown by about 65 per cent since 1980, while world trade, as we have seen, has grown by only 16 per cent. Another indication is given by a comparison between the foreign assets of deposit banks as a percentage of nominal GDP with foreign trade

(in goods and services) as a percentage of GDP, as in Table 7.2.

Table 7.2 Growth of foreign assets and foreign trade (% of GDP)

	1960	*1984*
Foreign assets of deposit banks as % of GDP	1.5%	17%
Foreign trade as % of GDP	11.3%	18.3%

Sources: OECD *Economic Studies* (Spring) 1987, and OECD *Historic Statistics*

The data, then, tend to show that non-trade-related monetary flows are now substantially larger than trade-related flows. If we are correct in believing that this was not so, or not so to the same degree, in the 1960s and well into the 1970s, then there has been a massive change in the international financial environment, almost 'geological' in its scale, which must have had a major impact on trade and investment. Experts agree on one thing – it is almost impossible to obtain accurate figures which will give a satisfactory measure of non-trade flows as compared with trade flows. However, while some would put the non-trade element as high as 80 per cent of the total, most are agreed that it must amount to at least 70 per cent.

Other figures relating to the financial markets themselves give some sort of indication of the growth in financing since 1980 (Table 7.3), which, when related to the figures for world trade already quoted, cannot in the main have been trade induced.

Table 7.3 Growth in financial instruments, 1980–6 ($ billion)

	1980	*1981*	*1982*	*1983*	*1984*	*1985*	*1986*
International bonds and notes*	39	44	72	72	108	164	220
Euro note facilities*			10	15	27	50	27
International interbank lending ($ trillion)†		1.5	1.25	1.60	1.70	1.85	2.35
Interest rate swaps†			10				310
Currency swaps†			10				40
Futures†	80						440

Sources: * Bank of International Settlements 1987, *Annual Report*
† *The Economist*, 21 March 1987

It is generally agreed that the extremely rapid growth in financial instruments depicted by figures such as these has been associated with, if not necessarily caused by, deregulation in many financial

markets since 1979 and striking financial innovation, encouraged by technological changes. It is interesting to look at what has been happening at the same time to exchange rates and interest rates.

As regards exchange rates, it is widely believed that since the mid-1970s volatility has increased, and there have also been longer periods of severe disequilibrium, with some currencies over- or undervalued when compared with others. The following figures (Table 7.4) for the key currencies – the US dollar, the Deutschemark and the yen – substantiate the view on volatility:

Table 7.4 Volatility of key currencies, 1960–86

		1960–73	*1973–79*	*1979–82*	*1982–86*
Average absolute Monthly % change	DM/$	1.05	2.40	2.75	2.80
	Yen/$	0.50	1.95	3.30	2.35

As for disequilibrium, it is now widely accepted that after the oil crisis in the 1970s, and all the attendant disturbances, the US dollar and pound sterling became overvalued by 40 per cent or more in the early 1980s, and the yen and the Deutschemark became undervalued by approximately the same amount. Since then cross-rates have moved more into line with appropriate, or competitive, values but the process of change has been by no means smooth.

The story on interest rates is given in Table 7.5 – in effect, a sharp increase in *real* rates since the 1980s.

Table 7.5 The increase in real interest rates, 1960–85

	1960–67	*1968–73*	*1974–79*	*1980–85*
US short-term rates	1.3	0.4	−1.0	4.5
Japanese short-term rates	0.6	−1.5	−2.6	3.3
German short-term rates	1.1	1.6	0.8	4.3

We will return to these developments again when we trace the history of the relationship between money and international trade in a more descriptive way, but for the time being it is fair to say that this statistical survey has indicated the following conclusions:

(1) World money supply grew in line with world trade up to about 1980, but then diverged, with money supply growing at a much faster rate.

(2) The same picture emerges if we look more strictly at monetary flows connected more directly with cross-border transactions – measured by the figures for 'cross-border interbank liabilities'. Again, there is a divergence after 1980, revealing a huge increase in non-trade-related monetary flows.
(3) This significant growth in non-trade flows from 1980 onwards is borne out by the statistics for the international bond and money markets, and for derivatives such as swaps and futures. This whole geological upheaval in the financial markets has been accompanied by increasing exchange rate volatility, by more pronounced currency disequilibrium, and by an increase in real interest rates.

II THE 1950s TO THE 1970s

At this point it seems that continued general analysis, particularly of the statistics which are available, will not yield much more of real interest on the relationship between money and international trade. However, the statistics do point to an interesting theme, namely: how was trade, and the investment which accompanied it, financed in the 1950s, 1960s and into the 1970s, and what has happened since as a result of all the changes we have been describing? What is needed is an account of how it was; of what has happened to the financial markets over the years; and how it has changed for the international or multinational trading company. This can best be done by attempting a historical and descriptive survey, based largely on experience. Much of what follows, therefore, is drawn from the financial history of ICI, a UK-based industrial and trading company, which has nevertheless been international and, more recently, multinational, in its activities throughout the period. The approach will be to describe the external financial environment, and the changes taking place within it, and then to try to explain their impact on the financing of international trade and its accompanying investment. Since we are concerned primarily with trade, the description of the financing problems begins in each case by looking at the short-term aspects, though reference will be made to long-term financial markets.

As a generalisation, the financial world of the 1950s and 1960s was a continuation of the 1930s, and not of the long period of economic freedom which had preceded that decade. In the 1930s exchange controls, trade barriers, protection and competitive devaluations had

disrupted and broken up the open and liberal economic world of the late 1800s and early 1900s. In spite of attempts to impose a more orderly system, such as Bretton Woods, by and large the practices and policies of the 1930s were picked up and applied again after 1945–6. For example, exchange controls were fairly common, and in all countries the financial services industry was regulated in one way or another.

It is worthwhile reminding ourselves briefly of the form these controls, regulations and barriers took, by looking at just two countries – the USA and the UK – before moving on to describe how the whole system began to change and eventually broke up. In the USA, the continuation into the 1950s and 1960s of the Federal Reserve system and SEC was not of great significance for trade, but the preservation of Glass Steagall, one state banking, and all the other regulations applicable to the financial services sector was. What it meant was that the world of 'finance' or 'money' had become separated into clearly defined segments which hardly competed with each other, and which were not really under any compulsion to change or innovate. There were clear divisions between commercial banking, investment banking and insurance. Interest rates were controlled, and so were the types of loans traders and manufacturers could use. From time to time there were controls over foreign investment, and also forms of exchange control, e.g. the Interest Equalisation Tax which, from 1963 to 1973, made it expensive for foreigners to borrow dollars in the USA and take the funds out of the country. In the UK, the system was broadly similar, except that exchange control was much more comprehensive and persistent. In banking and financial services generally, the Bank of England ensured that each player kept in his place and performed his proper role, and that outsiders were unwelcome. In commercial banking this meant that after a period of consolidation, there was little or no further development; in merchant banking it meant that, with limited capital bases, any aspirations that British investment banking might become a major force had to be abandoned in favour of a more limited domestic role; and on the Stock Exchange it meant that the jobber–broker system, plus fixed commissions, plus the exclusion of outsiders amounted to the preservation of a market which became increasingly disconnected from the developing world around it – for example, it almost completely ignored the burgeoning Eurobond market, even though it was largely based on London.

In spite of controls, restrictions and imperfections such as these, it

has to be said that the financial markets coped with rates of growth and expansion in the 1950s and 1960s which have not been experienced since. The average annual percentage growth rate, in real terms, of world GDP was 5.7 per cent from 1950 to 1973, and the figure for world trade (export volumes) was nearly 10 per cent. Since then growth rates have been considerably lower – 2.7 per cent (1973–9) and 2.0 per cent (1979–85) for GDP, and 6.0 per cent and 3.5 per cent for world trade for the same periods. During that earlier period, as we have suggested, capital and credit markets grew in line with economic growth and the expansion of trade. But the important point is that, as in earlier periods, the financial system continued to evolve and develop in order to serve needs in the real economy, i.e. trade in goods and services, and flows of direct investment. Throughout the 1960s, for example, and also well into the 1970s, the market behaved as though underlying real economic performance was a key determinant of financial trends, and in fact responded to economic announcements in the expected way. In the foreign exchange markets, the key figures on everyone's lips were those of balance of payments, and decisions as to whether or not to sell or buy forward or to wait and sell or buy spot turned very much on which forecast of the likely outcome in terms of the figures to be announced, and therefore also of the reaction in the market, one tended to adopt.

It would be entirely wrong, however, to give the impression that the period from the 1950s to the 1970s was one of quiet, regulated and harmonious stability in the financial markets. Massive changes were in fact taking place, and to the trader and manufacturer, trying to find the optimal methods of financing both trade and investment, the changes and upheavals often seemed more conspicuous than the stability which the regulated and steadily growing market places afforded. The most important of these, from our point of view, were the collapse of Bretton Woods, the growth of the Euro-markets and of the international capital market, the oil crises and the other major economic disturbances of the 1970s, and finally the onset of deregulation. After a brief description of each of these, we will look at the ways in which this whole financial background influenced and determined the ways in which trade and investment were financed during the period.

Bretton Woods – the fixed exchange rate system under which most currencies were tied to the US dollar – was the main plank of the post-war free world foreign exchange system. By 1958 reasonably balanced currency relationships had been re-established, and the

European currencies, for example, became freely convertible. The system contained a mechanism whereby currencies in disequilibrium could adjust, but had no means of coping with a situation in which the reserve currency itself – the US dollar – came under attack. From the early 1960s onwards the USA produced a series of balance of payments deficits and some central banks, which had tried to sell dollars for gold, agreed to permit a policy of 'benign neglect' which, in effect, meant a continuation of US deficits while surplus countries supported the US currency. This regime, however, could not last forever – by 1969 the effects of the Vietnam War and rising prices had produced a US deficit of $6 billion, and it was becoming increasingly obvious that either the USA, or Europe and Japan, would have to change their currencies' parities. In October 1969, the Deutschemark was revalued; in 1970, there was more pressure on the dollar; in May 1971, the Deutschemark was floated, and there followed further sales of dollars; in August 1971, gold sales were suspended, the USA imposed a 10 per cent surcharge on imports, and Europe and Japan 'floated'; in December 1971, the Smithsonian realignment revalued the yen by 17 per cent and the Deutschemark by 14 per cent; in June 1972, sterling came under attack and was floated; and finally, in early 1973, further speculation against the dollar (during which in just two days the Bundesbank purchased $6 billion!) culminated in full-scale floating and a new gold price of $42 per ounce. It was by this process that the world moved from 'Bretton Woods' to floating exchange rates, a regime under which currencies were supposed to adjust to underlying economic realities smoothly and rationally, and in which central banks were supposed to have much less need for reserves than in the past. As we shall see, neither of these forecasts proved correct. But the process of change briefly described here was obviously a stormy and turbulent one, and what made it worse was that it coincided with a period during which many international trading companies were trying to expand. In the case of UK-based companies, it came just after a massive devaluation of sterling in 1967. It formed, therefore, a powerful backcloth to the financial development of many trading companies, and significantly impacted on their financial systems.

Fortunately for some of the companies concerned, one development in the 1960s which was the product of the US balance of payments deficits, and of the regulatory system of the time, came to their aid – namely, the Euro-currency market. In the late 1950s US balance of payments deficits resulted in the accumulation of dollars in European

hands and London, as a centre of both trade and finance, saw the opportunity to create dollar deposits with banks in London, free of both US minimum reserve requirements and Regulation Q (which limited the rate of interest payable on deposits in the US). These deposits grew rapidly – by the late 1960s they amounted to $100 billion; by 1975 to $600 billion; and by 1985 to $2500 billion. The Eurodollar, and later Euro-currency, market quickly became of great value to traders and manufacturers, not only because it constituted a new source of finance in the form of short-term and, later, medium-term loans, but also because it facilitated currency hedging and made it easier for companies to cope with exchange controls. Dollar exposures arising from trading and also from investment, for example, could be hedged by borrowing Euro-dollars, and where the financing of overseas activities was not possible using sterling, again Euro-dollars came to the rescue. And there was no doubt that, in time, the Euro-dollar market acted also as a major stimulus in the development of the foreign exchange market. These forms of assistance to trade and industry were underlined by the subsequent development of the Eurobond market, initially in the form of US bond issues floated outside the USA and involving both non-US investors and borrowers. This market was also created largely by regulations attempting to restrict and control capital flows, e.g. the Interest Equalisation Tax, which made it costly for non-US borrowers to raise funds in the USA, and the 1968 measures which tried to stop US companies borrowing in the USA for investment overseas. Although this market grew much more slowly than the bank deposit market behind it, it became of considerable significance to companies.

One part of the whole Euro-currency market experienced particularly explosive growth during the 1970s, and that growth in turn helped to induce further instability and volatility in both money and currency markets. From 1968 onwards the major commercial banks began to make available large medium-term syndicated loans, and in the mid and late 1970s this process was intensified as a result of the oil shocks and the ensuing glut of so-called 'petro-dollars' in the hands of the oil producers. The recycling of these funds to the borrowers most in need at the time – the LDCs – has often been presented as a normal process of intermediation, but there can be little doubt that the existence of the large and flexible Euro-markets enabled a process to take place whereby lenders who might have been reluctant to lend on their own account to countries already massively in debt, were prepared to lend to banks which, in turn,

were prepared to lend in syndicates, or consortia, to those countries. And the syndicated loan system itself obscured the normal process in a bond or security market, under which borrowers are subjected to credit or rating analysis, and in time find it more and more difficult to borrow if their ability to repay declines. The whole system had become much more unstable on several counts – many of the major borrowers could not service their debt, and certainly were in no position to repay it; many major banks, whose assets contained large loans to LDCs, were as a result weakened; the corporate sector had suffered to some extent because of the emphasis on lending to LDCs, and was now experiencing the worst 'downturn' since the 1939–45 war; and exchange rates were no longer determined by trading balances and flows, but by the tendency on the part of depositors to change the currency designation of their massive deposits with the Euro-banks as they were renewed, according to their short-term views on the relative strengths and weaknesses of the major currencies. Again, the impact on the normal processes of trade and investment was massive.

In the meantime, the international financial system had been subjected to a number of other shocks. The demise of the Bretton Woods system, and the first days of the new floating rate regime had coincided with a worldwide boom and rising commodity prices. There followed the first great oil crisis of 1973–4, and in 1974 and 1975 rampant inflation – rates around 20 per cent per annum were experienced in the UK and Japan, and of 12½ per cent per annum in the OECD generally. The major economies drifted into a period of 'stagflation', characterised by budget deficits, until the second oil crisis of 1979–80 produced a second bout of inflation – 12½ per cent again in 1980. Then came the great reaction, especially in the USA from 1982 onwards – sound and tight money policies, producing a strong dollar (35 per cent appreciation between 1980 and 1985) and capital inflows. By 1986 inflation in the OECD was down to 3 per cent per annum, but the USA had maintained growth in the free world via its budget and current account deficits and, as a result, had become a debtor on a scale undreamt of even in the mid-1970s. By 1986 its position as a net creditor in 1980–1 to the tune of $100–150 billion had deteriorated to that of a net debtor to the rest of the world in the massive sum of over $260 billion! And this had occasioned a major shift in the direction of flows of 'capital' across the exchanges. Instead of funds flowing in the main from OPEC to the LDCs, as they did in the 1970s, the major flows now were from the surplus

economies of Japan and Europe to the United States. The great 'geological upheavel' in the international monetary system referred to earlier was now well under way, accompanied by increasing interest rates and increasing currency disequilibrium and instability.

At about the same time, the last of the major shocks to the post-war monetary system began to make itself felt – namely, 'deregulation'. Perhaps it was natural that after the collapse of the Bretton Woods system and the rise of a new regime of floating exchange rates, the emergence and rapid growth of new, unregulated money and capital markets (the Euro-markets), and the disturbances created by the two oil crises and the growth of the LDC debt problem, the controls and regulations on which that post-war system had been based should fall into disrepute. Perhaps it was something more, namely, a positive philosophy to the effect that these very controls had, for example, intensified inflationary pressures after the oil price shocks, and had made it more difficult, rather than easier, for governments to cope not only with inflation but also with mini recessions, deficits, and escalating interest rates. Whatever the cause of the change in approach, in economy after economy, controls began to be dismantled – in the USA, it was primarily a case of reversing the very considerable amount of legislation which restricted the activities of banks; in the UK, exchange controls were abolished in 1979, and then the capital market (the Stock Exchange) was fundamentally recast, resulting in a wave of mergers and acquisitions in the financial sector; in Germany, a number of fundamental tax changes on interest, dividends and on securities transactions were implemented; and in Japan, exchange controls were abolished and the new issues market liberalised. And this process materially helped others – 'securitisation' (or the increasing use of marketable paper, instead of the traditional bank loan, by borrowers), and 'innovation' (or the rapid growth of a wide range of derivatives, such as futures, swaps, options, forward rate agreements, caps, floors, and collars, and so on). These came to be used not only in connection with new financing operations but also to hedge or modify existing liability or asset situations, and thus intensified both activity in the financial markets and volatility of both interest and exchange rates.

Before we move on and look at the ways in which traders and manufacturers were forced to finance their operations during this long period to the end of the 1970s, it is worthwhile attempting some sort of summary. I have tried to do just that with the aid of three graphs – Figures 7.3, 7.4 and 7.5. They depict what happened to

three key financial indicators during the period 1960 to 1988 – exchange rates, short-term interest rates, and long-term interest rates.

Figure 7.3 charts the progress of two exchange rates of crucial significance to international traders based in the UK – sterling against the US dollar, and sterling against the Deutschemark. During the 1960s, with Bretton Woods still operating and exchange controls still effective in varying degrees, the rates were relatively stable, and up to 1967 there was little need to be concerned about wild fluctuations which might destroy trading margins altogether. The chief problem for the UK exporter and importer was whether or not he should sell forward his currency receipts or buy forward his currency requirements, or wait and deal spot at the time of completion of the transaction. With spot rates held within close margins to agreed 'fixed rates', and forward rates reflecting interest rate differentials, this did not present a *major* headache, except during periods when devaluations or revaluations were rumoured and expected – 1967 for example. The large devaluation of the pound in 1967 was followed by another period of relative stability, and then from the early 1970s onwards by the floating rate regime, characterised by instability,

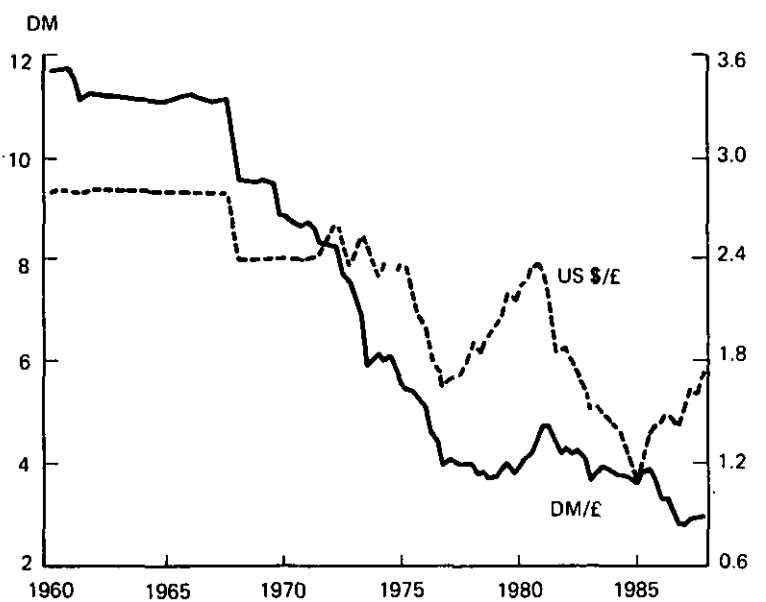

Figure 7.3 Exchange rates

volatility, and obviously also by periods when actual rates represented major variations from competitive or equilibrium rates. Looking at the sterling–dollar line, for example, it is clear that any trader who sold forward several months' worth of his dollar receipts from exports at any of the peaks during the 1970s would have been making a costly mistake. It needs to be remembered also that the graph is based on quarterly averages and therefore masks monthly, weekly and daily swings.

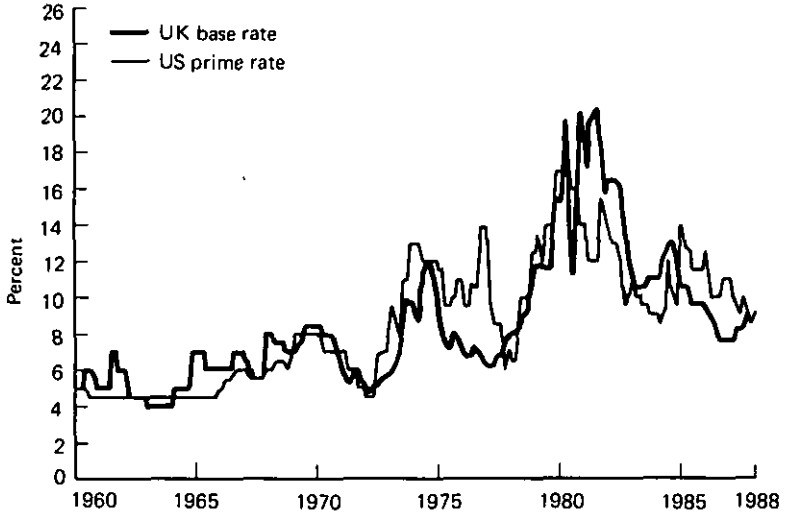

Figure 7.4 Short-term interest rates

Figure 7.4 traces the history of two key short-term interest rates – UK bank rate and base rate, and US prime rate. Short-term borrowing costs for an international trader based in the UK, but spreading his activities into Europe and the Americas, would be primarily related to these two official rates, because to the extent that borrowing had to be based on non-UK sources it would be the Euro-currency markets which would be the obvious choice, and there the principal currency has always been the US dollar. Again, the great contrast is between the pre-1970 and the post-1970 periods. The massive swings in short-term rates around the time of the two oil crises are particularly noticeable. It should also be remembered that during the 1970s 'floating rate' borrowing became more and more the norm. From the

early 1970s onwards, this obviously posed additional problems for corporate borrowers.

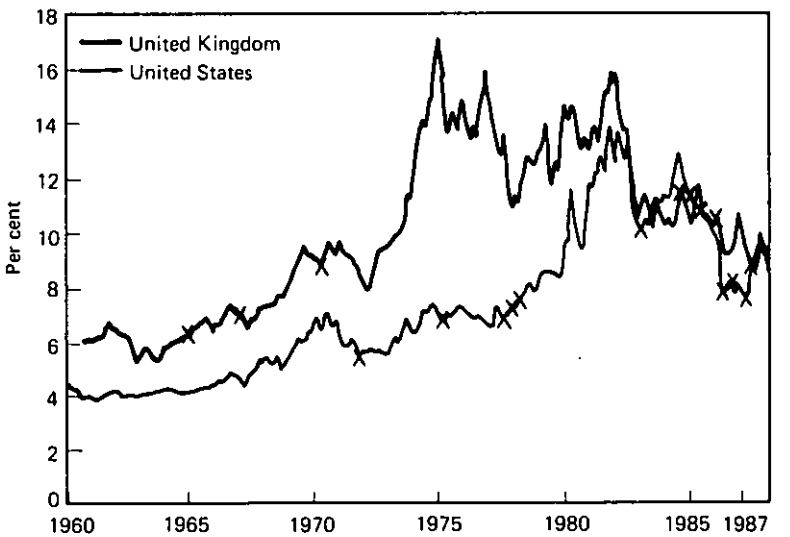

Figure 7.5 Long bond rates

Finally, Figure 7.5 shows what happened to long-term interest rates during the period, in the form of the twenty-year Gilt rate in the UK, and the thirty-year Treasury rate in the USA. Here again, we see a modest rate of increase in rates up to about 1970 and thereafter wide oscillations and much more rapid growth, especially in the UK in the mid-1970s, and also in the USA after 1980. It is obvious that long-term or medium-term financing at fixed rates became increasingly difficult, problematic, and even to some extent hazardous, from the late 1960s onwards. It was, of course, from about this time that many companies which had been benefiting from the relatively high growth rates of the 1960s attempted to expand. The large crosses on the lines mark the times at which ICI made large borrowings in sterling or dollars, and at fixed rates related to Gilts (twenty-year and ten-year) or US Treasuries (twenty-year and ten-year). Achievement of even such a moderate track record as that revealed in the figure obviously entailed relatively long periods of absence from the market, and therefore of dependence on other sources of finance.

III THE IMPLICATIONS FOR BUSINESS

Confronted with an international money system of this nature, how did the international trader and investor set about financing his business? What financial policies did he find it best to pursue, and what particular systems had to be developed to cope with the increasing instability in the system as the years passed by? Obviously, they varied a good deal from company to company. The more capital intensive businesses had to be concerned with the long-term securities markets as well as with the short-term and the currency markets, because of their need for long-term finance; companies exporting from a strong currency base, such as Germany, faced different problems from their UK rivals; and much depended on which particular foreign markets constituted the objective of the international trader. But all companies engaged in international trade and investment during the period, in varying degrees, had to cope with exchange controls, variable exchange rates, problems in covering or hedging their currency receipts and payments, fluctuating interest rates, and unstable loan and bond markets. It is, therefore, possible to generalise to some extent about their financing policies and systems.

Since we are concerned primarily with trade, it is convenient to start with the short-term financial problem, i.e. with exporting and importing, and coping with foreign currency receipts and payments, and with short-term financing – borrowing, management of cash, and so on. As far as the foreign currency problem is concerned, there has for many years existed an interesting, and inconclusive, debate over what constitutes the real currency exposure of a business trading abroad – does it consist simply of its transactions in foreign currencies (i.e. principally its exports and imports), or does it include also its profits and its net assets located outside its home base and denomination in foreign currencies (i.e. its 'translation exposure' in the parlance of the accountants). Fortunately, for many the problem was resolved in simple, if not always elegant, ways. For the exporter pure and simple the problem did not arise – he had only 'transactions' and no assets and profits located in other countries to worry him; for many others, such as UK businesses, exchange control permitted 'covering' or 'hedging' *only* of export and import payments, and other contractual cash flows, such as dividends and interest payments from foreign subsidiaries; and for others, the practical problems of coping with currency transactions proved large and difficult enough

without taking on board significant translation exposures as well. UK companies, engaged in building up exporting businesses, and subsequently involved in creating their own manufacturing establishments within foreign markets, found it necessary during the 1950s and 1960s to develop a system for coping with their foreign currency exposures under the all-embracing cover of exchange control. As the volatility and instability of currency cross-rates increased, such a system became more and more necessary, if only because a 5 or 10 per cent movement in rates could severely erode, or even completely destroy, the profit margin in an export transaction when the final sterling outcome was calculated. Yet at the same time, forward cover often looked expensive. The problem facing both the treasurer and the export manager – to use the forward market or not – became more and more difficult as the currency markets became less and less predictable, and only a comprehensive and consistent corporate system could help the company, at the very least, to avoid the worst pitfalls.

The main ingredients of most of the systems which were developed were roughly as follows:

(1) A central decision, consistently applied throughout the company or group, on which currencies should be used when invoicing export customers. For some time there were those who argued that the currency problem could be avoided by invoicing in sterling. But to attempt to be a major force in a foreign market, such as Germany, while insisting on selling in sterling, seemed almost a contradiction in terms. And once the company established its own manufacturing facilities, in the form of a local subsidiary company in that market, the problem was totally resolved – the German company had to sell its output in Deutschemarks, and one could not have some ICI products invoiced in sterling alongside others invoiced in domestic currency in the same market-place. So, by and large, the policy became: invoice in the currency of the customer (with some exceptions, e.g. the dollar or sterling in some very weak currency markets!). This centralised policy became the more important as large flows of goods began to cross frontiers within the group, i.e. intermediates moved from one company to another before being sold to the final outside customer as finished products. The group invoicing policy here was useful in that it could be used to ensure that the major currency risk was moved back to the centre of the

group, where it could best be managed as part of the central system, and leaving the foreign subsidiary companies as far as possible both buying and selling in their own local currencies.

(2) Accurate and comprehensive forecasts from the businesses of their expected foreign currency receipts and payments – primarily exports and imports, but also other items – for at least the next twelve months. The full forecast was vital, not only because it gave the centre a currency by currency picture of the company's exposure in the near future, but also because it made it possible to consider various forms of cover or hedging.

(3) Based on the currency forecasts detailed in (2) above, a banking system which would permit as much 'matching' as possible. Basically, this consisted of special accounts into which currency receipts were routed; these were then used to make currency payments, leaving only the net balances to be transformed into £s.

(4) Decisions on 'covering' or 'hedging' to be essentially 'central' decisions, made by the treasury team, looking at the whole company or group position, currency by currency, and embracing all the businesses within the group. Attempts to leave this to be dealt with on a decentralised basis, with each business or subsidiary making its own decisions, soon ran into real difficulties. To mention just a few – one part of the company might be selling, say, dollars at the same time as another was buying the currency; opportunities to 'match' receipts and payments before having to sell and buy the currencies concerned would be missed; a whole variety of views on the wisdom of forward cover could prevail, with no consistent policy emerging at all; and finally, close linkage with the short-term position of the group in financial terms, i.e. its liquidity and short-term borrowings, proved extremely difficult if not impossible. Therefore, again, centralisation proved optimal. What tended to emerge were arrangements under which the centralised currency operations, handled by the treasury team, acted as a sort of 'group banker' to the businesses, buying their currency receipts from them, and selling currency payments to them. Depending on the sophistication of the system, businesses were given varying degrees of freedom to negotiate with the centre, e.g. as regards forward sales or payments. The 'centre' was then left with the management of a stream of foreign currency

receipts and payments trying to improve on the sterling position established with the businesses. Thus the businessman's contact with the real world of the currency markets was maintained, even though it was largely by means of a dialogue with the central treasury experts.

(5) The centralised currency management system to be linked closely to, and indeed to be part of, the company's cash management system. In the early stages, when the emphasis was on the optimisation of the export–import cash flows, many of the lessons already learnt from the build up of a domestic cash system were applied. A whole variety of techniques began to be developed in order to optimise the cash position for the group. In particular, it became sensible to operate a central pool of short-term funds. In effect, subsidiaries with surpluses lent to the pool, and those funds were on lent to subsidiaries in need. Soon the 'central pool' was deciding where the group should borrow short term, and how those resources should be routed to the operations requiring short-term finance. It was, of course, never possible to achieve perfection – the group was always effectively borrowing somewhere, and in possession of liquid surpluses elsewhere, rather than being in a perfectly balanced position. This added a new currency problem, related to the group's short-term liquidity position. Since borrowings had to be denominated in one currency or another, and similarly with deposits or short-term investments, it became necessary to manage these short-term currency positions alongside the transactions exposures already mentioned. Short-term currency borrowings or investments could be switched into other currencies, if desired, via the forward market – in effect, a short-term swap market. It was, of course, vital that these manoeuvres should be managed on lines consistent with those being applied in the case of exports, imports, and other cash flows.

In the case of ICI this had been developed, by the early 1970s, into a short-term financing system essentially funding the group's international trading operations, which could best be summarised as one based on 'centralised coordination'. A central finance subsidiary – ICI Finance – had been created in the UK to carry out most of the operations itself, but in addition to coordinate financing activities by many of the overseas subsidiaries of the group. As far as trading

based on the UK assets was concerned, ICI Finance handled the export and import currency flows, as well as other currency receipts and payments; it was responsible for managing the bulk of the group's short- and medium-term borrowing facilities; and in the ways already described it collected and handled the bulk of the group's short term liquid resources. It was also in constant dialogue with overseas subsidiaries, advising and assisting them on their currency exposures and short-term positions. As a result, through the turbulent 1970s, it was possible to grapple with many of the group's problems arising from currency volatility, inflation and interest rate changes in one centralised operation. Many other companies involved in international trade had developed similar systems, or were in the process of doing so.

So far we have been concerned mainly with trade and the short-term financing problems associated with exporting and importing. But, as already mentioned, trade quickly engendered investment in the foreign markets being 'attacked' by the trader. At first, in the case of ICI, this took the form of the build-up of merchanting subsidiaries territory by territory – companies which took in ICI goods exported from the UK and resold them locally. As these operations expanded, they needed substantial amounts of finance for their debtors and stocks. And it was not long before manufacturing facilities were grafted on to them to produce plastics, fibres, pharmaceuticals, crop protection chemicals, and others, within the overseas markets, to supplement exports from the UK. These expansions materially changed the original merchanting subsidiaries, and involved the creation of others, and also of course called for the raising of substantial amounts of long- and medium-term finance. This began to emerge as a major problem in the early 1960s, and again exchange control had a major influence, if only because as far as the UK-based international trader was concerned, sterling could not be used to pay for investments outside the UK. Even though the UK part of the group possessed ample sterling liquid resources, or could borrow sterling, or could raise equity finance from its UK shareholders, it was forced by the controls to borrow foreign currencies. At first, in the early 1960s, ICI, like other UK-based businesses, tapped a number of foreign capital markets for this purpose, borrowing Swiss francs, Dutch guilders and US dollars in the main. Then, around 1963, the Interest Equalization Tax in the US closed that market, and the others began to exhibit signs of instability. Fortunately, the international bond market, initially in the form of US dollar

Eurobonds, came to the rescue, and quickly became a major source of medium- and long-term finance. Faced with this diversity and variety of sources and forms of finance, the international company had to resolve a whole range of new problems when compared with the more normal situation of raising finance domestically. Amongst the many such problems can be mentioned: which foreign capital markets to cultivate; which new financial advisers or investment bankers to use; which currencies to borrow; which types of loans to rely on – public issues, or private placements, and so on; which maturities to aim to achieve; and finally, which companies in the group to use as borrowing vehicles. In the case of ICI these considerations led fairly soon to the development of a long-term financing policy which bore a close resemblance to the short-term system described above in that it, too, was based on 'centralised coordination'. For a number of reasons – taxation, the inability of some of the overseas subsidiaries to service large borrowings themselves in the immediate future, the desire to achieve economies of scale, the long-term aim of managing the group's repayment schedule sensibly, and the wish, at times, to incorporate equity 'kickers' in bond issues – most of the large borrowings were effected centrally, in the name of the parent company or a special, 100 per cent owned, finance vehicle, rather than being floated 'locally' by the operating subsidiary needing finance at the time. It was felt that the latter method, while it would achieve matching of currency risks – the local borrowing would 'hedge' the local investment – could produce, in time, an extremely complex situation for the group as regards both cost of finance and the group repayment schedule. The other main consideration when borrowing – currency risk – was dealt with by trying, wherever possible, to achieve a broad matching relationship between the currencies borrowed and the investments made. This policy of centralisation of major group long-term borrowings produced additional benefits towards the end of the period, largely as a result of the company establishing a reputation in the major capital markets and becoming well known to international investors. For example, as the markets themselves became more volatile and unstable, the group was usually able to pick out at least one where finance was available on reasonable terms, and in addition was in a position to attempt optimum timing of its fund-raising by borrowing large amounts of money in one operation, often well in advance of actual expenditure. Figure 7.5 reinforces this point.

As one moved towards the end of the 1970s, into the second oil

crisis and the recession which followed it, the massive changes in the international monetary system which were described earlier meant that many of the salient features of these corporate financial systems also came under pressure for change. It is now necessary to look at what has been happening in the 1980s, and what is likely to happen in the future.

IV THE 1980s AND BEYOND

It must have become very obvious from both the statistical section at the beginning of this essay, and from the historical survey of the 1960s and 1970s recounted above, that by about 1980 a number of fundamental and massive changes had taken place in the international monetary system. A few figures help to summarise them and bring home their magnitude:

(1) *World or International Trade* – the IMF International Financial Statistics estimate this at $2175 billion for 1987. Some commentators have put the figure as high as $3000 billion – for our purposes a figure somewhere between $2000 and $3000 billion will suffice.
(2) *The Euro-currency Markets* – by 1987 these must have amounted to at least $3500 billion in size. But that figure is an estimate of the amount of the deposits in that market at any one time. Deposits are, of course, turned over several times a year, and are borrowed and reborrowed several times as well. The estimates of 'turnover' in the market vary considerably, but figures as high as $50 000 billion to $75 000 billion haved been mentioned.
(3) *The Currency Markets* – estimates of the volume of trading in the foreign exchange markets are often quoted in the region of $150 billion or more per day – say $37 500 billion per year. The figure is lower than the volume figure for the Euro-currency markets because not all Euro-currency deposits, when rolled over, involve a change of currency. On the other hand, in the foreign exchange markets one underlying transaction, e.g. a receipt of currency x in y months' time, can give rise to several transactions before completion. Nevertheless, if there is any approach to accuracy in these figures, and they are widely accepted, it is clear that both the Euro-currency and the foreign exchange markets are significantly larger than the volume of international trade would suggest. In fact the bulk of the trans-

actions in both markets can have little or nothing to do with basic buying and selling of goods and services across national frontiers.

(4) One or two other figures can be quoted to continue the story. The *Eurobond* or *International Bond Market* now manages new issues of the order of $250 billion per annum, and the turnover in the market is around $2000 billion, nearly as big as the figure for international trade. In size – the amount of outstanding issues held by investors – it is nearly 50 per cent bigger than the UK's bond market, including gilts. Interest rate swaps now amount to about $300 billion; currency swaps probably to $50 billion, and the figure for open positions in futures and options has been variously put between $450 and $700 billion. There can be doubts about the accuracy of some of these figures, but what they do point to is the upheaval of growth in the monetary markets in the late 1970s, the great bulk of which can have had little or nothing to do with developments in trade.

Indeed, one can with accuracy go back to a remark about the relationship between money and trade made earlier in this chapter, and now turn it on its head. It was argued earlier that in the 1960s and 1970s the global financial system existed to accommodate the real economy of trade in goods and services, and investment in plant and factories. For much of the period the financial indicators we observed – exchange rates, interest rates, and so on – responded to basic economic developments rather than acting as determining factors themselves. In the current world of international trade and money, to a very large degree this has been reversed. The financial markets are now so large in relation to the 'real markets' that the latter tend to respond to the former. Instead of basic economic transactions being expressed in money, money often determines economic trends and developments. Putting it another way, whereas in the 1960s and 1970 financial conditions were dependent on economic developments, now they have largely become independent variables which can often determine the current level, and the likely future growth rate of real international activity in both trade and investment. The causes of this reversal have already been chronicled – it remains merely to comment further on their nature and scope, and on their impact on the ways in which international traders and investors are likely to finance themselves.

This 'new financial system' has been more than adequately commented on by others, and I am sure will be much discussed in the next

few years. For the purposes of this study, however, its salient features can be said to be:

(a) Interest rates have risen to much higher levels than in the past. In recent years *real* interest rates have been as high as 8 per cent, and certainly over 4–5 per cent, significantly higher than the 1–2 per cent level of the 1960s and 1970s. With so much money sloshing around in the system, much of it unrelated to real transactions, one might ask why rates have moved so high. At first, inflationary expectations and the emergence of new financial markets may well have been causes but, of late, high interest rates have not been in anticipation of inflation, which has been falling, but rather a product of volatile markets, poor credit quality and short termism on the part of lenders.

(b) Outstanding debt is at much higher levels than either growing economies or expanding trade would suggest. Total debt in the USA, for example, has been calculated at approximately $1500 billion in 1970 and $7000 billion plus by 1984 – the latter figure being over twice as large as GNP (and it did not include futures, options, swaps, and lines of credit not being used!). These figures imply a growth rate of about 12 per cent per annum – the US economy grew at about 2½ per cent per annum *real* during the same period. Many reasons have been given for this phenomenon, but more important from our point of view are some of its effects. The rapid growth in debt helped to induce weakness in many financial institutions, as their capital became inadequate to support their expanding asset portfolios; at the same time, the clearly differentiated sectors of the financial system – banks, insurance companies, brokers, fund managers, and so on – became blurred, and gave way to composite financial institutions with inflation swollen balance sheets; and finally, we have already seen how the growth in debt also fuelled an explosion of derivatives in the form of futures, options, swaps, warrants and others.

(c) Both interest rates and exchange rates have become more volatile. To some extent this may have been due to the new monetarism which became popular after the inflationary outbreaks in the 1970s. After all, monetarism implies that interest rates should be allowed to climb to whatever levels are needed to control the money supply, and this tends to imply instability. Instability of interest rates in turn promotes the same phenomenon in exchange

rates. As a result the floating rate system has never produced that gradual progression of exchange rates towards competitive harmony and stability which was claimed for it at the outset.

(d) The banking system has become more international in scope, and at the same time more short term in its outlook. Cross-border bank to non-bank lending was just over $100 billion in 1973; by 1984 it had grown to nearly $750 billion. In between there had also been a massive growth in interbank lending for foreign exchange purposes, and of course the development of floating rate lending had caused increased interbank activity. By the early 1980s, deregulation was having its effects – commercial banks were beginning to try to compete with investment banks in the fields of securities trading and issues, and were showing signs that, in a deal-oriented world, short-term transactions were of more significance than long-term relationships.

By the time we get to the mid-1980s, the whole world of money and finance had become one in which the processes of deregulation, securitisation, innovation, and globalisation were in full swing, and dominating the scene. What did this mean, however, for the financing policies and systems which international traders had developed in the 1960s and 1970s? How did companies react and how are they continuing to cope with this new world of money and finance?

It is worthwhile leading into this question by looking, for a moment, at some of the detailed changes which have taken place in the last five years. Starting with international trade in its simplest form – exporting and importing – and the problem of dealing with foreign currencies on the part of the trader, in the early 1980s it was just not possible to transfer a currency exposure to New York, or to Tokyo, as London closed, for covering or hedging while the home team was resting; today that is possible, and is becoming common practice. In 1980–2, currency and interest rate swaps were in their infancy; today the markets in them are deep and fluid, and enable the trader to obtain short- to medium-term finance for his activities at costs below LIBOR. Similarly, five years ago the markets in options and futures were waiting to be explored by trading concerns; now they are widely used. Generally, in 1980–2 the short-term markets were much the same as they were in the 1970s; by 1987 the range of borrowing techniques had widened considerably, commercial paper was available in Euro-currencies as well as in dollars, medium-term notes were becoming widely used, and the big commercial banks were attempting

to become global in their operations. Finally, the longer term markets – in new bond issues, etc. – which were essentially still separated in the early 1980s, had by 1987 become much more global, and were providing finance in a range of 'new' currencies – yen, the ECU, Canadian, Australian and New Zealand dollars, French francs and the lira. In summary, financial markets had become deeper and wider and more integrated, and in many ways more efficient from the point of view of the borrower or user; but they had also become more complex, both in terms of the products provided and the intermediaries operating within them. And perhaps to some extent because of that, they had lost none of their uncertainty and volatility.

How have trading companies responded, in their financial systems, to these changes and challenges? Like the market operators, they have had to become more sophisticated, and develop expertise in a number of new areas. Deals have become much more complex, and also more competitive, and this has inevitably had an impact on relationships with banks. In fact, at one time they began to undergo a fundamental change, and there was much talk about a new transaction and deal-by-deal-oriented world, with the old pattern of long-term relationships between companies and banks receding into the background.

This is a view somewhat less prevalently held now, largely because the uncertainties and upheavals in the currency and money markets have induced a number of companies to seek continued special relationships with, and commitments from, small groups of major banks. Nevertheless, the very considerable variety, and complexity of new methods of hedging both currency and interest rate risks available today has meant that companies have had to be prepared to talk to a number of intermediaries in order to satisfy themselves that they are matching others in their performance. Above all, the back-office systems have had to be improved – complicated computer systems have had to be installed to analyse deals, record them, and track them through to the final accounting figures. These have proved invaluable from the point of view of control over the whole dealing operation. It has been vital to ensure that the system does not run away from itself, resulting in a plethora of transactions – forwards, futures, options and swaps – way out of line with the underlying company trading transactions. The corporate sector has not been without its own scandals, or at the very least, costly mistakes in this area, resulting in large losses in profit and loss accounts, and detailed control and checking systems are now quite common.

Even so, it is fair to say that so far the systems built up in the

1970s, and described above, have stood the test of the 1980s remarkably well. In particular, it has not proved necessary to change them fundamentally or to remould them completely. Generalised systems of 'centralised coordination' in both the short-term and long-term areas of finance have remained in place, and have been developed and extended in order to cope more effectively with the changes in the external world. In fact it would not be possible to use many of the new instruments – swaps, futures, options, and 'option products' such as FRAs – in highly decentralised and 'atomised' groups. In general, it can be said that the systems developed in the 1970s aimed at achieving three broad objectives – *control* over the company's cash flows and monetary risks, *mobility* and *flexibility* in the ways in which the company might cope with these risks. These objectives have remained at the forefront in the 1980s.

In some ways this continued predominance of the 'centralised' principle might seem at variance with other developments in the world of finance. In the USA, and to some extent in other countries, for example, the acquisition and divestment boom of the 1980s has resulted in the break up of large companies, 'sell-offs', 'spin-offs', and 'leveraged' and managerial 'buy-outs', and this has been widely regarded as the very opposite of centralisation. So has the widespread tendency to develop ways and means of cultivating entrepreneurship within large organisations, via options and other profit incentives, and so on. But the 'financial imperative' is a very strong one. International or global debt markets are best 'tapped' by a company on a centralised basis; so are global equity markets – witness the tendency to buy out public minorities in subsidiaries in order to emphasise the parent company's stock worldwide; and the instability and volatility of the currency and money markets continue to call for coherent and consistent policies and systems for the whole of the company or the group. And as we peer forward into the 1990s, and project that debt, because of 'securitisation' and tax inducements, is likely to continue to increase, that most of it, because of fears of inflation, is likely to be short and medium term rather than long, and expensive, and that one or more downturns may provoke significant adjustments as regards the poorer quality securities (the junk bonds?), and reregulation of, for example, the derivatives, we can only forecast that this imperative is likely to prevail. The influences of the international money system on trade and investment, which have been sketched above, are likely to remain just as powerful as in the recent past.

8 Money, Measurement and Accounting

Michael Bromwich and Christopher Noke

Accounting involves the expression of economic activity in monetary terms. Money as a unit of account provides a scale of measurement that makes it possible to assign numerals, according to the rules of accounting, to diverse activities and elements (such as sales, purchases, production, assets and liabilities) and enables us by mathematical manipulation to represent the results of these activities in terms of abstract concepts such as income and capital.

The eminent accounting historian, A. C. Littleton, has argued that the existence of money was one of seven prerequisites necessary for the evolution of systematic bookkeeping, 'since bookkeeping is unnecessary except as it reduces all transactions in properties or property rights to this common denominator'. Since the conventional accounting statements, the balance sheet and the profit and loss account, are founded on the workings of the double-entry bookkeeping system (which interlinks recorded changes in the resources and liabilities of the accounting entity), it follows that money is a prerequisite for systematic accounting as well as bookkeeping. The emphasis here must, however, be on 'systematic'. The ancient Greeks and Egyptians, for example, showed that it was quite feasible to maintain accounts in units other than money (grain, for example) or, indeed, without even employing a common unit of measurement. The building accounts of the Parthenon in 433–4 BC, though recording money amounts, happily recorded in one account receipts and payments made in Attican currency together with those made in the currencies of gold staters of Lampsacus and Cyzicus. Although the Parthenon account did not lend itself to addition or any other mathematical manipulation, none the less it served its purpose; that of recording all items that had passed through the hands of the accountant and for which he was accountable. These accounts served the purpose of control.

Systematic accounting goes beyond mere control. Although until

comparatively recently the principal purpose of annual reporting was regarded as 'accounting for stewardship' – whereby the directors reported on how the money raised by the company had been spent – modern accounting systems are now viewed primarily as systems for generating information for decision-making. Indeed, accounting has been described as 'the process of identifying, measuring and communicating economic information to permit informed judgement and decisions by users of the information' (American Accounting Association, 1966). Measurement is thus central to the accounting process, and in this chapter we examine some of the problems this raises.

We are concerned, however, only with economic information that can be measured and expressed in a common unit of account. But in any discussion of accounting and its expression in money terms, it is important to appreciate that the bottom line, with a £ sign by it, cannot tell the whole story.

We begin by illustrating some of the measurement problems associated with conventional accrual accounting – the method of accounting used, for example, for the annual financial accounts of limited companies. We then consider two major approaches to accounting theory which their proponents advance as a basis for resolving some of these problems – the 'economic income' approach and the 'user-needs' approach. Each of these introduces a new perspective on the relationship between money and accounting – the former by emphasising the concept of the time value of money, and the latter by stressing users' interest in forecasting future flows of money to and from the company. We then introduce the problem of inflation and discuss whether money is the most appropriate unit of measure in those circumstances. We are concerned throughout with the process of external corporate financial reporting, that is, with reporting the results of the company's activities to interested parties outside the company.

I ACCRUAL ACCOUNTING

There would be no measurement problems in accounting if accountants restricted their financial reporting to records of past cash flows to and from the company ('cash flow' being the term used by accountants for money receipts and payments) and counted the cash in hand at the end of the period. Cash is the only objective asset and cash flows the only economic event susceptible to exact measurement.

But reports of past cash movements are thought to provide inadequate information for users for decision-making purposes; in particular, they fail to match management efforts with its achievements. For example, money expended on producing stock for sale in future periods would appear as a cash outflow without any indication of the derived benefits in the future.

Instead of simply matching receipts and payments accrual accounting focuses on earnings, or income, by seeking to match in the profit and loss account revenues and expenses for the period irrespective of the date of receipt or payment of cash. It also recognises in the balance sheet that assets and liabilities are the results of events and activities which will have future cash consequences for the company. The process of matching or accruing thus implicitly incorporates in financial statements some information about management's expectations of the future.

But the accountant's 'measure' of assets and liabilities, based on expectations of the future, can only be a best estimate. The best the accountant can do is draw up a probability distribution of the likelihood of different states of the world occurring, in which, for example, varying percentages of debtors proved uncollectable, and choose some single point estimate, such as expected value, to report. It is true that the entire probability distribution could be reported. But if this were done for one item the accountant would logically have to report these distributions for all the assets and liabilities which are not subject to exact measurement, and (except under very unreal, static conditions) that is everything except cash. Such a policy of full disclosure would be expensive and its cost may well exceed its value, particularly if it led to information overload. Hence accrual accounting which focuses on just one number for each event, even though this will be subject to measurement errors, may be thought of as a cost-effective compromise between merely reporting cash flows and a policy of full disclosure.

The problems with conventional accrual accounting, however, stem not from measurement errors *per se*, but from the fact that the accrual process underlying the accounting system is ill-defined. Measurement has been defined 'in its broadest sense' as 'the assignment of numerals to objects or events according to rules' (Stevens, 1951), but the rules for assigning money symbols to accounting phenomena are not uniquely laid down. There is no unambiguous method of carrying out the matching of revenues and expenses and of establishing the amounts shown for assets and liabilities in the balance sheet.

Accountants have to choose some attribute or property of the event or object that they wish to measure and represent in the accounts; and as we will see, since some of these attributes may be thought to have no empirical existence outside the particular closed accounting world within which the accountant measures them, measurement may be arbitrary.

Traditionally, accountants have focused on 'historical cost' as the attribute to be measured: expenses are charged in the profit and loss account at the historical cost incurred; assets are represented in the balance sheet at their unexpired historical cost (i.e. the historical amount of the transaction that gave rise to them); revenues are recognised at the actual historical amount of proceeds received or receivable and liabilities at the amount incurred as the result of a past transaction. There are exceptions to this even within the general context of 'historical cost accounting'. For example, the concept of prudence requires the valuation of stocks at net realisable value if this is below cost. It is also common in the United Kingdom to incorporate a current market value for certain fixed assets, particularly land and buildings, but any increase in value would not be classed as 'income'; as part of the accountant's attempt to achieve the greatest possible objectivity in measurement, value changes are not recorded as 'income' until they are 'realised', normally by sale.

Later we touch on the possibility that attributes other than historical cost might be measured. For the moment, concentration on this attribute will illustrate the lack of definition inherent in accrual accounting. We will take the example of a company's stock in trade.

Measurement of the 'Cost' of Stock

Consider a company that trades in just one article. It buys 10 000 of them at a cost of £2 each and during the year sells 8000 of them. Accrual accounting requires us to match the 'cost' of the stock sold against the proceeds of sale (so as to measure the gross profit on sales) and to carry forward the 'cost' of the closing stock (assuming it to be above net realisable value) as a legacy to the next period where, when sold, it will be matched with the related revenue. In this simple case there are no problems. The goods sold cost £16 000 (8000 × £2) and closing stock cost £4000 (2000 × £2).

Consider next a similar company which makes two purchases of stocks during the year. It buys 10 000 units at £2 each and, following

a price rise, 10 000 of them at £3 each. It sells 15 000 units during the year. Again the accountant has to measure the 'cost' of goods sold and of closing stock. The measurement now depends upon the 'cost flow' assumptions made. Hypothetical assumptions must be made about the order in which the goods were sold; there is no pretence that these assumptions bear any resemblance to the actual physical flow of goods. Using the 'FIFO' ('first in, first out') convention, the accountant would assume that the earliest purchases were all sold first and would measure the cost of goods sold at (10 000 × £2) + (5000 × £3) = £35 000; closing stock would be reported at a cost of £15 000 (5000 × £3). Using the 'LIFO' ('last in, first out') convention, it would be assumed that the latest purchases were sold first and the cost of goods sold would be measured at (10 000 × £3) + (5000 × £2) = £40 000; closing stock would be reported at a cost of £10 000 (5000 × £2). Thus, in this simple case, there would be at least two possible measures of gross profit and of closing stock in the balance sheet.

Finally, consider a manufacturing company which purchases raw materials and processes them to produce finished goods for resale. To process the material it pays direct labour (labour directly involved with production) and direct expenses. The company also incurs various other expenses, including the wages of factory cleaners and maintenance staff, rent of the factory and offices, salaries and expenses of office staff and directors, selling and distribution expenses, depreciation of plant and machinery and of office equipment. As before, it is the job of the accountant to measure the attribute 'cost' of goods sold and 'cost' of closing stock.

Those outlays which form part of 'cost' must be identified. Some obviously do not. Selling and distribution expenses, for example, arise after production and do not enter into the cost of production. Raw materials, direct labour and direct expenses, on the other hand, do seem relevant. Direct labour and direct expenses are susceptible to fairly precise measurement, but ascertainment of the raw material content of total production involves the accountant, as before, in cost-flow assumptions about raw material consumption, so ambiguity in the figures is introduced at an early stage.

The accountant must also consider the other overheads. Some would argue that these overheads, whatever their nature, in so far as they do not vary with the level of production, should be expensed in the year in which they are incurred; they should be treated as a 'period' cost rather than a 'product' cost. Others would argue that

whilst this 'variable costing' approach is relevant for internal decision-making about, for example, levels of output, it does not show the 'cost' of closing stock and of goods sold; these figures must include all indirect overheads without which production would not have been possible. Factory cleaning and maintenance, factory rent, and depreciation of plant and machinery seem relevant; but what of the office rent, office salaries, directors' expenses, depreciation of office equipment, etc.? The answer is again not clear cut. For example, it might be argued that without office staff to process the order for raw materials on the office equipment, and to pay the wages of the production workers, there would be no production so that these overheads (and other similar ones) are a valid component of cost. And without a board of directors, complete with expense accounts, there would be no organisation to arrange the production. The question must be posed as to just how remote things have to be before it can be claimed that they do not form part of the cost of production.

Identifying relevant overheads is but one stage to their incorporation into the cost of stock. The next stage is determining how these overheads are to be allocated to units of production. Should this be on the basis of a predetermined absorption rate, itself based on assumptions about the 'normal' level of production, or on the basis of the actual production levels achieved? The problem of overhead allocation and absorption becomes yet greater when the company makes more than one product. Common overheads then have to be allocated between each product in order to measure separately the 'cost' of each product. But on what basis is that allocation to be made? It might, for example, be on the basis of labour hours or machine hours required by each product, or material cost, or some combination of these and others.

Nor can the measurement of the cost of manufactured stock be totally divorced from other measurement problems facing the accountant. As part of production overheads, for example, depreciation of factory plant and machinery is included. But how is that to be measured? In the context of traditional historical cost accounting, depreciation is simply a means of allocating the cost of a fixed asset over its useful economic life to the periods benefiting from its use; it is not an attempt to measure the fall in value over the period; it therefore follows that the 'net book value' appearing in the balance sheet, which equals cost less depreciation to date, is not a measure of the 'value' of the asset in any significant sense. However, the cost of a fixed asset can be allocated to time periods in one of several

different ways, by the straight line method or the reducing balance method, for example. Whichever method is used will produce a different charge for depreciation. The possibility of using different depreciation methods applies to all the fixed assets of the company, such as delivery vans and warehouses, not just those involved in production. But even looking at it purely from the point of view of measuring the 'cost' of manufactured stock, this clearly introduces a further degree of indeterminacy.

Thus, even when attempting to measure the same attribute with the same measuring unit – historical cost in money of account – different accountants may come up with different measurements, not because of errors but because the rules are not strictly defined. In fact, one may express doubt as to whether, in many cases, it is possible to talk properly of 'measurement' in relation, say, to costing stocks and fixed assets. Historical cost depreciation, for example, is one aspect of the attribute 'cost' that has no existence outside the closed accounting world; like the allocation of overheads it is purely arbitrary, and cannot be shown to be 'right' or 'wrong'. Even though it involves assigning a money symbol there is no empirical phenomenon (no object or event) to measure. And if we cannot be said to be 'measuring' cost in these cases, can we be said to be 'measuring' income, given that the income we purport to measure is dependent, *inter alia*, upon the amounts we report for closing stocks and depreciation?

Other Attributes

The lack of definition in accrual accounting becomes all the more flagrant once we admit of the possibility of measuring attributes other than historical cost. Recent years have seen increasing attention being paid to variants of current value as a more relevant attribute for accountants to measure – more relevant in the sense of having greater economic significance. But 'current value' is not one single attribute. It might be an 'entry value' – replacement cost. But replacement cost may mean the cost of replacing with an identical asset; or the cost of replacing with an asset best suited to the job done by the present asset; or the cost of replacing the service potential of the existing asset (which, in times of technological change, will differ from either of the above). Current value might be an 'exit' value – net realisable value. But if that attribute is selected what assumptions should be

made about the conditions of sale: an enforced liquidation or a leisurely sale in the normal course of business? Current value might be economic value or present value, which reflects its value in use. Or it might be a selective approach known as 'deprival value' (sometimes called 'value to the business') which chooses from among the other competing valuation techniques according to the circumstances, and is an attempt to measure the loss the company would suffer if it were deprived of an asset, assuming that it takes optimal action on deprival. The possibility of measuring these different attributes of accounting phenomena clearly increases the ambiguity of accrual accounting.

This ambiguity has not gone unrecognised, and accounting theorists have developed a number of approaches which either try to describe and explain the ambiguities, or else in an explicitly normative way try to determine which of the particular accrual methods (e.g. LIFO, FIFO) are in some sense 'best'. The earliest and most primitive approach to determining what is best, which is often categorised as the 'empirical inductive' approach, is based upon observation, generalisation and rationalisation of established practice. Such an approach has identified, for example, four 'fundamental accounting concepts' which are supposedly applied by accountants – the concepts of matching, prudence, going concern and consistency (between years and between companies). This approach becomes normative when accountants seek to ensure that accounting evolves in line with existing principles by applying these concepts. This type of approach seems often to have been implicitly adopted by standard-setting bodies when they have tried to reduce the ambiguity surrounding the treatment of specific issues.

The problem with the empirical inductive approach is that (since its observations are based on traditional historical cost accounting), it cannot provide guidance on which attributes of accounting phenomena should best be measured, nor can it satisfactorily resolve existing ambiguities. Recognising 'matching', for example, as a fundamental accounting concept does not help the accountant with the problem of how best to allocate depreciation, or the cost of goods sold, to match against revenues; straightline and reducing balance depreciation, LIFO and FIFO, and so on, all 'match' costs and revenues within the existing framework. The Accounting Standards Committee has attempted to lay down rules for the measurement of the 'actual cost' of stock, but while this reduces the variety of accounting practices there is at present no way of showing that these are 'best' practices.

Theorists who do not want to be constrained by the existing framework have often adopted a deductive approach to accounting theory, whereby accounting practices have been evaluated (or proposed) in terms of their ability to achieve some explicit goal. One of the major approaches of this kind is the 'true income' or 'economic income' approach which is based on the view that since accountants are trying to report economic events, the measures of income and capital they use should closely approximate those of the economists from whom they derived those concepts. Under this approach, those attributes and measures of accrual accounting which result in a reported profit closest to a measure of economic income would be deemed best.

II ECONOMIC INCOME AND THE TIME VALUE OF MONEY

The definition of income most often quoted by accountants of the economic income school is that of Sir John Hicks, who defined a man's income as 'the maximum value which he can consume during a week and still expect to be as well off at the end of the week as he was at the beginning'. The Sandilands Committee on Inflation Accounting, which based its proposals on an attempt to implement this central concept of income, redefined it for a company as: 'The maximum value which the company can distribute during the year and still expect to be as well off at the end of the year as it was at the beginning.' Hicks's definition makes income an increase in 'welloffness'. To make this operational it is necessary to define what is meant by 'welloffness'. Hicks attempts a number of 'approximations' to the central concept. It is his 'number 1' approximation that is normally regarded as a measure of economic income and the only one we will consider here: 'Income number 1 is the maximum amount which can be spent during a period if there is to be an expectation of maintaining intact the capital value of prospective money receipts.' For a company we may again substitute 'distribute' for 'spend'.

With this measure, the 'welloffness' to be maintained (referred to by accountants as the 'capital maintenance concept') is the present value of future money receipts, and income is the change in present value between two periods of time. Income is thus made very definitely a function of money, but in this case it is a function of

future money in exchange (the actual cash flow expected) rather than being the result of manipulating money of account recording past transactions. Moreover, it is a function of the time value of money; income and welloffness are connected directly through the rate of interest which measures preferences between consumption now and consumption in future time periods.

The concept of the time value of money makes use of the simple fact that cash which is receivable in one year's time is worth less than the same amount of cash receivable now. Similarly, cash receivable in two year's time is worth less than the same amount receivable in one year's time. The value – in present days terms – of the amount of cash receivable diminishes as the date of its receipt becomes more and more remote, reflecting the fact that cash receivable now has an opportunity value – the interest receivable on it by lending at the prevailing rate of interest.

The present value of the amounts receivable at future times may be found by the process of discounting. This process is familiar to businessmen who use the 'net present value' rule in their investment decisions, whereby the outlay on a project is compared with the periodic receipts, allowance being made for the time value of money through the discount rate. One of the advantages of the economic income approach is that it thus tries to relate financial reporting to financial decision-making. Assuming a rate of interest r, the present value at time 0 (PV_0) of a sum of cash D receivable at the end of each year for n years may be found as follows:

$$PV_0 = \frac{D_1}{(1+r)} + \frac{D_2}{(1+r)^2} + \frac{D_3}{(1+r)^3} \cdots \frac{D_n}{(1+r)^n}$$

or more succinctly

$$PV_0 = \sum_{t=1}^{n} \frac{Dt}{(1+r)^t}$$

Using this notation, the Hicks number 1 model may be expressed therefore as:

$$Y_0^1 = D_1 + \sum_{t=2}^{n} \frac{Dt}{(1+r)^{t-1}} - \sum_{t=1}^{n} \frac{Dt}{(1+r)^t}$$

where Y_0^1 signifies income for the first period.

We should note that the economic income measure as so far defined is an *ex ante* concept, that is, it is a forward-looking concept, a measure of income for the period from time 0 to time 1 made at time 0. Hicks rightly emphasises that for decision-making purposes it is the *ex ante* measure that matters. However, although financial directors make budgets for the company, and financial analysts make profit forecasts, financial accountants are accustomed to measuring profit at the end of the period rather than at the start, in the light of the events that have actually occurred. If accounting theory based on the economic income approach is not to demand a complete transformation of traditional accrual accounting, it must be able to cope with *ex post* measures, that is, measures of income made at the end of the period.

It can do this, by incorporating the actual cash flow that occurred during the period rather than the expected cash flow for the period, and by recalculating the closing capital value in the light of expectations then prevailing rather than in the light of the expectations at the start of the year. However, *ex post* measures can only be made at the expense (in times of uncertainty) of introducing some ambiguity into the measurement, so that the very concept of economic income itself becomes a little fuzzy. The problem is this: once it is perceived that the original expectations were wrong, should we recalculate the opening capital value in the light of the revised knowledge and expectations that we have at the end of the period? This process would exclude from income 'windfall' gains and losses arising from deviations from, and changes in, expectations. If we left the opening capital at its original amount, the windfalls would be included in income.

Leaving aside that particular problem, how might this model of economic income help resolve the dilemmas of accrual accounting? If the economic income model is accepted as an ideal, then measuring the assets and liabilities of the company in terms of the present value of their associated cash flows, and ascertaining the periodic change in present values, would give the income measure desired. But that is not feasible. It would ideally require management to forecast the amount and timing of future cash flows for very many years ahead – in principle into perpetuity. But as we do not possess the techniques for doing that with any degree of accuracy, the exercise is likely to flounder in subjectivity and bias (even though, as noted above, similar forecasts are required as part of the capital budgeting exercise). Moreover, if we wished to maintain the asset-by-asset approach of

conventional accrual accounting, problems would arise in ascribing cash flows to interdependent assets; in making arbitrary allocations we may end up with a system as ill-defined as the existing one.

As an alternative to using present values based on management forecasts, we might therefore seek some attributes of accounting phenomena which are more easily measured and which provide a suitable surrogate for present values of future cash flows. *A priori*, historical cost is unlikely to be such a surrogate. Current market prices may, however, at first sight appear more promising. In a perfect market the price of an asset will equal the market's expectations of future cash flows to the asset discounted at the expected rate of interest, and even if the market were slightly imperfect the market price would still at least partially reflect these factors. Using, in accounting reports, market values obtained in a very good market, and the realised cash flows of the period, would thus provide a measure of accounting income very close to that of *ex post* economic income (provided accountants were prepared to include all increases and decreases in market value in income, that is, to sacrifice their adherence to the realisation principle). However, markets of this type are likely to exist only for the assets of 'value based' companies such as investment or property companies.

Manufacturing companies often possess highly specific assets – specific to a particular process or to a particular company – so that market prices do not perfectly reflect the present value of future cash flows. There may therefore be wide discrepancies between asset replacement cost and net realisable value in the market and the problem would then be which market price to choose. In some cases, markets may be so imperfect as to be incomplete – there may be no market for some of the company's assets or output. Research and development expenditure, and goodwill, if they are regarded as assets, are rarely if ever traded as separable commodities; there may be a very restricted market, too, for trade debtors of the company, with factors demanding a high rate of discount.

Acceptable surrogates for economic measures are therefore unlikely to be found outside a world of perfectly competitive markets in equilibrium. The deprival value approach is sometimes justified on the grounds that, if not an exact surrogate, at least it produces figures representing the minimum value of future cash flows (because net realisable value and economic value in use measure the present value of future cash flows, and replacement cost cannot be greater than the present value of future cash flows otherwise replacement would

not be justified). However, the difference between 'minimum value' and 'actual value' of future cash flows might be considerable.

Moreover, once we admit the existence of imperfect and incomplete markets, the concept of economic income itself becomes muddied. The existence of incomplete markets means that some goods cannot be traded, and therefore cannot be unambiguously priced. Consider two mutually exclusive production plans for a company. We could compute the Hicksian economic income for each plan, by discounting cash flows, and the one with the higher income number would be preferred. However, because of market incompleteness each plan may also be associated with some non-marketable inputs and outputs, which may be valued differently by different shareholders in the company. Neither plan may therefore be unanimously preferred and the contribution to the 'welloffness' of the shareholders, which now becomes a matter of personal utility, cannot be objectively measured. Income has ceased to be well defined as it was in a setting of perfect and complete markets. This issue is further explored in Beaver and Demski (1979) and Bromwich and Wells (1983). Market imperfections such as the existence of multiple interest rates cause problems too, because it is not clear in those circumstances which rate of interest should be used to discount the cash flows to produce the measure of economic income.

The conclusion of this brief discussion of economic income is that, except in the unreal world of perfect and complete markets in full equilibrium (or in the case of a completed project where we can measure the cash in and the cash out so that accounting income and economic income would coincide), we are unlikely to be able to achieve a measure of economic income in our accounts. Advocates of these types of theory might still suggest that we should seek to judge accounting practices by how closely they approximate a measure of economic income. Others, including many practitioners, would argue that economic income and accounting income are very different concepts, and we should justify (and better define) the latter by appeal to some different criteria.

III A USER-NEEDS APPROACH TO ACCOUNTING THEORY

The goal of a user-needs theory is that financial reports should satisfy the informational needs of users for decision-making. Subsumed

under the broad heading of 'user needs' are two quite different approaches, the 'decision-maker' approach and the 'decision model' approach. The former concentrates on users' reactions to different types of accounting information measured by studies, for example, of security price changes; it embodies, therefore, a great deal of empirical market-based research. The latter concentrates on the inputs needed for users' decision models, to see which accounting information best meets the needs of users of these models. Although there will clearly be overlaps between these approaches, we will concentrate on the 'decision model' approach since, as generally developed, it highlights another aspect of the relationship between money and accounting.

This type of approach involves specifying as a first step the objectives of financial reporting and identifying the user groups whose needs we are trying to meet. At an applied level, this has been most fully developed by the Financial Accounting Standards Board in the United States in its Conceptual Framework project (which is an attempt to set out a coherent system of objectives and underlying concepts of financial accounting as the basis for the development of consistent financial reporting standards). As part of that project (FASB, 1978), the Board has identified as potential users of financial statements:

> owners, lenders, suppliers, potential investors and creditors, employees, management, directors, customers, financial analysts and advisors, brokers, underwriters, stock exchanges, lawyers, economists, taxing authorities, regulatory authorities, legislators, financial press and reporting agencies, labour unions, trade associations, business researchers, teachers and students, and the public.

FASB notes that all these groups share a common interest in the ability of a company to generate positive cash flows in the future. Investors, for example, are interested in the return from their investment through dividends and changes in market prices, and this depends on the ability of the company to generate cash; employees, similarly, are interested in the ability of the company to pay wages and salaries, and to provide job security, and this too is a function of favourable future cash flows. Whilst recognising this common interest, FASB (like most proponents of user needs) decided to focus specifically on the decisions and informational needs of investors and creditors because their decisions significantly affect the allocation of

society's resources. The board came to the conclusion that the primary purpose of financial reports was to 'provide information to help investors, creditors and others assess the amounts, timing and uncertainty of prospective net cash inflows to the related enterprise'.

By focusing on investors, FASB clearly sees accounting as serving what Keynes identified as the 'speculative motive' for holding money, by helping investors to predict the return to different investments. A new perspective on the relationship between money and accounting (as a predictor of future money flows) is established by the requirement for accountants to provide information to enable the assessment of future cash flows.

One of the problems of this approach is the possibility of conflict between user needs, for example, between present and potential investors, or between shareholders and employees. In choosing to satisfy some needs rather than others it may be necessary to make a social choice. Focusing on the needs of investors and creditors is just such a social choice.

Some accounting theorists would deduce from the objective specified by FASB that, far from trying to give better definition to accrual accounting, we should instead replace accrual accounting by a system of cash flow reporting that incorporates not merely reports of past cash flows but also forecasts of future cash flows for, say, the next five years. This would directly provide the information required, incorporating management's estimates of the future (since they are presumably best placed to make these forecasts). It would also avoid the problems of allocation inherent in accrual accounting. Focusing on the ability of companies to generate cash would highlight the similarities and differences between different companies which might presently be obscured by the variety of permissible accounting practices. Clearly, there would be problems with such reports; they would be highly subjective, difficult to audit and subject to potential bias in the same way as any attempt to measure economic income. Users might be able to correct for systematic bias but the presence of unsystematic bias could cause them loss.

FASB does not espouse cash flow reporting. It takes the view that, rather than making the predictions which cash flow reporting requires, accountants should instead provide the data to enable users to make their own predictions. FASB came to the conclusion that an interest in future cash flows is best served by providing information about earnings based on accrual accounting.

It is a common feature of theories based on user needs that they

incorporate a set of normative criteria about the characteristics of accounting information that will meet user needs. FASB's Concepts Statement on the *Qualitative Characteristics of Accounting Information* (1980) is an example; a hierarchy of qualities is drawn up, of which the relevance and reliability of information are given prominence. Relevance involves choice of the appropriate attribute of accounting phenomena that will help the decision-maker in decision-making. Reliability means that the user should be able to place confidence in the measurement of that attribute. There will often be a trade-off between these two characteristics, but recognising this does not tell us how to make the trade-off.

Relevance to user needs seems an obvious test of accounting practices. But recognising this has not so far conclusively determined whether, say, within the historical cost framework, LIFO, FIFO, straight line or reducing balance depreciation have greater relevance, or whether, as an attribute to be measured, historical cost or current value is to be preferred. Certainly, FASB's own findings on measurement and recognition in financial statements (1984) have failed (despite spending many millions of dollars) to resolve the issues satisfactorily, and contrasting the views of cash flow accountants with those of FASB shows that agreement on the end does not signify agreement on the means. Moreover, once we admit that there may be diversity of user needs beyond some common interest, we may be led to the conclusion that rather than trying to make accrual accounting 'all purpose' by focusing on just one attribute, we should recognise the possible relevance of 'different incomes for different purposes'. A measure of income for, say, taxation purposes, or for the purposes of dividend distribution consistent with the law's concern for creditor protection, might require income measurement on as objective a basis as possible (which will normally be historical cost, despite its limitations), while assessment of managerial performance might require the computation of income on some current value basis. It might be, too, that a balance sheet drawn up on the basis of net realisable values would be relevant to creditors looking to the cover for their claims, and to investors concerned with the funds available for investment elsewhere, irrespective of the basis of income measurement. But since user needs may never be fully satisfied the problem then would be knowing where to stop.

IV INFLATION AND MONEY AS A UNIT OF ACCOUNT

As we noted earlier, conventional accounting measures the attribute 'historical cost', using the money unit of account as the measuring rod, and income is similarly measured after charging for the historical cost of items sold or consumed. We have already observed the possibility of measuring some variant of current value. The case for selecting a current value is particularly strong during times of specific price change; that is, when the prices of the goods and services that the company owns or deals in are changing. This can occur whatever is happening to the general level of prices. When specific prices are changing a current value is *a priori* more likely to convey relevant information about future cash flows than is historical cost (which will give little information about the cash flows to be earned by the asset or the cash flows involved in replacing it). Ratios such as return on capital employed will give a better picture of comparative profitability when based on current values for capital employed. Moreover, those who are interested in the ability of the company to maintain its existing level of operations (what is called 'physical capital maintenance', which regards as income the surplus after providing for the maintenance of the operating capability of the company) will be interested in an operating profit figure arrived at after charging replacement cost, rather than historical cost, for items sold and consumed during the year.

We illustrate, using a simple example, a current value accounting (CVA) system in column 2 of Table 8.1. The current value is deemed to be replacement cost. The cost of sales is measured at replacement cost at the date of sale (60 × £1250), and the stock held at the year end is reported in the balance sheet at its replacement cost then (20 × £1250). The differences between these replacement costs and the corresponding historical costs represent 'holding gains' and in this example (following the concept of financial capital maintenance, which regards as income the difference between the opening shareholders' interest in money terms and the closing shareholders' interest in money terms) these holdings gains are included in the income statement. Money of account is the measuring unit. This system allows for the price changes affecting the specific assets of the company.

Despite the strength of the arguments in favour of some form of current value accounting, many businessmen resist it because of the

Table 8.1 A company was incorporated on 1 January with a share capital of £100 000 in cash. It immediately bought 80 units of stock at £1000 each; the remaining £20 000 was held as cash until the year end. It sold 60 units of stock for cash at £1400 each on 31 December. The company incurred no expenses and paid no dividends during the year. During the year to 31 December the general price index rose steadily from 100 to 110. On 31 December the replacement cost of stock was £1250 per unit.

Profit and Loss Accounts for Year 1	*1* *HC*	*2* *CVA*	*3* *CPP*	*4* *CVA/CPP*
	£	£	£C	£C
Sales	84 000	84 000	84 000	84 000
Cost of sales	60 000	75 000	66 000	75 000
Operating profit	24 000	9 000	18 000	9 000
Loss on holding money			2 000	2 000
Holding gain on stock sold		15 000		
Real holding gain on stock sold				9 000
Holding gain on stock held at year end		5 000		
Real holding gain on stock held at year end				3 000
Profit for year	24 000	29 000	16 000	19 000
Balance sheets at 31 December				
	£	£	£C	£C
Cash	104 000	104 000	104 000	104 000
Stock	20 000	25 000	22 000	25 000
	124 000	129 000	126 000	129 000
Share capital	100 000	100 000	110 000	110 000
Profit for year	24 000	29 000	16 000	19 000
Shareholders' interest	124 000	129 000	126 000	129 000

practical difficulties of measuring current value. When there were requirements for current cost information in the United Kingdom and the United States, measurement problems arose in many areas – for example, in ascertaining the replacement cost of growing timber, of oil and gas reserves, or of a ship whose cost depends on the level

of government subsidy. As the film industry asked, what is the replacement cost of *Gone with the Wind*? Some of these problems, and the other issues covered in this section, are further discussed in Noke (1985).

Issues quite distinct from those raised by special price changes are raised by conditions of inflation, that is, changes in the general level of prices. During times of inflation the monetary unit loses value in terms of its command over a basket of commodities; people who hold money become worse off in real terms.

A charge frequently levelled at historical cost accounting measured in money in times of inflation is the non-additivity of sums incurred at different dates. The balance sheet of a company at 30 June 1988 might, for example, include:

Factory (purchased 1960) cost	£100 000
Plant (purchased 1978–88) cost	£150 000
Stocks (purchased March 1988)	£100 000
	£350 000

Some accountants question whether we can validly sum these amounts. They would say, 'No; adding pounds of different dates is like adding together cats and dogs, or apples and bananas.' That line of argument, however, seems to be off the point. We can add together cats and dogs if we call them 'animals' no matter whether one is a kitten and one an old cur, and we can add together apples and bananas if we label them pieces of fruit irrespective of their degree of freshness. With this view a pound is a pound, irrespective of the date stamped on it (as a miser counting his coins would testify). The 'non-additivity' argument seems to be confusing money as a unit of account with money as a measure of value, or money in exchange.

The real question concerns not the additivity of the data, but whether the resultant total is useful to users of accounts. It is obvious that £100 000 in 1960 was *worth more* than £100 000 in 1988. The purchase of the factory represented the sacrifice by shareholders of £100 000 of 1960 purchasing power, and to appreciate the significance of that in 1988 it is necessary to express its equivalence in terms of 1988 purchasing power. Similarly, the historical cost profit and loss account will include depreciation of the plant and machinery based on its cost between 1978 and 1988; but to appreciate the significance of the sacrifice of purchasing power made by investment in that plant,

it is necessary to express its cost and depreciation in terms of 1988 purchasing power.

A method of accounting known as Current (or 'Constant') Purchasing Power accounting (CPP) has evolved to do just this. Instead of measuring in money units of account it measures in purchasing power units of account, usually the purchasing power of a pound (or dollar, or whatever) at the balance sheet date. All items in the profit and loss account are restated by the movement in a general index of prices since the date of the transaction to the balance sheet date; and in the balance sheet, all assets and liabilities are restated by the movement in the general index since the date of acquisition, except for those assets and liabilities which are fixed in monetary terms (cash, debtors, overdrafts, loan capital, for example) which remain at their monetary amount. Leaving monetary items at their face value throws up those losses in purchasing power familiar to anyone who holds cash in current accounts during times of inflation; perhaps less familiar, but just as logically, it highlights the gain in purchasing power arising as a result of repaying liabilities in depreciated pounds.

Column 3 of Table 8.1 illustrates CPP accounting, when the company in the example faces a 10 per cent change in the general price level. The symbol £C signifies constant £s. The cost of sales figure is arrived at by restating the historical amount by the movement in the general index (£60 000 $\times \frac{110}{100}$); the loss on cash represents the loss in purchasing power on the £20 000 held all year (£20 000 $\times \frac{110-100}{100}$); the balance sheet figure for stock is historical cost times the movement in the general index since it was bought (£20 000 $\times \frac{110}{100}$); and the share capital represents the amount shown in the historical cost accounts restated by the movement in the general index during the year (£100 000 $\times \frac{110}{100}$). Had the sales not taken place at the year end, that figure too would have been updated from the date it took place to the year end.

CPP accounting was the favoured system of both the Accounting Standards Committee in the United Kingdom and the Accounting Principles Board/Financial Accounting Standards Board in the United States when, in the early 1970s, they turned their attention to the problems of inflation accounting. The ASC was deflected from its path by the government-appointed Sandilands Committee which, in 1975, recommended a system of current cost accounting based on physical capital maintenance (with no general index adjustments) as the means of accounting for inflation, and FASB was partially deflected when the SEC in the United States started demanding

replacement cost information from companies. But in any case, not all accountants – and certainly not all businessmen – accepted the case for CPP accounting, for two main reasons.

First, it was argued that the system produces balance sheet figures that may bear no relationship to current values, and charges for items sold or consumed that do not reflect the current value of the sacrifice made. While this is true, this line of attack is wide of the target. CPP accounting does not claim to change the attribute being measured but to change the measuring unit to one more meaningful than money. We are still measuring historical cost, but in terms of current purchasing power. This attack can also be repulsed because, as we illustrate later, it is possible to incorporate CPP adjustments into any type of accounting system, including one where the attribute is 'current value'.

The second type of criticism of CPP accounting – whether used in combination with historical costs or current values – centres around the use of a general index of prices. Several issues have been raised. Sandilands, for example, argued that the concepts of general price change and of the general purchasing power of money were unquantifiable. However, whilst it is true that there are statistical problems in calculating a general index – what is to be included, for example, how the items should be weighted, or whether, say a Paasche or Laspeyres type index is used – people know that inflation exists and any attempt to ignore it is unreal. A slightly different line of argument points out that the indices available (such as the Retail Price Index) are not representative of the patterns of expenditure of different shareholders and are therefore of little relevance. But we may regard the RPI as a useful indication of the way things are going; it is used, for example, to uprate pensions and some social security benefits; and in any case empirical studies of the 'heterogeneity hypothesis' have suggested that, whatever their pattern of expenditure, the rate of inflation that different social/income groups actually face is very similar. A third variant of the attack on the use of a general index maintains that a general price index (based on jam, vegetables, videos, curtains, and the like) is irrelevant to adjusting the accounts of companies which spend their money on chemical plants, heavy engineering plants and office machinery. This objection, however, merely serves to highlight a question of indeterminacy in the whole process of financial reporting: whose income are we measuring in a set of company accounts, the shareholders' income or that of the company as a separate entity? If income is regarded – as seems most

sensible – as a change in the shareholders' interest in the company, as an alteration in their welloffness as a result of their investment, the general index does seem relevant since shareholders do purchase jam, videos, curtains, and so on.

The logical case for adjusting accounts for inflation by a method of CPP accounting therefore seems strong. The case for incorporating current values is also strong. However, as noted above, it is possible to apply CPP adjustments to a set of accounts based on current values. This is illustrated in column 4 of Table 8.1 where cost of sales is charged at replacement cost and the stock in the balance sheet is shown at year-end replacement cost. The loss on holding money is calculated as previously under the CPP system, but this system introduces two new elements of income, the real holdings gains on stock sold and on stock held at the year end. These real holding gains measure the extent to which the replacement cost of the stock exceeds its historical cost updated by the general index (£75 000 – £66 000) and (£25 000 – £22 000), that is, it shows how far the company has invested in assets which have 'beaten inflation'. This combination of current values with CPP adjustments is sometimes referred to as 'real terms accounting'. It shows the profit made by shareholders after maintaining the opening equity interest in real terms.

But the combination of current value and CPP accounting may be thought too complicated for practical use. A second best solution therefore might be to retain the money unit of account, but to make an appropriation out of profit to a capital maintenance reserve of a sum required to maintain the opening equity interest in *real* terms. Thus, continuing our example, profit on a current value basis, including money holding gains, amounts to £29 000; we could then make a transfer to reserve of £10 000 to compensate for effect of inflation on the opening capital.

At present, neither the United Kingdom nor the United States has any requirement for the production of accounts on either a current value or an inflation-adjusted basis. It is suggested by many that with current rates of inflation there is no necessity for departing from historical cost accounts. But that logic is false. First, it confuses inflation with specific price change; there may be considerable changes in the prices of specific assets even if the rate of inflation is zero, and as we suggest above in these circumstances current values are likely to convey greater information than historical cost. And secondly, even if inflation has slowed down, the cumulative effect of price change over a number of years on long life fixed assets and on original

subscribed capital could be considerable; if investors are interested in the maintenance of real capital it would be misleading to ignore this effect.

V SUMMARY

In this chapter we have examined some of the problems of expressing economic activity in monetary terms. We have seen that many of the measurement problems arise because the rules of accrual accounting are ill-defined, and attempts to resolve the problem through the development of accounting theories have not so far been very successful. Changing prices introduce further problems; should we change the focus of our measurement from historical cost to current value (with the subjectivity that may involve in a practical setting) and should we change the measuring unit from that of money to units of purchasing power?

The two types of accounting theory that we touched upon – economic income and user needs – illustrated further aspects of the relationship between money and accounting. However, as noted earlier, it must not be expected that all information of interest about the company can be expressed in money terms. User groups interested in the future of the company for investment decisions, for example, will perhaps be interested in the state of industrial relations or the health of the Chief Executive. These may well affect the ability of the company to generate future cash flows, but they do not lend themselves to expression in money terms. Hence the important *caveat* that accountants cannot encapsulate the whole of economic activity in a single figure.

References

American Accounting Association (1966) *A Statement of Basic Accounting Theory* (Sarasota).

Beaver, W. and J. Demski (1979) 'The Nature of Income Measurement', *Accounting Review* (January).

Bromwich, M. and M. C. Wells (1983) 'The Usefulness of a Measure of Wealth', *Abacus* (December).

Financial Accounting Standards Board (1978) *Statement of Financial Accounting Concepts No. 1, Objectives of Financial Reporting by Business Enterprises* (Stamford, Conn.).

Financial Accounting Standards Board (1980) *Statement of Financial Accounting Concepts No. 2, Qualitative Characteristics of Accounting Information* (Stamford, Conn.).

Financial Accounting Standards Board (1984) *Statement of Financial Accounting Concepts No. 5, Recognition and Measurement in Financial Statements of Business Enterprises* (Stamford, Conn.).

Hicks, J. R. (1946) *Value and Capital*, 2nd edn (Oxford: Clarendon Press).

Inflation Accounting: Report of the Inflation Accounting Committee (1975) ('Sandilands') Cmnd 6225 (London: HMSO).

Littleton, A. C. (1933) *Accounting Evolution to 1900* (New York: American Institute).

Noke, C. W. (1985) 'Financial Reporting and Changing Prices – A Review of the Issues', *Omega* (May).

Stevens, S. S. (1951) 'Mathematics, Measurement and Psychophysics', in Stevens, S. S. (ed.) *Handbook of Experimental Psychology* (New York: John Wiley).

This volume contains a series of, sometimes controversial, reflections on money and the conduct of monetary policy by eight distinguished economic analysts and practitioners. The varied topics include: the development of monetary theory; a spirited defence of monetarism; a review of recent monetary and credit trends and their impact on the economy; a discussion of money and stagflation; an analysis of the nature and causes of recent structural change in the British financial system; a survey of the role of money in international trade; and a review of some accounting aspects of inflation.

David T. Llewellyn, the editor, is Professor of Money and Banking and Head of the Economics Department at Loughborough University. He has previously held appointments at HM Treasury, Nottingham University and the International Monetary Fund in Washington, DC. He is currently Consultant Economist to money-brokers Butler Harlow Veda in London. He is the author of *International Financial Integration, Evolution of the British Financial System, The Regulation and Supervision of Financial Institutions* and editor of *Framework of UK Monetary Policy.*

ISBN 0-333-48529-7