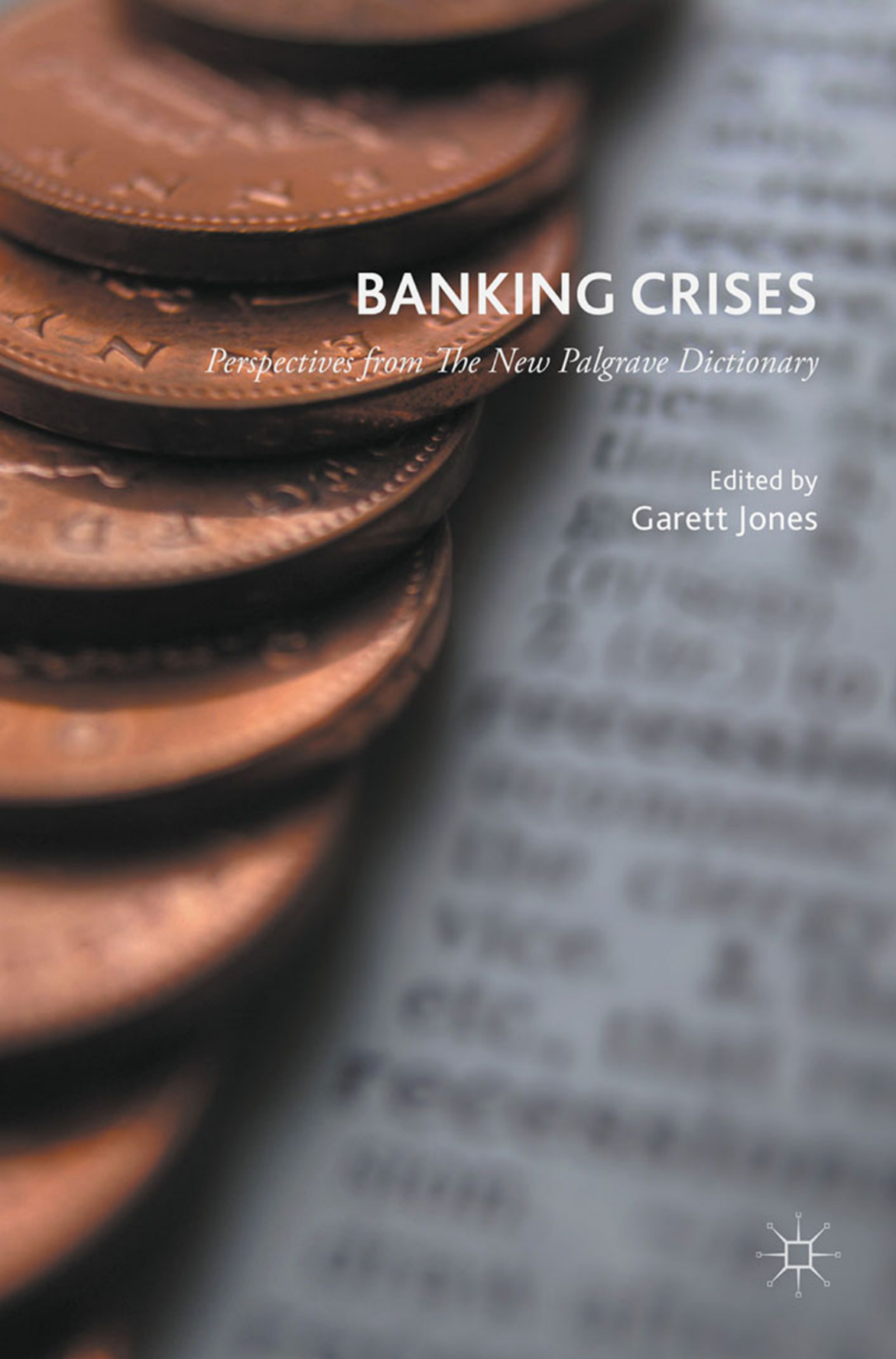


# BANKING CRISES

*Perspectives from The New Palgrave Dictionary*

Edited by
Garett Jones

# Banking Crises

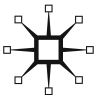
*This page intentionally left blank*

# Banking Crises

## Perspectives from *The New Palgrave Dictionary*

Edited by
Garett Jones
*George Mason University, USA*

palgrave
macmillan





All articles first published in *The New Palgrave Dictionary of Economics*, 2nd Edition
Edited by Steven N. Durlauf and Lawrence E. Blume
in eight volumes, 2008

or *The New Palgrave: A Dictionary of Economics*
Edited by John Eatwell, Murray Milgate and Peter Newman
in four volumes, 1987

Some entries in this volume have been updated since their original publication.



First published 2016 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited,
registered in England, company number 785998, of Houndmills, Basingstoke,
Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC,
175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has
companies and representatives throughout the world.



ISBN 978-1-349-55412-6 ISBN 978-1-137-55379-9 (eBook)
DOI 10.1057/9781137553799

This book is printed on paper suitable for recycling and made from fully managed and
sustained forest sources. Logging, pulping and manufacturing processes are expected to
conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Typeset by MPS Limited, Chennai, India.

# Contents

# List of Figures

# List of Tables

# Preface

Banking panics are nothing new: in rich and poor countries, the banking sector has long had a troubling tendency to collapse. In the aftermath workers lose jobs, investment collapses and the banking system tries to piece itself back together. Banking crises are a puzzle, and economists love nothing more than a hard puzzle.

This volume collects some remarkable articles from *The New Palgrave Dictionary of Economics* on banking crises, their historical background, the theories and laboratory experiments that have investigated their causes, and some policy recommendations that might make them less likely—or at least less destructive—in the future. About one third of the articles were written after 2008, and so aren't included in the print version of the most recent *New Palgrave*, although I should note that all are available through an online subscription to the Dictionary. Together with Palgrave editor Alison Howson, I commissioned many of these more recent articles in the wake of the global financial crisis. And through the efforts of Palgrave editor Rachel Sangster the present volume brings together many of those new articles plus more from the 2008 edition.

Allow me to point out three exceptional articles, all quite readable, that give a sense of the volume's scope:

1. A brief biography of Walter Bagehot, the Victorian-era editor of *The Economist* and author of one of the best and most enjoyable essays on banking crises and their cure, *Lombard Street*. The biography, written by the influential British historian and life peer Asa Briggs, along the way describes the early banking panics that have shaped how economists approach the topic.
2. Foote and Willen's article entitled "subprime mortgage crisis" is provocative and fact-driven. They argue the evidence doesn't easily fit the popular story of a financial crisis caused by Wall Street insiders. In that story, financial insiders lent money for mortgages that were "doomed to failure," and then sold off the resulting low-quality mortgage-backed securities to ignorant investors. Foote and Willen are both Federal Reserve economists; their article summarizes and synthesizes their own influential work on the topic, work that has substantially (though not entirely) shifted the narrative regarding the root causes of the crisis.
3. A thorough overview of Greece's recent economic crises by Nicos Christodoulakis, an economist and former Greek finance minister. This is just one sign of the volume's international scope, as well as a reminder that banking crises and debt crises are all too often simultaneous events.

We can break the articles into four categories: historical crises, the global financial crisis and aftermath, key economic institutions, and more theoretical, analytical pieces.

**Historical Financial Crises**
Bagehot, Walter
bubbles in history
gold standard
Great Depression
Great Depression, monetary and financial forces in
Kindleberger, Charles P.
Law, John
New Deal
South Sea bubble
tulipmania

**The Global Financial Crisis and Its Aftermath**
Credit Crunch Chronology
credit rating agencies
euro zone crisis 2010
Greek crisis in perspective
Lehman Brothers bankruptcy
quantitative easing by the major western central banks during the global financial crisis
Run on Northern Rock, the
shadow banking: a review of the literature
subprime mortgage crisis, the

**Key Financial Institutions**
Bank of England
banking industry
Federal Reserve System
International Monetary Fund
international monetary institutions

**Explanations and Solutions**
banking crises
bubbles
capital controls
currency crises
currency crises models
laboratory financial markets
liquidity trap
Minsky crisis
speculative bubbles

You'll note that the word "bubble" appears in the titles of four articles, and the concept is central to even more, including the article "laboratory financial markets," which summarizes the work of Nobel laureate Vernon Smith. Understanding financial market bubbles is central to understanding banking crises, and this volume offers a quite thorough and diverse survey of the topic.

Overall, the articles are intellectually serious without being heavily mathematical; in fact most contain no equations at all. While I am certainly not an unbiased observer, I believe it's a great volume for browsing, a good source of ideas and insights both to the undergraduate writing a term paper and the faculty member looking for inspiration. Banking crises are unlikely to vanish any time soon, so, alas, it's safe to forecast that this volume will be useful and relevant for decades to come.

Garett Jones, George Mason University

# List of Contributors


**TOBIAS ADRIAN**
Federal Reserve Bank of New York, USA

**CINZIA ALCIDI**
Centre for European Policy Studies, Brussels, Belgium

**ADAM B. ASHCRAFT**
Federal Reserve Bank of New York, USA

**J. ASHWORTH**
Morgan Stanley, UK

**STANLEY W. BLACK**
University of North Carolina, USA

**MICHAEL D. BORDO**
Rutgers University, USA

**ASA BRIGGS**
Honorary Fellow of Sidney Sussex College, Cambridge and Worcester College, Oxford, UK

**MARKUS K. BRUNNERMEIER**
Princeton University, USA

**CRAIG BURNSIDE**
Duke University, USA

**CHARLES W. CALOMIRIS**
Columbia University, USA

**SATYAJIT CHATTERJEE**
Federal Reserve Bank of Philadelphia, USA

**NICOS CHRISTODOULAKIS**
London School of Economics, UK

**P. DEAN CORBAE**
Wisconsin School of Business, USA

**W. MAX CORDEN**
Emeritus Professor, Johns Hopkins University, USA and Professorial Fellow, University of Melbourne, Australia

**GAUTI B. EGGERTSSON**
Brown University, USA

**MARTIN EICHENBAUM**
Northwestern University, USA

**THOMAS J. FITZPATRICK IV**
Federal Reserve Bank of Cleveland, USA

**PRICE V. FISHBACK**
University of Arizona, USA

**DARIO FOCARELLI**
Bank of Italy, Italy

**CHRISTOPHER L. FOOTE**
Federal Reserve Bank of Boston, USA

**KRISTIN J. FORBES**
MIT's Sloan School of Management, USA

**DANIEL FRIEDMAN**
University of California Santa Cruz, USA

**PETER GARBER**
National Bureau of Economic Research, USA

**CHARLES A. E. GOODHART**
London School of Economics, UK

**DANIEL GROS**
Centre for European Policy Studies, Brussels, Belgium

**DONALD D. HESTER**
Emeritus Professor, University of Wisconsin, USA

**BRETT HOUSE**
McGill University, USA

**MIGUEL A. IRAOLA**
Center for Economic Research, Mexico

**GRACIELA LAURA KAMINSKY**
George Washington University, USA

**CHARLES P. KINDLEBERGER**
Formerly Ford Professor of Economics, MIT, USA

**LAWRENCE H. OFFICER**
University of Illinois at Chicago, USA

**ALBERTO FRANCO POZZOLO**
University of Molise, Italy

**SERGIO REBELO**
Kellogg School of Management, Northwestern University, USA

**MANUEL S. SANTOS**
University of Miami School of Business Administration, USA

**JOEL SHAPIRO**
Saïd Business School, University of Oxford, UK

**PETER TEMIN**
Gray Professor Emeritus of Economics, MIT, USA

**JAMES B. THOMSON**
Duke University, USA

**DAVID VINES**
University of Oxford, UK

**HANS-JOACHIM VOTH**
Pompeu Fabra University, Spain

**PAUL S. WILLEN**
Federal Reserve Bank of Boston, USA

**L. RANDALL WRAY**
University of Missouri-Kansas City, USA

# Bagehot, Walter (1826–1877)

Editor and literary critic as well as banker and economist, Bagehot was described in retrospect by Lord Bryce as 'the most original mind of his generation' (Buchan, 1959, p. 260). It is a difficult claim to sustain, certainly as far as his scattered economic writings are concerned. There was no doubt, however, about his intellectual versatility: there was an immediacy, a clarity and an irony – what he said of his friend Arthur Hugh Clough's poems, 'a sort of truthful scepticism' – about Bagehot's essays in different fields which make them still pre-eminently readable. Bagehot saw connections, too, between economics, politics, psychology, anthropology and the natural sciences – 'mind and character' – refusing to draw rigid boundaries between most of these subjects and 'literary studies', while recognizing in his later years that the frontiers of political economy needed to be more carefully marked. 'Most original' or not, he was, as the historian G.M. Young (1948) has observed, *Victoranum maxime*, if not *Victoranum maximus*: 'he was in and of his age, and could have been of no other.' He pre-dated academic specialization and professionalization, and he was never didactic in his approach.

His first writing on economics, a revealing if not a searching review of John Stuart Mill's *Principles of Political Economy*, appeared in 1848 before the sense of a Victorian age had taken shape. His last and most voluminous writing on the subject appeared posthumously in a volume of essays, the first on 'the postulates of English political economy', which his editor-friend Richard Holt Hutton entitled *Economic Studies* (1879). By then the economic confidence of the mid-Victorian years was over, and there were many signs both of economic and social strain, some of which Bagehot had predicted. It was in 1859, the *annus mirabilis* of mid-Victorian England, however, the year of Darwin's *Origin of Species*, Mill's *On Liberty* and Smile's *Self Help*, that Bagehot became editor of *The Economist*, a periodical founded by his father-in-law James Wilson, and it was through his lively editorship, which continued until his death, that he was in regular touch with an interesting and influential, if limited, section of his contemporaries. 'The politics of the paper', he wrote simply, 'must be viewed mainly with reference to the tastes of men of business.'

The mid-Victorian years constituted, in his own phrase, 'a period singularly remarkable for its material progress, and almost marvellous in its banking development'. It was the latter aspect of the period which provided him with the theme of his best-known and brilliantly written book *Lombard Street*, which was begun in 1870 and appeared in 1873. It dealt, however, as it was bound to do, not only with the 'marvellous development', but with the 'panics' of 1857 and 1866 to which the Bank of England, the central institution in the system, had to respond. Indeed, the germ of *Lombard Street* was an article written in *The Economist* in 1857, 13 years after Peel's Bank Charter Act, and it was in 1866 that he took up the theme again.

Bagehot's conviction that the Bank of England neither fully understood nor fully lived up to its responsibilities was the product of years of experience which went back

to his own early life between 1852 and 1859 as a country banker with Stuckey's at Langport, his birthplace, in the West of England, where his father also was a banker. The chapter on deposit banking reflects this. So, too, does his complaint that the directors of the Bank of England were 'amateurs', and his insistence that the 'trained banking element' needed to be augmented.

*Lombard Street* is a book with a distinctive purpose rather than an essay in applied economics; and, as Schumpeter has observed, 'it does not contain anything that should have been new to any student of economics'. The main stress in it is on confidence as a necessary foundation of London's banking system. 'Credit – the disposition of one man to trust another – is singularly varying. In England after a great calamity, everybody is suspicious of everybody; as soon as that calamity is forgotten everybody again confides in everybody.' Bagehot underestimated the extent to which through joint stock banks' cheques trade was expanding without increases in note issue and the extent to which the Bank of England itself was beginning to develop techniques of influencing interest rates. He also overestimated the extent to which in 'rapidly growing districts' of the country 'almost any amount of money can be well employed'. In the last resort, too, his policy recommendations were deliberately restricted. He was disposed in principle to a 'natural system' in which each bank kept its own reserves of gold and legal tender, but in English circumstances he saw no more future in seeking to change the system fundamentally than in changing the political system. 'I propose to retain this system because I am quite sure that it is of no manner of use proposing to alter it.' With a characteristic glance across the Channel to France for a necessary comparison – things were done very differently there – he noted how the English system had 'slowly grown up' because it had 'suited itself to the course of business' and 'forced itself on the habits of men'. It would not be altered, therefore, 'because theorists disapprove of it, or because books are written against it'.

Bagehot had little use for 'theorists' and disdained the French for what he called their 'morbid appetite for exhaustive and original theories'. He described political economy 'as we have it in England' as 'the science of business' and did not object to the fact that it was 'insular'. Yet he talked of the 'laws of wealth' and believed that they had been arrived at in the same way as the 'laws of motion'. Free trade was such a law. It was impossible, he argued, to write the history of 'similar phenomena like those of Lombard Street' without 'a considerable accumulation of applicable doctrine': to do so would be like 'trying to explain the bursting of a boiler without knowing the theory of steam', a not very helpful analogy since the invention of the steam engine preceded the discovery of the laws of thermodynamics. Bagehot relied considerably on analogies. 'Panics', for example, were 'a species of neuralgia'. The 'unconscious "organization of capital"' in the City of London, described by Bagehot as a 'continental phrase', depended on the entry into City business of a 'dirty crowd of little men'; and this 'rough and vulgar structure of English commerce' was 'the secret of its life' because it contained 'the propensity to variation' which was 'the principle of progress' in the 'social as in the animal kingdom'.

Such an approach to political economy was radically different from that of W.S. Jevons who, like Bagehot, had been educated at University College, London, or 'M. Walras, of Lausanne' who, according to Bagehot himself, had worked out

'a mathematical theory' of political economy 'without communication and almost simultaneously'. There were however three defects, Bagehot maintained, in the British tradition of political economy, which started with Adam Smith but was sharpened and 'mapped' by David Ricardo. First, it was too culture-bound; for example, it took for granted the free circulation of labour, unknown in India. Second, its expositors did not always make it clear that they were dealing not with real men but with 'imaginary' ones. Abstract political economy did not focus on 'the entire man as we know him in fact, but . . . a man answering to pure definition from which all impairing and conflicting elements have been fined away'. It was not concerned with 'middle principles'. Third, considered as a body of knowledge, English political economy was 'not a questionable thing of unlimited extent but a most certain and useful thing of limited extent'. It was certainly not 'the highest study of the mind'. There were others 'which are much higher'.

Bagehot did not push such criticism far. He had much to say about primitive and pre-commercial economies, but he put forward no theory of economic development. Nor, despite an interest in methodology, did he draw out the full implications of his own behaviourist (and in places institutionalist) approach to economics. Finally, he offered no agenda for political economists in the future. He noted, as others noted, that during the 1870s political economy lay 'rather dead in the public mind. Not only does it not excite the same interest as it did formerly, but there is not exactly the same confidence in it.' His own precoccupations in that decade were more practical than theoretical despite the writing of such essays as 'The Postulates of English Political Economy', which first appeared in article form in the *Fortnightly* in 1876. He never completed a new essay on Mill, and an essay on Malthus, whom he took along with Smith, Ricardo and Mill to be the founders of British political economy, revealed more interest in the man than in his thought. In the year when the 'Postulates' appeared, he successfully suggested to the Chancellor of the Exchequer the value to the Treasury of short-term securities resembling as much as possible commercial bills of exchange. The result was the Treasury Bill. The fact that the Chancellor was then a Conservative mattered little to the liberal-conservative Bagehot, who was described by his Liberal admirer W.E. Gladstone as a 'sort of supplementary Chancellor of the Exchequer'.

Bagehot was as out of sympathy with the liberal radicals of the 1870s as he was with the bimetallists, and he had never shown any sympathy for socialist political economy. He saw the capitalist as 'the motive power in modern production' in the 'great commerce', the man who settled 'what goods shall be made, and what not'. Nonetheless, he stated explicitly in several places that he had 'no objection whatever to the aspiration of the workmen for more wages', and he came to appreciate more willingly than Jevons the role of trade unions and collective bargaining. In his first review of Mill in 1848 he had stated that 'the great problem for European and especially for English statesmen in the nineteenth century is how shall the [wage] rate be raised and how shall the lower orders be improved'. Some of the views he expressed on this subject – and on expectations – were not dissimilar to those of the neoclassical Alfred Marshall. He did not use the term 'classical' himself in charting the evolution of British political economy.

Bagehot left no school of disciples. He was content to persuade his contemporaries. His sinuous prose style was supremely persuasive. So, too, was his skill in sifting and

assessing inside economic intelligence. Yet while he devoted little attention to precise quantitative evidence in *Lombard Street* and, unlike Jevons, saw little point in developing economics in mathematical form, he was always interested in numbers as well as in words. One of his closest collaborators on the staff of *The Economist*, the statistician Robert Giffen, his first full-time assistant, paid tribute to 'his knowledge and feeling of the "how much" in dealing with the complex workings of economic tendencies'. 'He knew what tables could be made to say, and the value of simplicity in their construction.' Bagehot always maintained, however, that while 'theorists take a table of prices as facts settled by unalterable laws, a stockbroker will tell you such prices can be made'. Statistics were 'useful': they needed to be interpreted by 'men of business' who possessed the grasp of 'probabilities' and the 'solid judgement' which Bagehot most admired and which he sought to express. Indeed, business for him was 'really a profession often requiring for its practice quite as much knowledge, and quite as much skill, as law and medicine'. Businessmen did not go to political economy: political economy, as in the case of Ricardo, came to them.

ASA BRIGGS

### *Selected works*

All Bagehot's economic writings are collected in N. St. John Stevas, ed., *The Collected Works of Walter Bagehot*, vols 1–15 (1978–86), London: The Economist.

### *Bibliography*

Buchan, A. 1959. *The Spare Chancellor. The Life of Walter Bagehot*. London: Chatto & Windus.

Giffen, R. 1880. Bagehot as an economist. *The Fortnightly*, April, 549–67.

Young, G.M. 1948. The greatest Victorian. In G.M. Young, *Today and Yesterday*, London: Rupert Hart-Davis.

# Bank of England

The primary motivation for the establishment of the Bank of England was the need to raise funds to help the government finance the then current war against France, although the view had also developed that a bank could help to 'stabilize' financial activity in London given periodic fluctuations in the availability of currency and credit. An original proposal by William Paterson in 1693 for a government 'fund of perpetual interest' was turned down in favour of another proposal by Paterson in 1694 to establish a company known as the Governor and Company of the Bank of England, whose capital, once raised, would be lent in its entirety to the government.

An ordinary finance act, now known as the Bank of England Act (1694), stipulated that the Bank was to be established via stock subscriptions which were to be lent to the government. A governor, deputy governor and 24 directors were to be elected by stockholders (holding £500 or more of stock).

### The evolution of the Bank's objectives and functions, 1694–1914

Under its original charter the Bank was allowed to issue bank notes, redeemable in silver coin, as well as to trade in bills and bullion. The notes of the Bank competed with other paper media of exchange, which comprised notes issued by the Exchequer and by private financial companies. In addition, customers could maintain deposit accounts with the Bank, which were transferable to other parties via notes drawn against deposit receipts (known as accomptable notes), thus providing an early form of cheque.

An early customer of the Bank was the Royal Bank of Scotland, which made arrangements to keep cash at the Bank from its outset in 1727. Loans were extended, predominantly in the form of discounting of bills, to individuals and companies, and the Bank undertook a large amount of lending (often via overdrafts) to the Dutch East India Company and, from 1711, to the South Sea Company. The Bank also acted as a mortgage lender, although this business never took off, and ceased some years later. Finally, an important function of the Bank was the remittance of cash to Flanders and elsewhere for the wars against Louis XIV, which was facilitated through correspondent arrangements with banks in Holland.

In 1697 the renewal of the Bank's charter for another ten years involved the passage of a second Bank Act, which increased the capital of the Bank and prohibited any other banks from being chartered in England and Wales. This monopoly was strengthened at the next renewal of the Bank's charter in 1708, when any association of six or more persons was forbidden to engage in banking activity, thereby precluding the establishment of any other joint stock banks. The Bank's position as banker to the government was consolidated in 1715 when it was decided that subscriptions for government debt issues would be paid to the Bank, and further that the Bank was to manage the government debt (the Ways and Means Act). The Bank then acted as manager of the government's debts from that date until 1997.

The Bank also encouraged the use of its own notes in preference to other media of exchange by persuading the Treasury to increase the denomination of Exchequer bills. By 1725 the Bank's notes had become sufficiently widely used as to be pre-printed for the first time. Although a number of private banks had developed by 1750, both within and outside London, none competed seriously with the Bank in the issue of notes. By 1770 most London bankers had ceased to issue notes, using Bank of England notes (and cheques) to settle balances among themselves in what had become a well-developed clearing system. Furthermore, in 1775 Parliament raised the minimum denomination for any non-Bank of England notes to one pound and, two years later, to five pounds, effectively guaranteeing the use of Bank of England notes as the dominant form of currency. Problems relating to counterfeiting, and to the harsh treatment of those caught in the act, were, however, perennial.

In Scotland, by contrast, no note issuing monopoly existed, and banks were free to issue notes, although two banks dominated, namely, the Bank of Scotland and the Royal Bank of Scotland. Furthermore, several private note-issuing banks were in business in Ireland, and the Bank of Ireland was established in 1783. These banks relied on the Bank of England to obtain silver and gold, particularly during times of financial stress, such as 1783 and 1793.

Following a dramatic rise in government expenditures after 1793 due to the war against France, which caused a large rise in the Bank's note issue, the Bank's gold holdings fell sharply. After a scare about a French invasion convertibility was suspended in 1797, and resumed only in 1821. In view of the financial exigencies of the war, and the fact that there was in such circumstances no limit to the expansion of its note issue, now effectively legal tender, by the Bank, a privately owned company, what is in retrospect surprising about the period of suspension is how comparatively low the resulting inflation was. Even so, it was high enough to set off a major debate on its causation, for example in the Parliamentary Committee on the High Price of Bullion (1810). This period saw a further consolidation of the Bank as a note issuer, since it began to issue small denomination notes (given the shortage of silver and gold coin), which became legal tender in 1812. Furthermore, in 1816 silver coin ceased to be legal tender for small payments. The government also moved most of its accounts to the Bank in 1805 (in 1834 all government accounts were finally moved to the Bank).

During the 18th century and early part of the 19th century, smaller country banks had proliferated throughout England and Wales, many issuing their own notes. Given the prohibition on joint stock banking, the capital of these banks was usually small, and they regularly became insolvent, especially when the demand for cash (coin) became strong. This contrasted sharply with Scotland, where joint stock banking and branch banking were permitted, and relatively few failures occurred. Following a severe banking crisis in 1825, during which many English country banks failed, an Act renewing the Bank's charter (in 1826) abolished the restrictions on banking activity more than 65 miles outside of London. This led to the establishment of several joint stock banks, while the Bank countered by opening several branches throughout England.

Thus, a semblance of a banking 'system' began to emerge by 1830, with the Bank of England as the 'central' bank. By far the best book on such nascent central banking at this time was that written by Henry Thornton, *An Enquiry into the Nature and Effects*

*of the Paper Credit of Great Britain* (1802). The practice of banks placing surplus funds with bill brokers also developed, with the Bank beginning to extend secured loans to these brokers on a more or less regular basis. In 1833 joint stock banks were finally allowed to operate in London, although they were not permitted to issue notes and thus were essentially deposit-taking banks only. The same Act specified that Bank of England notes were legal tender, and the Bank was also given the freedom to raise its discount rate freely (until then usury laws had placed a ceiling on interest rates) in response to cash outflows. The Bank's reaction (an early reaction function), in varying its interest rate, to cash inflows and outflows became codified around this time in what became known as the Palmer rule, after Horsley Palmer, Governor 1830–33, though the rule itself is usually dated from 1827.

The position of Bank of England notes was consolidated in an important Act, passed in 1844, generally known as the Bank Charter Act, preventing all note issuers from expanding their note issue above existing levels, and prohibiting the establishment of any new note-issuing banks. The 1844 Act also separated the issue and banking functions of the Bank into different departments, and required the Bank to publish a weekly summary of accounts.

Given that it did not pay interest on its deposits, the deposit activity of the Bank could never really compete with that of other banks, which expanded rapidly from 1850 onwards. In 1854, joint stock banks in London joined the London Clearing House, and it was agreed that clearing by transfer of Bank of England notes would be abandoned in favour of cheques drawn on bank accounts held at the Bank. Ten years later the Bank of England itself entered this clearing arrangement, and cheques drawn on bankers' accounts at the Bank became considered as paid.

Although the Bank had, from the beginning of the 19th century, periodically bought or sold exchequer bills to influence the note circulation, explicit open-market borrowing operations to support its discount rate began in 1847. From 1873 until 1890 the Bank almost always acted as a borrower rather than a lender of funds, as there were typically cash surpluses. As a result, the Bank introduced the systematic issue of Treasury bills via a regular tender offer in 1877. Treasury bills had a much shorter maturity (three to twelve months) than Exchequer bills (five or more years), and were to play an important role in raising funds from the outset of the First World War onwards.

By 1890, the Bank's role as lender of last resort became undisputed when it orchestrated the rescue of Baring Brothers and Co., a bank whose solvency had become suspect, threatening to cause systemic problems. Earlier, in 1866, the failure of a discount house, Overend, Gurney and Co., had precipitated a financial panic, during which the Bank discounted large amounts of bills and extended considerable loans. The Bank, however, was criticized for not doing more to prevent the onset of such a panic, not least by Walter Bagehot in his famous book *Lombard Street* (1873).

Throughout the 19th century, the Bank streamlined its discount facilities. In 1851 it overhauled its discount rules, stipulating that only those parties having a discount account could present bills, and that these bills had to have a maturity of fewer than 95 days and be endorsed by two creditworthy firms. In the latter part of the century, however, the Bank gradually came to favour discount houses, often by presenting them

with better rates of discount, and the range of firms doing discount business with the Bank declined. Discount houses were favoured because there was tension then between the Bank and the rapidly growing commercial banks – there was much banking consolidation via mergers between the 1870s and 1914 – and dealing via the intermediation of the discount houses enabled the Bank to influence market rates without having to interact directly with the joint-stock banks as counterparties.

Until the First World War the Bank pursued a discount policy which was primarily aimed at maintaining its gold reserves (as noted earlier) and which was conducted largely independently of the government. During the First World War, however, a clash occurred between the Bank Governor (Cunliffe) and the Chancellor (Law), during which the government made clear that it bore the ultimate responsibility for monetary policy, and that the Bank was expected to act on its direction.

## A subservient Bank, 1914–1992

The First World War was a major watershed not only in the history of the Bank but in the world more widely. It ushered in a half-century of increasing government intervention in every country, of a move towards socialist economies in most, and of communism in a wide swathe of countries. Under these circumstances the Bank became increasingly subservient to the government, in practice to the Chancellor of the Exchequer and to the Treasury, in the conduct of macro-monetary policy, its previous primary function.

Initially, however, there was little perception that the war and the rise of socialist ideas had irretrievably altered the context for policy. There was a desire to return to the previous regime, the gold standard, with its tried and true verities, as expressed in the Cunliffe Committee Report (the first report of the Committee on Currency and Foreign Exchange, 1919). That was probably inevitable under the circumstances, but a much more questionable decision was to return at the pre-war parity (against gold) despite the war-induced loss of markets (especially for the UK's main staples, textiles, coal, and iron and steel) and of competitiveness. Several of the other belligerent states, notably France, had inflated, and allowed their exchange to float downwards by so much that they did not seek to re-peg at the previous parity, but could choose a more suitable and competitive rate. While the decision to return to gold at the pre-war parity, steadfastly supported by the Bank, has been much criticized, the modern theory of time inconsistency provides some defence, namely, if the Bank had started to change the chosen rate to suit the immediate conjuncture it would have been expected to do so again in future, making commitment to the regime less credible.

Be that as it may, conditions after the First World War, with a weak balance of payments and a massively inflated money stock and floating debt, were hardly conducive to the re-establishment of gold standard conditions. Indeed, the authorities initially felt forced to move in the other direction, to unpeg the sterling–dollar rate that had been established since 1916 and formally to leave the gold standard in March 1919. The ending of the war led then to an extremely sharp and short boom and bust, in which tight monetary policy played a major role in the subsequent deflation (see Howson, 1975). From then until the return to gold at the pre-war parity of $4.86 to the pound in 1925,

the Bank advocated keeping the Bank rate high enough to facilitate that regime change, but decisions on Bank rate and on the conduct of monetary policy were joint, in that no proposal by the Bank could be activated without the agreement of the Chancellor and HM Treasury; the Treasury view, however, then was in line with classical thought, namely, that monetary policy could and should impinge primarily on nominal prices, with real output affected by real factors.

Despite the boom in the USA, growth in the UK was perceived as remaining low and unemployment high, at least as compared with its main comparator countries, in the 1920s. This was in part due to the continuing problems of restoring a successful economic regime in Europe, wherein German reparations had a malign effect. Although the Bank had lost much of its power to direct domestic monetary policy (to Whitehall), the Bank and its Governor, Montagu Norman, played a leading role in the various international exercises to try to restore Europe to normality and to the gold standard, (Sayers, 1976, ch. 8); and Sir Otto Niemeyer, a top Bank official, spread the gospel of establishing central banks to maintain price stability to the Dominions.

This whole structure came apart in the crisis that started in the USA in 1929 and then engulfed the rest of the world progressively through the subsequent four years. How far that collapse was itself exacerbated by the attempt to restore the gold standard has been explored by Eichengreen (1992). The UK was not in a strong economic position to avoid the world recession, but suffered a much smaller decline in output than in the USA or much of Continental Europe. The struggle to maintain the gold standard had required the maintenance of high interest rates, despite the imposition of controls on new issues in sterling by foreign governments. Despite high unemployment, wages and prices remained too sticky to allow the restoration of international competitiveness, though quite why this was so remains a debated issue.

With the gold standard collapsing in Europe and social pressures rising in the UK, there was diminishing political will to take the measures that appeared necessary to maintain the gold standard. The government decided to abandon it (in Norman's absence) in September 1931. From that moment onwards, until May 1997, the decision to alter the Bank rate moved decisively to Whitehall, effectively into the hands of the Chancellor, advised by HM Treasury. Of course, the Bank could, and did, make suggestions and played a major role in all the discussions, but the Chancellor took the decisions. Indeed, from June 1932 until November 1951 a policy of cheap money was followed whereby Bank rate was held constant at two per cent. Norman stated in 1937, 'I am an instrument of the Treasury'.

Meanwhile, the Bank was becoming more professional. The old system of circulating the Governor's chair in turn among the directors of the Bank, who were appointed from city (but not commercial bank) institutions, was superseded by the continuing governorship of Montagu Norman from 1920 until 1944. While this arose by happenstance rather than intention (see Sayers, 1976, ch. 22), it gave the Bank highly skilled, even if also highly idiosyncratic, leadership. Moreover, Norman introduced economists and other able officials into both the staff and the Court (the largely ceremonial board) of the Bank, although it is (apocryphally) recorded that Norman told one such economist, 'You are not here to tell me what to do, but to explain why I have done what I have already decided to do.'

In effect, the Bank had already become nationalized by the end of the Second World War. So the formal act of nationalization in 1946 brought about no real substantive changes, except that the Governor and his deputy (there has as yet been no woman Governor, although Rachel Lomax became the first female Deputy Governor in 2003), were appointed by the government for five years, renewable once more in most cases. Indeed, the more profound changes were brought about by Governor Gordon Richardson (1973–83) in the early 1980s. Until then, the Governor had been rather akin to a chairman, with the deputy and other internal directors as members of the board, setting strategy. Much of the executive power still lay with the Chief Cashier, who acted as leader of the heads of department, who ran the Bank. There was a clear break, a division, between the staff in the departments on the one hand and the Governors and Directors on the other. Richardson changed all that, concentrating power in the Governors' hands, sharply demoting the role of Chief Cashier, and underlining the precedence of (internal) directors over heads of department in all policy matters.

So, as power to decide the course of monetary policy – and to set the Bank rate – passed to Whitehall, what did these professional central bank officials do? The Bank came to have three main areas of responsibility. The first was the management of markets, notably the money market, the bond (gilts) market and the foreign exchange market. The UK had come out of the Second World War with a massively inflated ratio of debt to GDP, and its management had remained difficult and delicate, at least until after the War Loan Conversion of 1932. No sooner, however, had debt management been thereby put on a sounder foundation than the Second World War led to a further upsurge in the debt ratio, which led once again to debt management becoming a major preoccupation of policy. Thereafter, a combination of generally prudent fiscal policies, so that the debt ratio fell steadily, and then unexpected inflation in the 1970s, which accelerated the decline in the debt ratio, and market reforms in the 1980s, enabled the procedures of debt management to become simpler and standardized. Similarly, the floating exchange rate in the 1930s, followed by attempts to maintain pegged exchange rates both during the Second World War and thereafter under the Bretton Woods system, against a background of perennially weak balance of payments conditions, made the management of the UK's foreign exchange reserves and intervention on the foreign exchange market a crucial function of the Bank until 1992, when the UK was forced out of the European exchange rate mechanism. During crises the officials in charge of such foreign exchange operations were in telephone communication with the Chancellor and, occasionally, the Prime Minister at frequent intervals.

The Bank held that such market operations required a special professional expertise (though HM Treasury remained sceptical). The Bank threw itself into such activities with enthusiasm, and defended its pre-eminent role in this respect stoutly against all outside encroachment or criticism. Indeed, its market 'savvy' was its most powerful lever to persuade the Chancellor to its views in any debate; 'I am sorry, Chancellor, but the market will not accept that policy' was the strongest card it had to play, and that card was played often and with alacrity.

Although ultra-cheap money, with Bank rate held at two per cent, was abandoned in 1951, when the Conservative Party was returned to office, monetary policy in

general, and interest rates in particular, were still seen as both more ineffective and uncertain in their impact on domestic demand than the supposedly more reliable fiscal policy, a conclusion upheld by the controversial Radcliffe Report (1959). Consequently, fiscal policy was used to try to steer domestic demand while interest rates were raised to protect the balance of payments during the regular bouts of external weakness, and otherwise held low both to ease government finance and to support fixed investment. The outcome was a system in which inflationary pressures regularly threatened both the internal and external value of the currency. The chosen solution was to supplement market measures by direct interventions, in the case of external pressure via exchange controls, in the case of monetary expansion via direct controls on bank lending to the private sector. In both instances the Bank acted as the administrative agent of HM Treasury.

Such direct controls were introduced (on bank lending), or greatly extended and tightened (exchange controls), with the onset of the Second World War in 1939, but were continued, for the reasons outlined above, until 1971 for bank lending and 1979 for exchange controls. The administration of exchange controls required a large staff, but, unlike with its market operations, the Bank had little enthusiasm for acting in this guise. The Bank hoped to restore London to its former role as an international financial centre. While it succeeded in this through its encouragement of the Eurodollar market, aided by inept US policies, the continued administration of exchange controls remained an unwelcome burden. The same was true for direct controls on bank lending. Such controls were regarded by politicians as a comparatively painless way of dampening demand and inflation, while they were resented by commercial bankers. The Bank found itself in the middle of these disputes, and grew painfully aware of such controls' stultifying effect on efficiency, dynamism and growth. The Bank, inspired by John Fforde (the then executive director in charge of domestic finance, and subsequent Bank historian), pressed hard for these controls to be dismantled, and succeeded with the liberalizing reform of Competition and Credit Control (Bank of England, 1971).

As with many other cases of banking liberalization, such as in Scandinavia at the end of the 1980s, this was followed by an expansionary boom and then a bust, the fringe (secondary) bank crisis of 1973/74 (Reid, 1982). While there remain questions about how monetary policy could have been better applied to prevent the prior monetary boom (1972/73), there was no question but that the financial crisis found both the Bank and the banks unskilled in risk management and unprepared for adverse shocks to financial stability. The long period of financial repression – that is, controls on bank lending to the private sector and force-feeding with government debt – had had the by-product of making the (core) commercial banking system safe between the mid-1930s and the early 1970s. The central banking function of maintaining financial stability, via regulation and supervision, had atrophied.

This had not been so earlier, and the Bank had been closely involved in the rescue of Williams Deacon's Bank by the Royal Bank of Scotland in 1930 (Sayers, 1976, ch. 10), and in helping to shape the structure of both the commercial banking system and the London Discount Market Association. Williams Deacon's had got into trouble largely because of bad debts from Lancashire cotton companies. Norman, and the

Bank, extended their structural interventions beyond banking to try to encourage strategic amalgamations to shore up the positions of weakened companies in a variety of industries, such as cotton, steel, shipping, armaments (Sayers, 1976, ch. 14). The Bank's involvement in structural matters outside of banking itself was episodic depending on both circumstances and personalities. Another example of such Bank involvement was the considerable role it played in the reform of the UK capital market in the 1980s, more familiarly known as 'Big-Bang'. But views on whether the Bank has any locus in such wider structural issues vary over time; the early 2000s saw a major withdrawal by the Bank from any such involvement.

The fringe bank crisis in the early 1970s was, however, a clarion call to put more emphasis on its third main function, bank supervision and regulation. The immediate result was a reorganization in the Bank. Initially a nucleus of a new specialized department was established in the Discount Office where the limited staff assigned to this role had sat, which rapidly absorbed staff and resources. Thereafter this became a separate department devoted to banking supervision and regulation (its first head was George Blunden, later to become Deputy Governor, who handed it on to Peter Cooke in 1976). Its position was regularized in the Banking Act (1979) which gave formal powers to the Bank to authorize, monitor, supervise, control and, under certain circumstances, withdraw prior authorization (tantamount to closure) for banks. No such powers had been available before that date. Meanwhile, other financial intermediaries, such as building societies or insurance companies, remained (lightly) regulated by various government departments.

The fringe bank crisis was almost entirely domestic, confined to British headquartered companies. Meanwhile, however, the onwards march of liberalization (involving the removal of direct controls, notably exchange controls in 1979) and of information technology were leading to a growing internationalization of financial business. For a variety of reasons, mostly relating to the innovation of the Eurodollar and Euromarkets, London regained its role as an international financial centre in the 1960s, and thus international monetary problems became of particular importance to the Bank, which took a leading role in such matters from the 1970s onwards.

Central bankers had met regularly at the headquarters of the Bank for International Settlements (BIS) in Basel for many years. It was, therefore, a logical step for supervisory officials also to come together at Basel on regular occasions to discuss matters of common interest. Thus was born (in 1974), as a result of an initiative from Gordon Richardson, the Basel Committee on Banking Regulation and Supervisory Practices. For the first 15 years of its existence it was chaired by the participant from the Bank of England, and was usually known by his name; thus, the Blunden Committee (1974–77) gave way in due course to the Cooke Committee (1977–88). The failures of Franklin National and Herstatt prompted the First Basel Concordat, which allocated responsibility for supervising internationally active banks to home and host authorities.

So by the mid-1970s, a need was perceived for banking supervision at both the domestic and, via consolidation, at the international levels. The purpose of these initiatives was to clarify where responsibility lay for the supervision of international banks, to prevent fragile, and possibly fraudulent, banking leading to avoidable failures and potential systemic crises.

Despite the growing number of bank supervisors, and notable success in reversing prior declines in capital ratios, the history of banking in the subsequent decades in the UK was spotted by occasional bank failures. Unlike the fringe bank crisis, none was, or was allowed to become, systemic, nor did individual depositors lose any money, except in the case of Bank of Credit and Commerce International (BCCI), and even in that case the deposit protection scheme provided some relief. The failures of Johnson–Matthey (in 1984), BCCI (in 1991) and Barings (in 1995) were all isolated cases of bad, in some respects fraudulent, banking.

The main problem of the 1970s and 1980s was, however, that of combating inflation, which soared to heights previously unknown, not only in peacetime but even in wartime, during the 1970s, up to 25 per cent per annum. There were three main theories, though divisions between them were never completely distinct. The first was the cost-push theory, that inflation was driven by over-mighty trade unions, seeking to increase the relative real pay of their members; the appropriate remedy was then prices and incomes policies plus reform (and constraint) of trades unions. The second was the (vertical) Phillips curve analysis; the remedy here was to raise unemployment above the 'natural' rate to reduce inflation. The third was that inflation was a monetary phenomenon; the remedy was to control the rate of growth of the (appropriate) monetary aggregate.

Until the mid-1970s, both major political parties, the Bank and HM Treasury all professed some combination of theories 1 (cost-push) and 2 (Phillips curve). Left-leaning politicians, academics and officials tended to put more weight on cost-push. In the 1960 and 1970s the third, monetarist, view seemed to explain events better and gained strength, not only in the USA (Milton Friedman) but also in the UK. In particular, the surge in inflation in the UK in 1973–75 followed closely behind the rapid expansion of broad (but not narrow) money in 1972–73. So, when in opposition, the leading Conservative politicians Keith Joseph and Margaret Thatcher embraced a version of monetarism.

When they came to power in 1979, they tried to commit monetary policy to follow a target for broad money, via the Medium Term Financial Strategy. In order to achieve this, nominal, and real, interest rates were kept high, and the exchange rate appreciated sharply, partly under the influence of North Sea oil and confidence in Thatcherite policies. Inflation duly declined, as planned, but broad money growth did not. This latter was partly due to the abolition of the 'corset' in 1980. The 'corset' was a reformulated, and somewhat disguised, direct control over commercial bank expansion that had been pressed into service on several occasions during the 1970s. The Bank was glad to see the end of exchange controls and direct controls over bank lending, but had never shared the government's monetarist faith in trying to set, and stick to, targets for the growth of (the various) monetary aggregates.

The empirical demonstration of the unpredictability of the relationship between (broad) money and nominal incomes in the early 1980s soon weakened the government's own faith. After moving from one monetary target to several joint targets, and an attempt to hit the broad money target by 'overfunding', an exercise criticized by many as artificial, the government abandoned its monetary targetry in 1986.

That left the question of how monetary policy, and with it control of inflation, was to be managed or, in the standard phrase, 'anchored'. The then Chancellor, Nigel Lawson,

wanted to 'anchor' by joining the exchange rate mechanism (ERM) of the European Monetary System and leaving the steering of monetary policy to the Bundesbank. The Prime Minister, Mrs Thatcher, and her adviser, Alan Walters, were opposed, both on economic grounds (that such a pegged system was 'half-baked') and for wider political reasons. There was a battle royal in which the Bank was left on the sidelines. Lawson was sacked, but eventually Mrs Thatcher was, grudgingly, persuaded to allow the UK to join the ERM in October 1990.

This was in the aftermath of German reunification, and the expenditures connected with that led the Bundesbank to keep interest rates higher than was tolerable for the UK (or Italy). The UK was in the throes of a sharp downturn in housing prices, following an unstable housing boom in the late 1980s. With the Conservatives having become politically weaker, there was just no stomach to raise interest rates to the levels necessary to sustain the ERM. The UK was forced out in September 1992.

## Independent and focused, 1992–

The ejection of the UK from the ERM left the government and HM Treasury with the recurrent problem of how to manage, to 'anchor', monetary policy. Both monetary and exchange rate targets had been tried, and both had been found wanting. While the economic experience of the 1980s was better than that of the stagflationary 1970s, it was hardly stellar, with a boom–bust cycle at the end of the decade.

Meanwhile, a new approach had been adopted in New Zealand, whereby the central bank was given administrative freedom to vary interest rates for the purpose of hitting a target for the inflation rate, jointly set by the government and the central bank: that is, inflation targetry. This obviated one of the shortcomings of monetary targetry, namely, the unpredictability of the velocity of money; it left setting the goals of policy, the overall strategy, in the hands of government, but shifted the (constrained) discretion to vary interest rates to the professional and technical judgement of the central bank. This procedure soon generated a strong body of academic support (for example, Fischer, 1994).

Although Conservative Chancellors (both Lawson and Lamont) had toyed with the idea of giving the Bank operational independence, consecutive Prime Ministers (Thatcher and Major) refused, primarily on political grounds. Nevertheless Lamont wanted to move to an inflation target. But there was a problem of governmental credibility. To foster credibility, Lamont now encouraged (in 1992/93) the Bank to prepare and to publish an independent forecast of the likely projection for inflation, the *Inflation Report* (on the assumption of unchanged policies); this was a reversal of prior habits whereby HM Treasury and Ministers customarily censored Bank publications and discouraged any publication of internal Bank forecasts. The process of gradually giving the Bank a more independent role in setting monetary policy took a step further when the next Chancellor, Clarke, not only held a meeting with the Governor, and the Bank, to discuss future changes in interest rates, but published the minutes of the meeting, including the Governor's initial statement, verbatim; this was termed the Ken (Clarke) and Eddie (George) show. That said, Clarke had strong views on the appropriate policy and on a couple of occasions overruled the Governor's suggestions.

At that time – the mid-1990s – there were still question marks over the Labour Party's ability to manage the economy; financial markets are inherently suspicious of left-leaning governments. So Labour had more to gain (than the Conservatives), in terms of confidence and lower interest rates, by granting operational independence (back) to the Bank. In advance of the 1997 election the then shadow Chancellor, Gordon Brown, was cautious; while indicating general support for both inflation targetry and operational independence, he stated that he wanted time to see how well the Bank performed before granting such independence. But, within days of winning the election, he made that strategic change to the monetary regime.

This was, of course, a great prize for the Bank, but it did not come without cost. In the same month as operational independence was awarded to the Bank, both debt management and banking supervision were hived off, to a separate Debt Management Office (DMO) and Financial Services Authority (FSA) respectively. With the government debt to GDP ratio having declined and capital markets strengthened, debt management had become more of a routine and standardized exercise. Nevertheless, its departure to the DMO, and the fact that the float of the exchange rate after 1992 was kept 'clean', that is, without intervention, meant that much of the market operations which had been so central to the Bank in the post-Second World War period disappeared, though its money market operations, of course, continued. The administration of direct controls had gone at the beginning of the 1980s. And now banking supervision was also taken away. This meant that almost *all* the prime functions that the Bank had undertaken in its post-Second World War period of subservience had now gone. Instead, the Bank was now focused on varying interest rates to achieve the inflation target set for it by the Chancellor.

There are numerous arguments, quite evenly balanced, for whether bank supervision should be kept within a central bank or put with a separate Financial Services Authority (FSA), covering both banks and other financial intermediaries (see Goodhart, 2000). Be that as it may, there are various aspects of the financial system, such as oversight of the payments' system, and of crisis management, such as lender of last resort functions, which cannot be delegated to an FSA. Moreover, the achievement of price stability is likely to be seriously compromised by any serious bout of financial instability – and vice versa, with financial stability adversely affected by price instability. So the removal of individual bank supervision does not absolve the Bank from concern with financial stability issues more widely; indeed, the Bank is specifically charged with maintaining overall systemic stability in the financial system. But exactly what that means when responsibility for the conduct of individual bank supervision is located elsewhere is not yet entirely clear.

What it certainly does mean is that the FSA, the Bank, and the political authorities as the ultimate source of any needed fiscal support have to work extremely closely together, in advising on any new regulations (whether domestic or international), in monitoring developments (as in the Financial Stability Review), and in crisis management. This latter task would be done via the Tripartite Standing Committee (FSA, Bank, and HM Treasury), set up in 1997, although so far no such financial (as contrasted with simulated 'war games') crisis has occurred, though the Committee did meet after the terrorist attacks on 7 July, 2005. How successful crisis management by such a committee may be has yet to be seen.

The monetary policy function of the Bank, now its central preoccupation, has, however, been very successful by all the usual criteria. In several papers Luca Benati (for example, Benati, 2005) has demonstrated that the variance of both GDP and of inflation around its target has been lower under the inflation targetry regime (whether taken as starting in 1992 or in 1997) than under any previous historical regime. The procedures of having a Monetary Policy Committee consisting of five senior Bank officials and four outside experts (appointed by the Chancellor), with the Committee serviced by Bank staff, has worked generally smoothly and well. So the Bank's reputation and credibility have rarely been higher, although now tightly focused on one main function.

CHARLES A. E. GOODHART

*See Also* **banking crises; bullionist controversies (empirical evidence); gold standard; inflation targeting; monetary policy, history of.**

### *Bibliography*

Acres, W. 1931. *The Bank of England from Within*. London: Oxford University Press.

Andréadès, A. 1909. *A History of the Bank of England*. London: P. S. King and Sons.

Bagehot, W. 1873. *Lombard Street*. London: Kegan, Paul and Co.

Bank of England. 1971. *Competition and Credit Control*. London: Bank of England.

Bank for International Settlements. 1963. 'Bank of England', in *Eight European Central Banks*. Basle: Bank for International Settlements.

Benati, L. 2005. The Inflation-targeting framework from an historical perspective. *Bank of England Quarterly Bulletin* 45(2), 160–8.

Bowman, W. 1937. *The Story of the Bank of England: From its Foundation in 1694 until the Present Day*. London: Herbert Jenkins.

Chapham, R. 1968. *Decision Making: A Case Study of the Decision to Raise the Bank Rate in September 1957*. London: Routledge and Kegan Paul.

Clapham, J. 1944. *The Bank of England: A History*. Cambridge: Cambridge University Press.

Clay, H. 1957. *Lord Norman*. London: Macmillan.

Committee on Currency and Foreign Exchange After the War (Cunliffe Committee). 1918. *First Interim Report*, Cmnd. 9182; and 1919. *Final Report*, Cmnd 464, London: HMSO.

Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression*. New York: Oxford University Press.

Feavearyear, A. 1963. *The Pound Sterling: A History of English Money*, 2nd edn, rev. E. Morgan. Oxford: Clarendon.

Fforde, J. 1992. *The Bank of England and Public Policy 1941–1958*. Cambridge: Cambridge University Press.

Fischer, S. 1994. 'Modern central banking', in F. Capie, C. Goodhart, S. Fischer and N. Schnadt, *The Future of Central Banking*. Cambridge: Cambridge University Press.

Geddes, P. 1987. *Inside the Bank of England*. London: Boxtree.

Giuseppi, J. 1966. *The Bank of England: A History from its Foundation in 1694*. London: Evans Brothers Limited.

Hennessey, E. 1992. *A Domestic History of the Bank of England 1930–1960*. Cambridge: Cambridge University Press.

Goodhart, C. 2000. The organisational structure of banking supervision. Special Paper No. 127. London: Financial Markets Group Research Centre, London School of Economics. Subsequently published in *Economic Notes* 31, 1–32.

Howson, S. 1975. *Domestic Monetary Management in Britain, 1919–38*. Cambridge: Cambridge University Press.

Radcliffe Report. 1959. *Report: Committee on the Working of the Monetary System*. Cmnd 827. London: HMSO.
Reid, M. 1982. *The Secondary Banking Crisis, 1973–75: Its Causes and Course*. London: Macmillan.
Richards, R. 1929. *The Early History of Banking in England*. London: Frank Cass and Co.
Rogers, J. 1887. *The First Nine Years of the Bank of England*. Oxford: Clarendon.
Sayers, R. 1936. *Bank of England Operations, 1890–1914*. London: P.S. King and Son.
Sayers, R. 1957. *Central Banking After Bagehot*. Oxford: Clarendon.
Sayers, R. 1976. *The Bank of England, 1891–1944*. Cambridge: Cambridge University Press.
Smith, V. 1936. *The Rationale of Central Banking*. London: P.S. King and Son.
Steele, H. and Yerbury, F. 1930. *The Old Bank of England*. London: Ernest Benn.
Stockdale, E. 1967. *The Bank of England in 1934*. London: Eastern Press.
Thornton, H. 1802. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*. New York: Kelley, 1962.
Ziegler, D. 1990. *Central Bank, Peripheral Industry: The Bank of England in the Provinces 1826–1913*. London: Leicester University Press.

## banking crises

There are two distinct phenomena associated with banking system distress: exogenous shocks that produce insolvency, and depositor withdrawals during 'panics'. These two contributors to distress often do not coincide. For example, in the rural United States during the 1920s many banks failed, often with high losses to depositors, but those failures were not associated with systemic panics. In 1907, the United States experienced a systemic panic, originating in New York. Although some banks failed in 1907, failures and depositor losses were not much higher than in normal times. As the crisis worsened, banks suspended convertibility until uncertainty about the incidence of the shock had been resolved.

The central differences between these two episodes relate to the commonality of information regarding the shocks producing loan losses. In the 1920s, the shocks were loan losses in agricultural banks, geographically isolated and fairly transparent. Banks failed without resulting in system-wide concerns. During 1907, the ultimate losses for New York banks were small, but the incidence was unclear *ex ante* (loan losses reflected complex connections to securities market transactions, with uncertain consequences for some New York banks). This confusion hit the financial system at a time of low liquidity, reflecting prior unrelated disturbances in the balance of payments (Bruner and Carr, 2007).

Sometimes, large loan losses, and confusion regarding their incidence, occurred together. In Chicago in mid-1932 losses resulted in many failures and also in widespread withdrawals from banks that did not ultimately fail. Research has shown that the banks that failed were exogenously insolvent; solvent Chicago banks experiencing withdrawals did not fail. In other episodes, however, bank failures may reflect illiquidity resulting from runs, rather than exogenous insolvency.

Banking crises can differ according to whether they coincide with other financial events. Banking crises coinciding with currency collapse are called 'twin' crises (as in Argentina in 1890 and 2001, Mexico in 1995, and Thailand, Indonesia and Korea in 1997). A twin crisis can reflect two different chains of causation: an expected devaluation may encourage deposit withdrawal to convert to hard currency before devaluation (as in the United States in early 1933); or, a banking crisis can cause devaluation, either through its adverse effects on aggregate demand or by affecting the supply of money (when a costly bank bail-out prompts monetization of government bail-out costs). Sovereign debt crises can also contribute to bank distress when banks hold large amounts of government debt (for example, in the banking crises in the United States in 1861, and in Argentina in 2001).

The consensus views regarding banking crises' origins (fundamental shocks versus confusion), the extent to which crises result from unwarranted runs on solvent banks, the social costs attending runs, and the appropriate policies to limit the costs of banking crises (government safety nets and prudential regulation) have changed dramatically, and more than once, over the course of the 19th and 20th centuries. Historical

experience played a large role in changing perspectives toward crises, and the US experience had a disproportionate influence on thinking. Although panics were observed throughout world history (in Hellenistic Greece, and in Rome in AD 33), prior to the 1930s, in most of the world, banks were perceived as stable, large losses from failed banks were uncommon, banking panics were not seen as a great risk, and there was little perceived need for formal safety nets (for example, deposit insurance, or programmes to recapitalize banks). In many countries, ad hoc policies among banks, and sometimes including central banks, to coordinate bank responses to liquidity crises (as, for example, during the failure of Barings investment bank in London in 1890), seemed adequate for preventing systemic costs from bank instability.

## Unusual historical instability of US banks

The unusual experience of the United States was a contributor to changes in thinking that led to growing concerns about banks runs, and the need for aggressive safety net policies to prevent or mitigate runs. In retrospect, the extent to which US banking instability informed thinking and policy outside the United States seems best explained by the size and pervasive influence of the United States; in fact, the US crises were unique and reflected peculiar features of US law and banking structure.

The US panic of 1907 (the last of a series of similar US events, including 1857, 1873, 1884, 1890, 1893, and 1896) precipitated the creation of the Federal Reserve System in 1913 as a means of enhancing systemic liquidity, reducing the probability of systemic depositor runs, and mitigating the costs of such events. This innovation was specific to the United States (other countries either had established central banks long before, often with other purposes in mind, or had not established central banks), and reflected the unique US experience with panics – a phenomenon that the rest of the world had not experienced since 1866, the date of the last British banking panic (Bordo, 1985).

For example, Canada did not suffer panics like those of the United States and did not establish a central bank until 1935. Canada's early decision to permit branch banking throughout the country ensured that banks were geographically diversified and thus resilient to large sectoral shocks (like those to agriculture in the 1920s and 1930s), able to compete through the establishment of branches in rural areas (because of the low overhead costs of establishing additional branches), and able to coordinate the banking system's response in moments of confusion to avoid depositor runs (the number of banks was small, and assets were highly concentrated in several nationwide institutions). Outside the United States, coordination among banks facilitated systemic stability by allowing banks to manage incipient panic episodes to prevent widespread bank runs. In Canada, the Bank of Montreal would occasionally coordinate actions by the large Canadian banks to stop crises before the public was even aware of a possible threat.

The United States, however, was unable to mimic this behaviour on a national or regional scale (Calomiris, 2000; Calomiris and Schweikart, 1991). US law prohibited nationwide branching, and most states prohibited or limited within-state branching. US banks, in contrast to banks elsewhere, were numerous (for example, numbering more

than 29,000 in 1920), undiversified, insulated from competition, and unable to coordinate their response to panics (US banks established clearing houses, which facilitated local responses to panics beginning in the 1850s, as emphasized by Gorton, 1985).

The structure of US banking explains why the United States uniquely had banking panics in which runs occurred despite the health of the vast majority of banks. The major US banking panics of the post-bellum era (listed above) all occurred at business cycle peaks, and were preceded by spikes in the liabilities of failed businesses and declines in stock prices; indeed, whenever a sufficient combination of stock price decline and rising liabilities of failed businesses occurred, a panic *always* resulted (Calomiris and Gorton, 1991). Owing to the US banking structure, panics were a predictable result of business cycle contractions that, in other countries, resulted in an orderly process of financial readjustment.

The United States, however, was not the only economy to experience occasional waves of bank failures before the First World War. Nor did it experience the highest bank failure rates, or bank failure losses. None of the US banking panics of the pre-First World War era saw nationwide banking distress (measured by the negative net worth of failed banks relative to annual GDP) greater than the 0.1 per cent loss of 1893. Losses were generally modest elsewhere, but Argentina in 1890 and Australia in 1893, where the most severe cases of banking distress occurred during this era, suffered losses of roughly ten per cent of GDP. Losses in Norway in 1900 were three per cent and in Italy in 1893 one per cent of GDP, but with the possible exception of Brazil (for which data do not exist to measure losses), there were no other cases in 1875–1913 in which banking loss exceeded one per cent of GDP.

Loss rates tended to be low because banks structured themselves to limit their risk of loss, by maintaining adequate equity-to-assets ratios, sufficiently low asset risk, and adequate asset liquidity. Market discipline (the fear that depositors would withdraw their funds) provided incentives for banks to behave prudently. The picture of small depositors lining up around the block to withdraw funds has received much attention, but perhaps the more important source of market discipline was the threat of an informed (often 'silent') run by large depositors (often other banks). Banks maintained relationships with each other through interbank deposits and the clearing of public deposits, notes and bankers' bills. Banks often belonged to clearing houses that set regulations and monitored members' behaviour. A bank that lost the trust of its fellow bankers could not long survive.

## Changing perceptions of banking instability

This perception of banks as stable, as disciplined by depositors and interbank arrangements to act prudently, and as unlikely to fail was common prior to the 1930s. The banking crises of the Great Depression changed this perception. US Bank failures resulted in losses to depositors in the 1930s in excess of three per cent of GDP. Bank runs, bank holidays (local and national government-decreed periods of bank closure to attempt to calm markets and depositors), and widespread bank closure suggested a chaotic and vulnerable system in need of reform. The Great Depression saw an unusual raft of banking regulations, especially in the United States, including restrictions on

bank activities (the separation of commercial and investment banking, subsequently reversed in the 1980s and 1990s), targeted bank recapitalizations (the Reconstruction Finance Corporation), and limited government insurance of deposits.

Academic perspectives on the Depression fuelled the portrayal of banks as crisis-prone. The most important of these was the treatment of the 1930s banking crises by Milton Friedman and Anna Schwartz in their book, *A Monetary History of the United States* (1963). Friedman and Schwartz argued that many solvent banks were forced to close as the result of panics, and that fear spread from some bank failures to produce failures elsewhere. They saw the early failure of the Bank of United States in 1930 as a major cause of subsequent bank failures and monetary contraction. They lauded deposit insurance: 'federal deposit insurance, to 1960 at least, has succeeded in achieving what had been a major objective of banking reform for at least a century, namely, the prevention of banking panics'. Their views that banks were inherently unstable, that irrational depositor runs could ruin a banking system, and that deposit insurance was a success, were particularly influential coming from economists known for their scepticism of government interventions.

Since the publication of *A Monetary History of the United States*, however, other scholarship (notably, the work of Elmus Wicker, 1996, and Charles Calomiris and Joseph Mason, 1997; 2003a) has led to important qualifications of the Friedman–Schwartz view of bank distress during the 1930s, and particularly of the role of panic in producing distress. Detailed studies of particular regions and banks' experiences do not confirm the view that panics were a nationwide phenomenon during 1930 or early 1931, or an important contributor to nationwide distress until very late in the Depression (that is, early 1933). Regional bank distress was often localized and traceable to fundamental shocks to the values of bank loans. Not only does it appear that the failure of the Bank of United States had little effect on banks nationwide in 1930, one scholar has argued that there is evidence that the bank was, in fact, insolvent when it failed (Lucia, 1985).

Other recent research on banking distress during the pre-Depression era has also de-emphasized inherent instability, and focused on the historical peculiarity of the US banking structure and panic experience, noted above. Furthermore, recent research on the destabilizing effects of bank safety nets has been informed by the experience of the US Savings and Loan industry debacle of the 1980s, the banking collapses in Japan and Scandinavia during the 1990s, and similar banking system debacles occurring in 140 developing countries in the last quarter of the 20th century, all of which experienced banking system losses in excess of one per cent of GDP, and more than 20 of which experienced losses in excess of ten per cent of GDP (data are from Caprio and Klingebiel, 1996, updated in private correspondence with these authors). Empirical studies of these unprecedented losses concluded that deposit insurance and other policies that protect banks from market discipline, intended as a cure for instability, have become instead the single greatest source of banking instability.

The theory behind the problem of destabilizing protection has been well known for over a century, and was the basis for US President Franklin Roosevelt's opposition to deposit insurance in 1933 (an opposition shared by many). Deposit insurance was seen as undesirable special interest legislation designed to benefit small banks. Numerous

attempts to introduce it failed to attract support in Congress (Calomiris and White, 1994). Deposit insurance removes depositors' incentives to monitor and discipline banks, and frees bankers to take imprudent risks (especially when they have little or no remaining equity at stake, and see an advantage in 'resurrection risk taking'). The absence of discipline also promotes banker incompetence, which leads to unwitting risk taking.

Empirical research on late 20th-century banking collapses has produced a consensus that the greater the protection offered by a country's bank safety net, the greater the risk of a banking collapse (see, for example, Caprio and Klingebiel, 1996, and the papers from a 2000 World Bank conference on bank instability listed in the bibliography). Empirical research on prudential bank regulation emphasizes the importance of subjecting some bank liabilities to the risk of loss to promote discipline and limit risk taking (Shadow Financial Regulatory Committee, 2000; Mishkin, 2001; Barth, Caprio and Levine, 2006).

Studies of historical deposit insurance reinforce these conclusions (Calomiris, 1990). The basis for the opposition to deposit insurance in the 1930s was the disastrous experimentation with insurance in several US states during the early 20th century, which resulted in banking collapses in all the states that adopted insurance. Government protection had played a similarly destabilizing role in Argentina in the 1880s (leading to the 1890 collapse) and in Italy (leading to its 1893 crisis). In retrospect, the successful period of US deposit insurance, from 1933 to the 1960s, to which Friedman and Schwartz referred, was an aberration, reflecting limited insurance during those years (insurance limits were subsequently increased), and the unusual macroeconomic stability of the era.

Models of banking crises followed trends in the empirical literature. The understanding of bank contracting structures, in light of potential crises, has been a consistent theme. Banks predominantly hold illiquid assets ('opaque,' non-marketable loans), and finance those assets mainly with deposits withdrawable on demand. Banks are not subject to bankruptcy preference law, but rather, apply a first-come, first-served rule to failed bank depositors (depositors who are first in line keep the cash paid out to them). These attributes magnify incentives to run banks. An early theoretical contribution, by Douglas Diamond and Philip Dybvig (1983), posited a banking system susceptible to the constant threat of runs, with multiple equilibria, where runs can occur irrespective of problems in bank portfolios or any fundamental demand for liquidity by depositors. They modelled deposit insurance as a means of avoiding the bad (bank run) equilibrium. Over time, other models of banks and depositor behaviour developed different implications, emphasizing banks' abilities to manage risk effectively, and the beneficial incentives of demand deposits in motivating the monitoring of banks in the presence of illiquid bank loans (Calomiris and Kahn, 1991).

The literatures on banking crises also rediscovered an older line of thought emphasized by John Maynard Keynes (1931) and Irving Fisher (1933): market discipline implies links between increases in bank risk, depositor withdrawals and macroeconomic decline. As banks respond to losses and increased risk by curtailing the supply of credit, they can aggravate the cyclical downturn, magnifying declines in investment, production, and asset prices, whether or not bank failures occur (Bernanke, 1983; Bernanke and Gertler, 1990; Calomiris and Mason, 2003b; Allen and Gale, 2004; Von Peter,

2004; Calomiris and Wilson, 2004). New research explores general equilibrium linkages among bank credit supply, asset prices and economic activity, and adverse macroeconomic consequences of 'credit crunches' that result from banks' attempts to limit their risk of failure. This new generation of models provides a rational-expectations, 'shock-and-propagation' approach to understanding the contribution of financial crises to business cycles, offering an alternative to the endogenous-cycles, myopic-expectations view pioneered by Hyman Minsky (1975) and Charles Kindleberger (1978).

CHARLES W. CALOMIRIS

*See Also* **credit rationing; currency crises; deposit insurance; Great Depression; moral hazard.**

### *Bibliography*


Allen, F. and Gale, D. 2004. Financial fragility, liquidity, and asset prices. *Journal of the European Economic Association* 2, 1015–48.

Barth, J.R., Caprio, G. and Levine, R. 2006. *Rethinking Bank Regulation: Till Angels Govern*. Cambridge: Cambridge University Press.

Bernanke, B.S. 1983. Nonmonetary effects of the financial crisis in the propagation of the great depression. *American Economic Review* 73, 257–76.

Bernanke, B.S. and Gertler, M. 1990. Financial fragility and economic performance. *Quarterly Journal of Economics* 105, 87–114.

Bordo, M. 1985. The impact and international transmission of financial crises: some historical evidence, 1870–1933. *Revista di Storia Economica* 2(2d), 41–78.

Boyd, J., Gomis, P., Kwak, S. and Smith, B. 2000. A user's guide to banking crises. Conference paper. Washington, DC: World Bank.

Bruner, R.F. and Carr, S.D. 2007. Money Panic: Lessonsfrom the Financial Crisis of 1907. New York: Wiley.

Calomiris, C.W. 1990. Is deposit insurance necessary? A historical perspective. *Journal of Economic History* 50, 283–95.

Calomiris, C.W. 2000. *U.S. Bank Deregulation in Historical Perspective*. Cambridge: Cambridge University Press.

Calomiris, C.W. and Gorton, G. 1991. The origins of banking panics: models, facts, and bank regulation. In *Financial Markets and Financial Crises*, ed. R.G. Hubbard. Chicago: University of Chicago Press.

Calomiris, C.W. and Kahn, C.M. 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81, 497–513.

Calomiris, C.W. and Mason, J.R. 1997. Contagion and bank failures during the great depression: the June 1932 Chicago banking panic. *American Economic Review* 87, 863–83.

Calomiris, C.W. and Mason, J.R. 2003a. Fundamentals, panics and bank distress during the depression. *American Economic Review* 93, 1615–47.

Calomiris, C.W. and Mason, J.R. 2003b. Consequences of bank distress during the great depression. *American Economic Review* 93, 937–47.

Calomiris, C.W. and Schweikart, L. 1991. The panic of 1857: origins, transmission, and containment. *Journal of Economic History* 51, 807–34.

Calomiris, C.W. and White, E.N. 1994. The origins of federal deposit insurance. In *The Regulated Economy: A Historical Approach to Political Economy*, ed. C. Goldin and G. Libecap. Chicago: University of Chicago Press.

Calomiris, C.W. and Wilson, B. 2004. Bank capital and portfolio management: the 1930s 'capital crunch' and scramble to shed risk. *Journal of Business* 77, 421–55.

Caprio, G. and Klingebiel, D. 1996. Bank insolvencies: cross country experience. Working Paper No. 1620. Washington, DC: World Bank.

Cull, R., Senbet, L. and Sorge, M. 2000. Deposit insurance and financial development. Conference paper. Washington, DC: World Bank.
Demirguc-Kunt, A. and Detragiache, E. 2000. Does deposit insurance increase banking system stability? Conference paper. Washington, DC: World Bank.
Demirguc-Kunt, A. and Huizinga, H. 2000. Market discipline and financial safety net design. Conference paper. Washington, DC: World Bank.
Diamond, D. and Dybvig, P. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91, 401–19.
Fisher, I. 1933. The debt deflation theory of great depressions. *Econometrica* 1, 337–57.
Friedman, M. and Schwartz, A.J. 1963. *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton University Press.
Gorton, G. 1985. Clearing houses and the origin of central banking in the United States. *Journal of Economic History* 45, 277–83.
Honohan, P. and Klingebiel, D. 2000. Controlling fiscal costs of banking crises. Conference paper. Washington, DC: World Bank.
Keynes, J.M. 1931. The consequences to the banks of the collapse of money values. In *Essays in Persuasion*. New York: W.W. Norton, 1963.
Kindleberger, C.P. 1978. *Manias, Panics, and Crashes: A History of Financial Crises*. New York: Basic Books.
Lucia, J.L. 1985. The failure of the bank of United States: a reappraisal. *Explorations in Economic History* 22, 402–16.
Minsky, H.P. 1975. *John Maynard Keynes*. New York: Columbia University Press.
Mishkin, F.S. 2001. *Prudential Supervision: What Works and What Doesn't*. Chicago: University of Chicago Press.
Shadow Financial Regulatory Committee. 2000. *Reforming Bank Capital Regulation*. Washington, DC: American Enterprise Institute.
Von Peter, G. 2004. Asset prices and banking distress: a macroeconomic approach. BIS Working Paper No. 167. Basel: Bank of International Settlements.
Wicker, E. 1996. *The Banking Panics of the Great Depression*. Cambridge: Cambridge University Press.

## banking industry

The distinctive function of banks is the transformation of short-term deposits into longer-term, less liquid and riskier loans (Fama, 1980; 1985; Diamond and Rajan, 2001; Gorton and Winton, 2003). By raising funds from depositors and providing credit, banks avoid the duplication of monitoring, which reduces the overall cost of transferring funds from capital suppliers to its users (Leland and Pyle, 1977; Diamond, 1984). At the same time, however, the greater liquidity of liabilities than of assets, which are typically longer-term and riskier, makes bank balance sheets vulnerable. Not only may banks fail if they are unable to obtain repayment of their loans, but depositors might even decide to withdraw their assets simply anticipating that others will do so. Such a 'bank run' can drive an otherwise sound bank to insolvency (Diamond and Dybvig, 1983). The need to protect depositors and so guarantee a stable monetary transaction system explains why the banking industry is so heavily regulated. It is harder for a depositor to protect his interests than for an average investor, because judging the financial condition of a bank is difficult and costly, even for specialists. For this reason, the typical instruments adopted by bank regulators include restrictions on the amount of risk that a bank can take, and compulsory deposit insurance schemes that prevent runs.

Regulatory intervention affects the shape of the banking industry and its degree of competition. Until the mid-1960s, governments deliberately limited competition in the interest of 'safety and soundness' by regulating deposit rates, entry, branching and mergers. The traditional view is of a trade-off between soundness and competition, with more intense competition reducing franchise values and increasing incentives to take on risky projects, since forgone future profits in the case of bankruptcy are lower (Keeley, 1990). By increasing the equity at risk, capital controls reduce (although perhaps not entirely) excessive risk-taking (Hellman, Murdock and Stiglitz, 2000).

Recently, a more comprehensive view has been put forward, suggesting that regulation interacts dynamically with pervasive information asymmetries, and that the relationship between competition and stability is accordingly complex and multifaceted (Allen and Gale, 2003). The cost of acquiring information in order to mitigate moral hazard and adverse selection is a strong endogenous barrier to the entry of new banks, allowing incumbents to gain monopoly rents (Broecker, 1990), making competitive equilibria unsustainable (Dell'Ariccia, 2001; Dell'Ariccia, Friedman and Marquez, 1999), and forcing new entrants to take a higher-risk clientele (Shaffer, 1998).

The problems of information asymmetries can be attenuated if a bank deals repeatedly with the same customer, a practice known as 'relationship lending'. However, as Sharpe (1990) and Rajan (1992) show, this gives relationship banks a monopoly on information about their borrowers, further reducing competition, especially in the short run (Petersen and Rajan, 1995). In this case, deregulation aimed at fostering inter-bank competition in transaction lending could have the effect of augmenting the scope for relationship banking, which permits banks to retain some monopoly power. As Boot and Thakor (2000) show, this is not the case if stronger

competition comes from capital market financing, which drives some banks out of the market, reducing competition and consequently relationship lending.

Since the mid-1980s, the banking industry has been transformed by a series of events: deregulation of deposit accounts, which forced US banks to compete on interest rates; branching liberalization, which led to a sharp decline in the number of banks; the changes in capital requirements introduced with the Basel accords of 1988, which pushed banks towards newer and less regulated off-balance-sheet activities; the introduction of the euro, which created a unique wholesale banking market within Europe (Berger, Kashyap and Scalise, 1995); and the substantial repeal of the Glass–Steagall Act of 1933, allowing banks to supply financial services previously offered only by other intermediaries, such as investment firms and insurance companies.

One of most important consequences of deregulation has been the unprecedented numbers of mergers and acquisitions during the 1990s, which sharply reduced the number of banks in many industrial countries and often heightened concern over possible anti-competitive effects. However, there is no clear evidence that the consolidations have harmed consumers or diminished competition, as would have been predicted from the observed negative correlation between the degree of concentration in local banking markets and the level of deposit rates (Berger and Hannan, 1989). Rather, the available evidence indicates a positive effect stemming from the larger and more efficient banks taking over the smaller and less efficient (Berger, Kashyap and Scalise, 1995; Focarelli, Panetta and Salleo, 2002). And while there may be some contraction of credit to smaller clients due to consolidation, this effect appears to be largely offset by increased lending by other banks (Berger et al., 1998). Indeed, there is evidence that in the medium term mergers increase the efficiency of the target bank, benefiting depositors (Focarelli and Panetta, 2003).

The future of the banking industry is likely be determined by the interaction of three major forces: international competition, innovation in information technology and regulation. At present, all three factors are heightening competition in banking. International competition, while still limited, tends to display the same pattern as domestic consolidation, with larger and more efficient banks in more developed countries taking over less efficient banks in financially less developed areas (Focarelli and Pozzolo, 2005). Technological innovation is lessening the importance of close lending relationships, enlarging the size of local credit markets and further reducing the role of small banks (Petersen and Rajan, 2002). Worldwide regulatory systems are moving to allow more competition and to assign a more important role to market evaluation (Basel Committee on Banking Supervision, 2005).

DARIO FOCARELLI AND ALBERTO FRANCO POZZOLO

*See Also* **agency problems; banking crises; financial intermediation; market structure; merger analysis (United States); micro-credit; payment systems.**

### *Bibliography*

Allen, F. and Gale, D. 2003. Competition and financial stability. *Journal of Money, Credit, and Banking* 36, 433–80.

Basel Committee on Banking Supervision. 2005. *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Basel: BIS.

Berger, A. and Hannan, T. 1989. The price–concentration relationship in banking. *Review of Economics and Statistics* 71, 291–9.
Berger, A., Kashyap, A. and Scalise, J. 1995. The transformation of the US banking industry: what a long trip it's been. *Brookings Papers on Economic Activity* 1995(2), 55–201.
Berger, A., Saunders, A., Scalise, J. and Udell, G. 1998. The effects of bank mergers and acquisitions on small business lending. *Journal of Financial Economics* 50, 187–229.
Boot, A. and Thakor, A. 2000. Can relationship banking survive competition? *Journal of Finance* 55, 679–713.
Broecker, T. 1990. Credit-worthiness tests and interbank competition. *Econometrica* 58, 429–52.
Dell'Ariccia, G. 2001. Asymmetric information and the structure of the banking industry. *European Economic Review* 45, 1957–80.
Dell'Ariccia, G., Friedman, E. and Marquez, R. 1999. Adverse selection as a barrier to entry in the banking industry. *RAND Journal of Economics* 30, 515–34.
Diamond, D. 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51, 393–414.
Diamond, D. and Dybvig, P. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91, 401–19.
Diamond, D. and Rajan, R. 2001. Liquidity risk, liquidity creation and financial fragility: a theory of banking. *Journal of Political Economy* 109, 287–327.
Fama, E. 1980. Banking in the theory of finance. *Journal of Monetary Economics* 6, 39–57.
Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15, 29–34.
Focarelli, D. and Panetta, F. 2003. Are mergers beneficial to consumers? Evidence from the market for bank deposits. *American Economic Review* 93, 1152–72.
Focarelli, D., Panetta, F. and Salleo, C. 2002. Why do banks merge? *Journal of Money, Credit, and Banking* 34, 784–803.
Focarelli, D. and Pozzolo, A. 2005. Where do banks expand abroad? An empirical analysis. *Journal of Business* 78, 2435–64.
Gorton, G. and Winton, A. 2003. Financial intermediation. In *Handbook of the Economics of Finance*, vol. 1, ed. G. Constantinides, M. Harris and R. Stulz. Amsterdam: North-Holland.
Hellman, T., Murdock, K. and Stiglitz, J. 2000. Liberalization, moral hazard in banking and prudential regulation: are capital requirements enough? *American Economic Review* 90, 147–65.
Leland, H. and Pyle, D. 1977. Informational asymmetries, financial structure and financial intermediation. *Journal of Finance* 32, 371–87.
Keeley, M. 1990. Deposit insurance, risk, and market power in banking. *American Economic Review* 80, 1183–200.
Petersen, M. and Rajan, R. 1995. The effect of credit market competition on lending relationships. *Quarterly Journal of Economics* 110, 407–43.
Petersen, M. and Rajan, R. 2002. Does distance still matter? The information revolution in small business lending. *Journal of Finance* 57, 2533–70.
Rajan, R. 1992. Insiders and outsiders: the choice between relationship and arms length debt. *Journal of Finance* 47, 1367–400.
Shaffer, S. 1998. The winner's curse in banking. *Journal of Financial Intermediation* 7, 359–92.
Sharpe, S. 1990. Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships. *Journal of Finance* 45, 1069–87.

# bubbles

Bubbles are typically associated with dramatic asset price increases followed by a collapse. Bubbles arise if the price exceeds the asset's fundamental value. This can occur if investors hold the asset because they believe that they can sell it at a higher price than some other investor even though the asset's price exceeds its fundamental value. Famous historical examples are the Dutch tulip mania (1634–7), the Mississippi Bubble (1719–20), the South Sea Bubble (1720), and the 'Roaring '20s' that preceded the 1929 crash. More recently, up to March 2000 Internet share prices (CBOE Internet Index) surged to astronomical heights before plummeting by more than 75 per cent by the end of 2000.

Since asset prices affect the real allocation of an economy, it is important to understand the circumstances under which these prices can deviate from their fundamental value. Bubbles have long intrigued economists and led to several strands of models, empirical tests and experimental studies.

We can broadly divide the literature into four groups. The first two groups of models analyse bubbles within the rational expectations paradigm, but differ in their assumption as to whether all investors have the same information or are asymmetrically informed. A third group of models focuses on the interaction between rational and non-rational (behavioural) investors. In the final group of models traders' prior beliefs are heterogeneous, possibly due to psychological biases, and consequently they agree to disagree about the fundamental value of the asset.

### Rational bubbles under symmetric information

Rational bubbles under symmetric information are studied in settings in which all agents have rational expectations and share the same information. There are several theoretical arguments that allow us to rule out rational bubbles under certain conditions. Tirole (1982) uses a general equilibrium reasoning to argue that bubbles cannot exist if it is commonly known that the initial allocation is interim Pareto efficient. A bubble would make the seller of the 'bubble asset' better off, which – due to interim Pareto efficiency of the initial allocation – has to make the buyer of the asset worse off. Hence, no individual would be willing to buy the asset. Partial equilibrium arguments alone are also useful in ruling out bubbles. Simply rearranging the definition of (net) return, $r_{t+1,s} = (p_{t+1,s} + d_{t+1,s})/p_t - 1$, where $p_{t,s}$ is the price and $d_{t,s}$ is the dividend payment at time $t$ and state $s$, and taking rational expectations yields

$$p_t = E_t\left[\frac{p_{t+1} + d_{t+1}}{1 + r_{t+1}}\right]. \tag{1}$$

That is, the current price is just the discounted expected future price and dividend payment in the next period. For tractability assume that the expected return that the marginal rational trader requires in order to hold the asset is constant over time, $E_t[r_{t+1}] = r$, for all $t$. In solving the above difference equation forward, that is, in

replacing $p_{t+1}$ with $E_{t+1}[p_{t+2} + d_{t+2}]/(1+r)$ in eq. (1) versus Equation (2) below and then $p_{t+2}$ and so on, and using the law of iterated expectations, one obtains after $T-t-1$ iterations

$$p_t = E_t\left[\sum_{\tau=1}^{T-t} \frac{1}{(1+r)^{\tau}} d_{t+\tau}\right] + E_t\left[\frac{1}{(1+r)^{T-t}} p_T\right].$$

The equilibrium price is given by the expected discounted value of the future dividend stream paid from $t+1$ to $T$ plus the expected discounted value of the price at $T$. For *securities with finite maturity*, the price after maturity, say $T$, is zero, $p_T = 0$. Hence, the price of the asset, $p_t$, is unique and simply coincides with the expected future discounted dividend stream until maturity. Put differently, finite horizon bubbles cannot arise as long as rational investors are unconstrained from selling the desired number of shares in all future contingencies. For *securities with infinite maturity*, $T \to \infty$, the price $p_t$ only coincides with the expected discounted value of the future dividend stream, call it fundamental value, $v_t$, if the so-called transversality condition, $lim_{T\to\infty} E_t[\frac{1}{(1+r)^{T-t}} p_T] = 0$, holds. Without imposing the transversality condition, $p_t = v_t$ is only one of many possible prices that solve the above expectational difference equation. Any price $p_t = v_t + b_t$, decomposed in the fundamental value, $v_t$, and a bubble component, $b_t$, such that

$$b_t = E_t\left[\frac{1}{(1+r)} b_{t+1}\right], \tag{2}$$

is also a solution. Equation (2) versus eq. (1) needs to be made consistent. Equation (2) highlights that the bubble component $b_t$ has to 'grow' in expectations exactly at a rate of $r$. A nice example of these 'rational bubbles' is provided in Blanchard and Watson (1982), where the bubble persists in each period only with probability $\pi$ and bursts with probability $(1-\pi)$. If the bubble continues, it has to grow in expectation by a factor $(1+r)/\pi$. This faster bubble growth rate (conditional on not bursting) is necessary to achieve an expected growth rate of $r$. In general, the bubble component may be stochastic. A specific example of a stochastic bubble is an intrinsic bubble, where the bubble component is assumed to be deterministically related to a stochastic dividend process.

The fact that any bubble has to grow at an expected rate of $r$ allows one to eliminate many potential rational bubbles. For example, a positive bubble cannot emerge if there is an upper limit on the size of the bubble. That is, for example, the case with potential bubbles on commodities with close substitutes. An ever-growing 'commodity bubble' would make the commodity so expensive that it would be substituted with some other good. Similarly, a bubble on a non-zero net supply asset cannot arise if the required return $r$ exceeds the growth rate of the economy, since the bubble would outgrow the aggregate wealth in the economy. Hence, bubbles can only exist in a world in which the required return is lower than or equal to the growth rate of the economy. In addition, rational bubbles can persist if the pure existence of the bubble enables trading opportunities that lead to a different equilibrium allocation. Fiat money in an overlapping generations (OLG) model is probably the most famous example of such a bubble.

The intrinsic value of fiat money is zero, yet it has a positive price. Moreover, only when the price is positive, does it allow wealth transfers across generations (that might not even be born yet). A negative bubble, $b_t < 0$, on a limited-liability asset cannot arise since the bubble would imply that the asset price has to become negative in expectation at some point in time. This result, together with eq. (2), implies that if the bubble vanishes at any point it has to remain zero from that point onwards. That is, rational bubbles can never emerge within an asset-pricing model; they must already be present when the asset starts trading.

*Empirically testing* for rational bubbles under symmetric information is a challenging task. The literature has developed three types of tests: regression analysis, variance bounds tests and experimental tests. Initial tests proposed by Flood and Garber (1980) exploit the fact that bubbles cannot start within a rational asset-pricing model and hence at any point in time the price must have a non-zero part that grows at an expected rate of $r$. However using this approach, inference is difficult due to an exploding regressor problem. That is, as time $t$ increases, the regressor explodes and the coefficient estimate relies primarily on the most recent data points. More precisely, the ratio of the information content of the most recent data point to the information content of all previous observations never goes to zero. This implies that as time $t$ increases, the time series sample remains essentially small and the central limit theorem does not apply. Diba and Grossman (1988) test for bubbles by checking whether the stock price is more explosive than the dividend process. Note that if the dividend process follows a linear unit-root process (for example, a random walk), then the price process has a unit root as well. However the change in price, $\Delta p_t$, and the spread between the price and the discounted expected dividend stream, $p_t - d_t/r$, are stationary under the no-bubbles hypothesis. That is, $p_t$ and $d_t/r$ are co-integrated. Diba and Grossman test this hypothesis using a series of unit root tests, autocorrelation patterns, and co-integration tests. They conclude that the no-bubble hypothesis cannot be rejected. However, Evans (1991) shows that these standard linear econometric methods may fail to detect the explosive nonlinear patterns of periodically collapsing bubbles. West (1987) proposes a different test that exploits the fact that one can estimate the parameters needed to calculate the expected discounted value of dividends in two different ways. One way of estimating them is not affected by the bubble, the other is. Note that the accounting identity (1) can be rewritten as $p_t = \frac{1}{1+r}(p_{t+1} + d_{t+1}) - \frac{1}{1+r}(p_{t+1} + d_{t+1} - E_t[p_{t+1} + d_{t+1}])$. Hence, in an instrumental variables regression of $p_t$ on $(p_{t+1} + d_{t+1})$ – using for example $d_t$ as an instrument – one obtains an estimate for $r$ that is independent of the existence of a rational bubble. Second, if, for example, the dividend process follows a stationary AR(1) process, $d_{t+1} = \varphi d_t + \eta_{t+1}$, with independent noise $\eta_{t+1}$, one can easily estimate $\varphi$. Furthermore, the expected discounted value of future dividends is $v_t = (\varphi/(1 + r - \varphi))d_t$. Hence, under the null-hypothesis of no bubble, that is $p_t = v_t$, the coefficient estimate of the regression of $p_t$ on $d_t$ provides a second estimate of $\varphi/(1 + r - \varphi)$. In a final step, West uses a Hausman specification test to test whether both estimates coincide. He finds that the US stock market data usually reject the null hypothesis of no bubble.

Excessive volatility in the stock market seems to provide further evidence in favour of stock market bubbles. LeRoy and Porter (1981) and Shiller (1981) introduced variance bounds that indicate that the stock market is too volatile to be justified by the volatility of the discounted dividend stream. However, the variance bounds test is controversial (see, for example, Kleidon, 1986). Also, this test, as well as all the aforementioned bubble tests, assumes that the required expected returns, $r$, are constant over time. In a setting in which the required expected returns can be time-varying, the empirical evidence favouring excess volatility is less clear-cut. Furthermore, time-varying expected returns can also rationalize the long-horizon predictability of stock returns. For example, a high price–dividend ratio predicts low subsequent stock returns with a high $R^2$ (Campbell and Shiller, 1988).

Finally, it is important to recall that the theoretical arguments that rule out rational bubbles as well as several empirical bubble tests rely heavily on backward induction. Since a bubble cannot grow from time $T$ onwards, there cannot be a bubble of this size at time $T-1$, which rules out this bubble at $T-2$, and so on. However, there is ample *experimental evidence* that individuals violate the backward induction principle. Most convincing are experiments on the centipede game (Rosenthal, 1981). In this simple game, two players alternatively decide whether to continue or stop the game for a finite number of periods. On any move, a player is better off stopping the game than continuing if the other player stops immediately afterwards, but is worse off stopping than continuing if the other player continues afterwards. This game has only a single subgame perfect equilibrium that follows directly from backward induction reasoning. Each player's strategy is to stop the game whenever it is his or her turn to move. Hence, the first player should immediately stop the game and the game should never get off the ground. However, in experiments players initially continue to play the game – a violation of the backward induction principle (see for example, McKelvey and Palfrey, 1992). These experimental findings question the theoretical reasoning used to rule out rational bubbles under symmetric information. More experimental evidence on bubbles in general is provided in the final section.

In a rational bubble setting an investor only holds a bubble asset if the bubble grows in expectations ad infinitum. In contrast, in the following models an investor might hold an overpriced asset if he thinks he can resell it in the future to a less informed trader or someone who holds biased beliefs. In Kindleberger's (2000) terms, the investor thinks he can sell the asset to a greater fool.

## Asymmetric information bubbles

Asymmetric information bubbles can occur in a setting in which investors have different information, but still share a common prior distribution. In these models prices have a dual role: they are an index of scarcity and informative signals, since they aggregate and partially reveal other traders' aggregate information (see for example Brunnermeier, 2001 for an overview). In contrast to the symmetric information case, the presence of a bubble need not be commonly known. For example, it might be the case that everybody knows the price exceeds the value of any possible dividend stream, but it is not the case that everybody knows that all the other investors also know this

fact. It is this lack of higher-order mutual knowledge that makes it possible for finite bubbles to exist under certain necessary conditions (Allen, Morris and Postlewaite, 1993). First, it is crucial that investors remain asymmetrically informed even after inferring information from prices and net trades. This implies that prices cannot be fully revealing. Second, investors must be constrained from (short) selling their desired number of shares in at least one future contingency for finite bubbles to persist. Third, it cannot be common knowledge that the initial allocation is interim Pareto efficient, since then it would be commonly known that there are no gains from trade and hence the buyer of an overpriced 'bubble asset' would be aware that the rational seller gains at his expense (Tirole, 1982). In other words, there have to be gains from trade or at least some investors have to think that there might be gains from trade. There are various mechanisms that lead to these. For example, fund managers who invest on behalf of their clients can gain from buying overpriced bubble assets, since trading allows them to fool their clients into believing that they have superior trading information. A fund manager who does not trade would reveal that he does not have private information. Consequently, bad fund managers churn bubbles at the expense of their uninformed client investors (Allen and Gorton, 1993). Furthermore, fund managers with limited liability might trade bubble assets due to classic risk-shifting incentives, since they participate on the potential upside of a trade but not on the downside risk.

### Bubbles due to limited arbitrage

Bubbles due to limited arbitrage arise in models in which rational, well-informed and sophisticated investors interact with behavioural market participants whose trading motives are influenced by psychological biases. Proponents of the 'efficient markets hypothesis' argue that bubbles cannot persist since well-informed sophisticated investors will undo the price impact of behavioural non-rational traders. Thus, rational investors should go against the bubble even before it emerges. The literature on limits to arbitrage challenges this view. It argues that bubbles can persist, and provides three channels that prevent rational arbitrageurs from fully correcting the mispricing. First, *fundamental risk* makes it risky to short a bubble asset since a subsequent positive shift in fundamentals might *ex post* undo the initial overpricing. Risk aversion limits the aggressiveness of rational traders if close substitutes and close hedges are unavailable. Second, rational traders also face *noise trader risk* (DeLong et al., 1990). Leaning against the bubble is risky even without fundamental risk, since irrational noise traders might push up the price even further in the future and temporarily widen the mispricing. Rational traders with short horizons care about prices in the near future in addition to the long-run fundamental value and only partially correct the mispricing. For example, in a world with delegated portfolio management, fund managers are often concerned about short-run price movements, because temporary losses instigate fund outflows (Shleifer and Vishny, 1997). A temporary widening of the mispricing and the subsequent outflow of funds force fund managers to unwind their positions exactly when the mispricing is the largest. Anticipating this possible scenario, mutual fund managers trade less aggressively against the mispricing. Similarly, hedge funds face a high flow-performance sensitivity, despite some arrangements designed to prevent

outflows (for example, lock-up provisions). Third, rational traders face *synchronization risk* (Abreu and Brunnermeier, 2002, 2003). Since a single trader alone cannot typically bring the market down by himself, coordination among rational traders is required and a synchronization problem arises. Each rational trader faces the following trade-off: if he attacks the bubble too early, he forgoes profits from the subsequent run-up caused by behavioural momentum traders; if he attacks too late and remains invested in the bubble asset, he will suffer from the subsequent crash. Each trader tries to forecast when other rational traders will go against the bubble. Timing other traders' moves is difficult because traders become sequentially aware of the bubble, and they do not know where in the queue they are. Because of this 'sequential awareness', it is never common knowledge that a bubble has emerged. It is precisely this lack of common knowledge that removes the bite of the standard backward induction argument. Since there is no commonly known point in time from which one could start backward induction, even finite horizon bubbles can persist. The other important message of the theoretical work on synchronization risk is that relatively insignificant news events can trigger large price movements, because even unimportant news events allow traders to synchronize their sell strategies. Unlike the earlier limits to arbitrage models, in which rational traders do not trade aggressively enough to completely eradicate the bubble but still short an overpriced bubble asset, in Abreu and Brunnermeier (2003) rational traders prefer to ride the bubble rather than attack it. The incentive to ride the bubble stems from a predictable 'sentiment' in the form of continuing bubble growth.

*Empirically*, there is supportive evidence in favour of the 'bubble-riding hypothesis'. For example, between 1998 and 2000 hedge funds were heavily tilted towards highly priced technology stocks (Brunnermeier and Nagel, 2004). Contrary to the efficient markets hypothesis, hedge funds were not a price-correcting force even though they are among the most sophisticated investors and are arguably closer to the ideal of 'rational arbitrageurs' than any other class of investors. Similarly, Temin and Voth (2004) document that Hoares Bank was profitably riding the South Sea bubble in 1719–20, despite giving numerous indications that it believed the stock to be overvalued. Many other investors, including Isaac Newton, also tried to ride the South Sea bubble but with less success. Frustrated with his trading experience, Isaac Newton concluded 'I can calculate the motions of the heavenly bodies, but not the madness of people' (Kindleberger, 2005, p. 41).

### Heterogeneous beliefs bubbles

Bubbles can also emerge when investors have heterogeneous beliefs and face short-sale constraints. Investors' beliefs are heterogeneous if they start with different prior belief distributions that can be due to psychological biases. For example, if investors are overconfident about their own signals, they have a different prior distribution (with lower variance) about the signals' noise term. Investors with non-common priors can agree to disagree even after they share all their information. Also, in contrast to an asymmetric information setting, investors do not try to infer other traders' information from prices. Combining heterogeneous beliefs with short-sale constraints can result in overpricing since optimists push up the asset price, while pessimists cannot counterbalance it since they face short-sale constraints (Miller, 1977). Ofek and Richardson (2003) link this

argument to the Internet bubble of the late 1990s. In a dynamic model, the asset price can even exceed the valuation of the most optimistic investor in the economy. This is possible, since the currently optimistic investors – the current owners of the asset – have the option to resell the asset in the future at a high price whenever they become less optimistic. At that point other traders will be more optimistic, and hence be willing to buy the asset since optimism is assumed to oscillate across different investor groups (Harrison and Kreps, 1978). It is essential that less optimistic investors, who would like to short the asset, are prevented from doing so by the short-sale constraint. Heterogeneous belief bubbles are accompanied by large trading volume and high price volatility (Scheinkman and Xiong, 2003).

### Experimental evidence

Many theoretical arguments in favour of or against bubbles are difficult to test with (confounded) field data. Laboratory experiments have the advantage that they allow the researcher to isolate and test specific mechanisms and theoretical arguments. For example, the aforementioned experimental evidence on centipede games questions the validity of backward induction. There is a large and growing literature that examines bubbles in a laboratory setting. For example, Smith, Suchanek and Williams (1988) study a double-auction setting, in which a risky asset pays a uniformly distributed random dividend of $d \in \{0, d_1, d_2, d_3\}$ in each of the 15 periods. Hence, the fundamental value for a risk-neutral trader is initially $15\Sigma_i \frac{1}{4} d_i$ and declines by $\Sigma_i \frac{1}{4} d_i$ in each period. Even though there is no asymmetric information and the probability distribution is commonly known, there is vigorous trading, and prices initially rise despite the fact that the fundamental value steadily declines. More specifically, the time-series of asset prices in the experiments are characterized by three phases. An initial boom phase is followed by a period during which the price exceeds the fundamental value, before the price collapses towards the end. These findings are in sharp contrast to any theoretical prediction and seem very robust across various treatments. A string of subsequent articles show that bubbles still emerge after allowing for short sales, after introducing trading fees, and when using professional business people as subjects. Only the introduction of futures markets and the repeated experience of a bubble reduce the size of the bubble. Researchers have speculated that bubbles emerge because each trader hopes to outwit others and to pass the asset on to some less rational trader in the final trading rounds. However, more recent research has revealed that the lack of common knowledge of rationality is not the cause of bubbles. Even when investors have no resale option and are forced to hold the asset until the end, bubbles still emerge (Lei, Noussair and Plott, 2001).

In summary, the literature on bubbles has taken giant strides since the 1970s that led to several classes of models with distinct empirical tests. However, many questions remain unresolved. For example, we do not have many convincing models that explain when and why bubbles start. Also, in most models bubbles burst, while in reality bubbles seem to deflate over several weeks or even months. While we have a much better idea of why rational traders are unable to eradicate the mispricing introduced by behavioural traders, our understanding of behavioural biases and belief distortions is less advanced. From a policy perspective, it is interesting to answer the question whether

central banks actively try to burst bubbles. I suspect that future research will place greater emphasis on these open issues.

MARKUS K. BRUNNERMEIER

*See Also* **behavioural finance; Kindleberger, Charles P.; South Sea bubble; speculative bubbles; tulipmania.**

## *Bibliography*


Abreu, D. and Brunnermeier, M.K. 2002. Synchronization Risk and Delayed Arbitrage. *Journal of Financial Economics* 66, 341–60.

Abreu, D. and Brunnermeier, M.K. 2003. Bubbles and crashes. *Econometrica* 71, 173–204.

Allen, F. and Gorton, G. 1993. Churning bubbles. *Review of Economic Studies* 60, 813–36.

Allen, F., Morris, S. and Postlewaite, A. 1993. Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory* 61, 206–29.

Blanchard, O.J. and Watson, M.W. 1982. Bubbles, rational expectations, and financial markets. In *Crisis in the Economic and Financial Structure*, ed. P. Wachtel. Lexington, MA: Lexington.

Brunnermeier, M.K. 2001. *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis and Herding*. Oxford: Oxford University Press.

Brunnermeier, M.K. and Nagel, S. 2004. Hedge funds and the technology bubble. *Journal of Finance* 59, 2013–40.

Campbell, J.Y. and Shiller, R.J. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195–228.

DeLong, J.B., Shleifer, A., Summers, L.H. and Waldmann, R.J. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98, 703–38.

Diba, B.T. and Grossman, H.I. 1988. The theory of rational bubbles in stock prices. *Economic Journal* 98, 746–54.

Evans, G.W. 1991. Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review* 81, 922–30.

Flood, R.P. and Garber, P.M. 1980. Market fundamentals versus price-level bubbles: the first tests. *Journal of Political Economy* 88, 745–70.

Harrison, J.M. and Kreps, D. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 89, 323–36.

Kindleberger, C.P. 2005. *Manias, Panics and Crashes: A History of Financial Crises*, 5th edn. New york: Wiley.

Kleidon, A.W. 1986. Variance bounds tests and stock price valuation models. *Journal of Political Economy* 94, 953–1001.

Lei, V., Noussair, C.N. and Plott, C.R. 2001. Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs. actual irrationality. *Econometrica* 69, 831–59.

LeRoy, S.F. and Porter, R.D. 1981. The present value relation: tests based on implied variance bounds. *Econometrica* 64, 555–74.

McKelvey, R.D. and Palfrey, T.R. 1992. An experimental study of the centipede game. *Econometrica* 60, 803–36.

Miller, E.M. 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32, 1151–68.

Ofek, E. and Richardson, M. 2003. DotCom mania: the rise and fall of Internet stocks. Working Paper No. FIN-01-037 58(3), 1113–38. New York University, Stern School.

Rosenthal, R. 1981. Games of perfect information, predatory pricing, and the chain-store paradox. *Journal of Economic Theory* 25, 92–100.

Scheinkman, J. and Xiong, W. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111, 1183–219.

Shiller, R.J. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–36.

Shleifer, A. and Vishny, R.W. 1997. The limits of arbitrage. *Journal of Finance* 52, 35–55.
Smith, V.L., Suchanek, G.L. and Williams, A.W. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica* 56, 1119–51.
Temin, P. and Voth, H.-J. 2004. Riding the South Sea Bubble. *American Economic Review* 94, 1654–68.
Tirole, J. 1982. On the possibility of speculation under rational expectations. *Econometrica* 50, 1163–82.
West, K.D. 1987. A specification test for speculative bubbles. *Quarterly Journal of Economics* 102, 553–80.

## bubbles in history

A bubble may be defined loosely as a sharp rise in price of an asset or a range of assets in a continuous process, with the initial rise generating expectations of further rises and attracting new buyers – generally speculators interested in profits from trading in the asset rather than its use or earning capacity. The rise is usually followed by a reversal of expectations and a sharp decline in price often resulting in financial crisis. A boom is a more extended and gentler rise in prices, production and profits than a bubble, and may be followed by crisis, sometimes taking the form of a crash (or panic) or alternatively by a gentle subsidence of the boom without crisis.

Bubbles have existed historically, at least in the eyes of contemporary observers, as well as booms so intense and excited that they have been called 'manias'. The most notable bubbles were the Mississippi bubble in Paris in 1719–20, set in motion by John Law, founder of the *Banque Générale* and the *Banque Royale*, and the contemporaneous and related South Sea bubble in London. Most famous of the manias were the Tulip mania in Holland in 1636, and the Railway mania in England in 1846–7. It is sometimes debated whether a particular sharp rise and fall in prices, such as the German hyperinflation from 1920 to 1923, or the rise and fall in commodity and share prices in London and New York in 1919–21, the rise of gold of $850 an ounce in 1982 and its subsequent fall to the $350 level, were or were not bubbles. Some theorists go further and question whether bubbles are possible with rational markets, which they assume exist (see e.g. Flood and Garber, 1980).

Rational expectations theory holds that prices are formed within the limits of available information by market participants using standard economic models appropriate to the circumstances. As such, it is claimed, market prices cannot diverge from fundamental values unless the information proves to have been widely wrong. The theoretical literature uses the assumption of the market having one mind and one purpose, whereas it is observed historically that market participants are often moved by different purposes, operate with different wealth and information and calculate within different time horizons. In early railway investment, for example, initial investors were persons doing business along the rights of way who sought benefits from the railroad for their other concerns. They were followed by a second group of investors interested in the profits the railroad would earn, and by a third group, made up of speculators who, seeing the rise in the railroad's shares, borrowed money or paid for the initial instalments with no intention of completing the purchase, to make a profit on resale.

The objects of speculation resulting in bubbles or booms and ending in numerous cases, but not all, in financial crisis, change from time to time and include commodities, domestic bonds, domestic shares, foreign bonds, foreign shares, urban and suburban real estate, rural land, leisure homes, shopping centres, Real Estate Investment Trusts, 747 aircraft, supertankers, so-called 'collectibles' such as paintings, jewellery, stamps, coins, antiques etc. and, most recently, syndicated bank loans to developing countries. Within these relatively broad categories, speculation may fix on

particular objects – insurance shares, South American mining stocks, cotton-growing land, Paris real estate, Post-Impressionist art, and the like.

At the time of writing, the theoretical literature has yet to converge on an agreed definition of bubbles, and on whether they are possible. Virtually the same authors who could not reject the no-bubbles hypothesis in the German inflation of 1923 one year, managed to do so a year later (Flood and Garber, 1980). Another pair of theorists has demonstrated mathematically that rational bubbles can exist after putting aside 'irrational bubbles' on the grounds not of their non-existence but of the difficulty of the mathematics involved (Blanchard and Watson, 1982).

Short of bubbles, manias and irrationality are periods of euphoria which produce positive feedback, price increases greater than justified by market fundamentals, and booms of such dimensions as to threaten financial crisis, with possibilities of a crash or panic. Minsky (1982a, 1982b) has discussed how after an exogenous change in economic circumstances has altered profit opportunities and expectations, bank lending can become increasingly lax by rigorous standards. Critical exception has been taken to his taxonomy dividing bank lending into hedge finance, to be repaid out of anticipated cash flows; speculative finance, requiring later refinancing because the term of the loan is less than the project's payoff; and Ponzi finance, in which the borrower expects to pay off his loan with the proceeds of sale of an asset. It is objected especially that Carlo Ponzi was a swindler and that many loans of the third type, for example those to finance construction, are entirely legitimate (Flemming, Goldsmith and Melitz, 1982). Nonetheless, the suggestion that lending standards grow more lax during a boom and that the banking system on that account becomes more fragile has strong historical support. It is attested, and the contrary rational-expectations view of financial markets is falsified, by the experience of such a money and capital market as London having successive booms, followed by crisis, the latter in 1810, 1819, 1825, 1836, 1847, 1857, 1866, 1890, 1900, 1921 – a powerful record of failing to learn from experience (Kindleberger, 1978).

CHARLES P. KINDLEBERGER

*See Also* **tulipmania.**

### *Bibliography*

Blanchard, O. and Watson, M.W. 1982. Bubbles, rational expectations and financial markets. In *Crises in the Economic and Financial Structure*, ed. P. Wachtel. Lexington, Mass.: Heath.

Flemming, J.S., Goldsmith, R.W. and Melitz, J. 1982. Comment. In *Financial Crises: Theory, History and Policy*, ed. C.P. Kindleberger and J.-P. Laffargue. Cambridge: Cambridge University Press.

Flood, R.P. and Garber, P.M. 1980. Market fundamentals versus price-level bubbles: the first tests. *Journal of Political Economy* 88, 745–70.

Kindleberger, C.P. 1978. *Manias, Panics and Crashes: A History of Financial Crises*. New York: Basic Books.

Minsky, H.P. 1982a. *Can 'It' Happen Again?: Essays on Instability and Finance*. Armonk: Sharpe.

Minsky, H.P. 1982b. The financial instability hypothesis. In *Financial Crises: Theory, History and Policy*, ed. C.P. Kindleberger and J.-P. Laffargue. Cambridge: Cambridge University Press.

# capital controls

Capital controls are any restrictions on the movement of capital into or out of a country. Capital controls can take a wide variety of forms. For example, capital controls can be quantity-based or price-based, or apply to only capital inflows, only capital outflows, or all types of capital flows. Capital controls can also be directed at different types of capital flows (such as at bank loans, foreign direct investment or portfolio investment) or at different types of actors (such as at companies, banks, governments or individuals).

Most developed countries believe that the benefits from the free movement of capital across borders outweigh the costs, and therefore have very limited (if any) capital controls in place today. For emerging markets and developing economies, however, there has been a long-standing debate on the desirability of capital controls. Assessing the impact of capital controls is complicated due to a number of factors, including the various forms in which they can be structured. This article discusses the recent debate on capital controls, focusing on the theoretical arguments for and against controls and the existing empirical evidence on their impact.

## History of the debate

Throughout the 20th century, economists have regularly expressed concerns about international capital flows. For example, in the 1940s Ragnar Nurkse worried about 'destabilizing capital flows' and in the 1970s Charles Kindleberger described the role of capital in driving 'manias, panics and crashes' (see Nurkse, 1944; Kindleberger, 1978). When the world's leading economies met at Bretton Woods in 1944 to formulate rules governing the international financial system, John Maynard Keynes and other delegates debated the role of capital controls. The resulting compromise required that members of the International Monetary Fund (IMF), one of the newly created international monetary institutions, allow capital to be freely exchanged and convertible across countries for the purpose of all current account transactions, but permitted members to implement capital controls for financial account transactions. Most countries had capital controls in place at this time.

Over the following years, however, many developed countries gradually removed their capital controls, so that by the 1980s most had few controls in place. In the early and mid-1990s, many emerging markets and developing countries also began to lift their capital controls. The impact initially appeared to be positive – capital flowed into countries with liberalized capital accounts, investment and growth increased, and asset prices rose. In fact, support for lifting capital controls was so widespread that in 1996–7 leading policymakers discussed amending the rules agreed to at Bretton Woods to extend the IMF's jurisdiction to include capital movements and make capital account liberalization a goal of the IMF. In mid-1997, however, a series of financial crises started in Asia and spread across the world, appearing to disproportionately affect emerging markets that had recently liberalized their capital accounts. This series of

crises sparked a reassessment of the desirability of capital controls for emerging markets and developing economies.

In a sharp sea change, many leading policymakers and economists began to support the use of capital controls for emerging markets in some circumstances, especially taxes on capital inflows. Much of this support was based on the belief that controls on capital inflows could reduce a country's vulnerability to financial crises. From 2002 to 2005, several emerging markets (such as Colombia, Russia and Venezuela) also implemented new controls on capital inflows, largely to reduce the appreciations of their currencies. Over the same period, however, several large emerging markets (such as India and China) moved in the opposite direction and lifted many of their existing controls.

## Benefits and costs of capital controls

The free movement of capital across borders can have widespread benefits. Capital inflows can provide financing for high-return investment, thereby raising growth rates. Capital inflows – especially in the form of direct investment – often bring improved technology, management techniques, and access to international networks, all of which further raise productivity and growth. Capital outflows can allow domestic citizens and companies to earn higher returns and better diversify risk, thereby reducing volatility in consumption and income. Capital inflows and outflows can increase market discipline, thereby leading to a more efficient allocation of resources and higher productivity growth. Implementing capital controls can reduce a country's ability to realise these multifaceted benefits.

On the other hand, the free movement of capital across borders can also have costs. Countries reliant on foreign financing will be more vulnerable to 'sudden stops' in capital inflows, which can cause financial crises and/or major currency depreciations. Large volumes of capital inflows can cause currencies to appreciate and undermine export competitiveness, causing what is often called the 'Dutch disease'. The free movement of capital can also complicate a country's ability to pursue an independent monetary policy, especially when combined with a fixed exchange rate. Finally, capital inflows may be invested inefficiently due to a number of market distortions, thereby leading to overinvestment and bubbles that create additional challenges. Capital controls could potentially reduce these costs from the free movement of capital.

## Empirical evidence on capital controls

Since capital controls can have costs and benefits, evaluating the desirability and aggregate impact of capital controls is largely an empirical question. (See Eichengreen, 2003, on the potential costs and benefits of capital controls.) Not surprisingly, an extensive literature has attempted to measure and assess the effects of capital controls.

The most studied experience with capital controls is the Chilean *encaje* – a market-based tax on capital inflows from 1991 to 1998 so structured that the magnitude of the tax decreased with the maturity of the capital flow. Chile's experience with capital controls is generally viewed positively, largely due to Chile's strong economic performance during the period the controls were in place. Empirical studies of the impact of Chile's capital controls, however, have reached several general conclusions. First, there is no

evidence that the capital controls moderated the appreciation of Chile's currency (which was the primary purpose of the capital controls). Second, there is little evidence that the controls protected Chile from external shocks. Third, there is some evidence that the controls raised domestic interest rates (at least in the short term). Fourth, there is some evidence that the controls did not affect the volume of capital inflows, but did lengthen the maturity of capital inflows. Finally, the capital controls significantly raised the cost of financing for small and medium-sized firms and distorted the mechanisms by which Chilean companies procured financing. The general conclusion from this work is that Chile's strong economic performance during the 1990s resulted from sound macroeconomic and financial policies, not the capital controls, and that the capital controls had both costs and benefits. (See Forbes, 2007, for more information on this literature and the Chilean capital controls.)

A second major branch of literature examining the impact of capital controls focuses on the effects of lifting capital controls (that is, capital account liberalization). The majority of this work uses macroeconomic data, typically focusing on how capital account liberalization raises economic growth using cross-country growth regressions. Prasad et al. (2003) is a detailed survey of this literature and shows that, although several papers find a robust, positive effect of capital account liberalization on growth, other papers find no significant effect, and most papers find mixed evidence. This literature is generally read as showing weak evidence that lifting capital controls may have some positive effect on growth.

There are several explanations for the inconclusive results in this macroeconomic literature assessing the impact of capital controls. First, it is extremely difficult to measure capital account openness and to capture the various types of capital controls in a simple measure that can be used for empirical analysis. Second, different types of capital flows and controls may have different effects on growth and other macroeconomic variables. For example, controls on portfolio investment may be more beneficial than other types of capital controls. Third, the impact of removing capital controls could depend on a range of other factors that are difficult to capture in cross-country regressions, such as a country's institutions, financial system, corporate governance or even the sequence in which different controls are removed. Fourth, capital controls can be very difficult to enforce (especially for countries with undeveloped financial markets) so the same capital control may have different degrees of effectiveness in different countries. Finally, most countries that remove their capital controls undertake simultaneously a range of reforms and undergo structural changes, so that it can be difficult to isolate the impact of removing the controls. (For additional details on the challenges in measuring the impact of capital controls, see Eichengreen, 2003; Forbes, 2006; Magud and Reinhart, 2006; and Prasad et al., 2003.)

Given these challenges in measuring the impact of capital controls, it is not surprising that the empirical literature has had difficulty documenting their effects on growth at the macroeconomic level. To put these results in perspective, however, the current status of this literature is similar to the literature in the 1980s and 1990s on how trade liberalization affects economic growth. Economists generally believe that trade openness raises growth, but most of the initial work on this topic also focused on cross-country, macroeconomic studies and reached inconclusive results. At a much earlier date,

however, several papers using microeconomic data and case studies found compelling evidence that trade liberalization raises productivity and growth.

Similarly, recent work based on microeconomic data has been much more successful than the macroeconomic literature in documenting the effects of capital controls. Forbes (2006) surveys this new literature, which covers a variety of countries and periods, uses a range of approaches and methodologies, and builds on several different fields. This literature has, to date, reached five general results. First, capital controls reduce the supply of capital, raise the cost of financing, and increase financial constraints – especially for smaller firms and firms without access to international capital markets. Second, capital controls reduce market discipline in financial markets and the government, leading to a more inefficient allocation of capital and resources. Third, capital controls distort decision-making by firms and individuals as they attempt to minimize the costs of the controls, or even evade them outright. Fourth, the effects of capital controls vary across different types of firms and countries, reflecting different pre-existing economic distortions. Finally, capital controls can be difficult and costly to enforce, even in countries with sound institutions and low levels of corruption. Therefore, this series of microeconomic studies suggests that capital controls have widespread and pervasive costs, but has not yet provided significant evidence of the benefits of capital controls.

## Conclusions

The debate on the effects and desirability of capital controls is likely to continue and to motivate new academic research. Most economists agree that countries should gradually lift their capital controls as they grow and develop, and that developed countries should have few (if any) capital controls in place. Most economists also believe that the free movement of capital can have widespread benefits, but that in countries with weak financial systems, poorly developed institutions, and vulnerable macroeconomies the free movement of capital can also generate distortions and increase a country's vulnerability. As a result, emerging markets and developing countries that currently have capital controls should work to address the shortcomings in their economies as they liberalize their capital accounts. There continues to be widespread disagreement, however, on the exact sequencing of these reforms and the optimal pace of capital account liberalization for emerging markets and developing economies.

KRISTIN J. FORBES

*See Also* **international capital flows; international monetary institutions; Kindleberger, Charles P.; Nurkse, Ragnar.**

### *Bibliography*

Eichengreen, B. 2003. *Capital Flows and Crises*. Cambridge, MA: MIT Press.

Forbes, K. 2006. The microeconomic evidence on capital controls: no free lunch. In *Capital Controls and Capital Flows in Emerging Economies: Policies, Practices and Consequences*, ed. S. Edwards. Chicago: University of Chicago Press.

Forbes, K. 2007. One cost of the Chilean capital controls: increased financial constraints for smaller traded firms. *Journal of International Economics*.

Kindleberger, C. 1978. *Manias, Panics and Crashes: A History of Financial Crises*. New York: John Wiley & Sons.

Magud, N. and Reinhart, C. 2006. Capital controls: an evaluation. Working Paper No. 11973. Cambridge, MA: NBER.

Nurkse, R. 1944. *International Currency Experience: Lessons of the Interwar Experience*. Geneva: League of Nations.

Prasad, E., Rogoff, K., Wei, S.-J. and Kose, M. 2003. Effects of financial globalization on developing countries: some empirical evidence. Occasional Paper No. 220. Washington, DC: IMF.

# Credit Crunch Chronology: April 2007–September 2009

**April 2007**

*2nd* – New Century Financial, based in California and second only to HSBC in the US sub-prime mortgage market, filed for Chapter 11 bankruptcy protection, making over 3,200 employees redundant.

**May 2007**

*3rd* – Dillon Read Capital Management, a hedge fund, was forced to shut down following a SFr150m. (US$123m.) first-quarter loss on US sub-prime mortgage investments.

**June 2007**

*25th* – Queen's Walk Investment announced a loss of €67.7m. (US$91m.) in the year ending 31 March, reflecting a decline in the value of its UK and US mortgage-linked securities holdings.

*28th* – Caliber Global Investment, a London-listed fund, announced it would wind down over twelve months following a £4.4m. (US$8.8m.) loss from sub-prime investments.

*29th* – US investment bank Bear Stearns replaced the chairman and chief executive of its asset management business in an effort to restore investor confidence following the collapse of two of its hedge funds invested in the sub-prime mortgage market.

**July 2007**

*3rd* – United Capital Asset Management, a Florida-based hedge fund, suspended investor redemptions following heavy losses in sub-prime bonds and derivatives.

*11th* – Braddock Financial, based in Denver, Colorado closed its US$300m. Galena fund owing to sub-prime losses.

*19th* – Ben Bernanke, chairman of the Federal Reserve, warned that the sub-prime crisis in the USA could cost up to US$100bn.

*27th* – Absolute Capital, an Australian hedge fund, temporarily suspended redemptions for two of its funds.

*31st* – After losing over 50% of its capital, Boston-based hedge fund, Sowood Capital Management, was bought by larger rival, Citadel.

**August 2007**

*1st* – Shares in Australia's Macquarie Bank fell by more than 10% after a warning to investors that its two Fortress funds could lose more than $A300m. (US$256m.).

*1st* – Bear Stearns halted redemptions in a third hedge fund, Asset-Backed Securities, following a rush of withdrawals.

*1st* – German bank IKB was bailed out by rival banks for €8bn. after it was exposed to losses in the US sub-prime sector.

*6th* – American Home Mortgage Investment (AHM), the tenth biggest home loan lender in the USA, filed for Chapter 11 bankruptcy protection.
*9th* – France's largest bank, BNP Paribas, suspended three of its funds exposed to the US sub-prime mortgage market.
*9th* – The European Central Bank (ECB) injected €94.8bn. into the eurozone banking market to stabilize overnight interest rates. The Fed quickly followed the ECB by announcing that it would provide US$12bn. of temporary reserves to the American banking system.
*10th* – Continuing turmoil in the markets forced action from the world's central banks. In total US$120bn. of extra liquidity was pumped into financial markets.
*10th* – The FTSE 100 Index fell by 3.7%, its largest drop in four years.
*13th* – Investment bank Goldman Sachs injected US$3bn. into its Global Equity Opportunities hedge fund.
*16th* – The USA's largest mortgage lender, Countrywide Financial, received an US$11.5bn. lifeline from 40 of the world's largest banks.
*17th* – The US Federal Reserve cut its primary discount rate, the rate at which it lends money to banks, by half a point from 6.25% to 5.75%.
*22nd* – Countrywide Financial received a US$2bn. capital injection from the Bank of America.
*23rd* – US and European banks, including the Bank of America, Citigroup, JP Morgan Chase and Germany's Deutsche Bank, borrowed US$2bn. from the US Federal Reserve to improve credit access.
*23rd* – Lehman Brothers closed its sub-prime mortgage unit, BNC Mortgage, releasing 1,200 workers.
*31st* – President George W. Bush announced plans to help struggling sub-prime mortgage borrowers. Federal Reserve chairman Ben Bernanke pledged to take action to protect the wider economy from market turmoil.

**September 2007**

*6th* – The US Federal Reserve added US$31.25bn. to the US money markets and the ECB lent an extra €42.2bn. to banks.
*10th* – Victoria Mortgages, owned by US private equity group Venturion Capital, was forced into administration, becoming the first UK casualty of the sub-prime crisis.
*13th* – The Bank of England provided emergency financial support to Northern Rock, the UK's fifth largest mortgage lender.
*17th* – UK Chancellor Alistair Darling guaranteed Northern Rock's savings accounts, following several days of a run on the bank's deposits.
*18th* – The US Federal Reserve cut interest rates by half a point from 5.25% to 4.75%.
*20th* – Goldman Sachs announced record profits after hedging that the value of mortgage bonds would fall, despite losing US$1.5bn. from the sub-prime crisis.
*26th* – UK banks shunned the Bank of England's auction of £10bn. worth of three-month loans, an emergency funding facility introduced by Governor Mervyn King.

**October 2007**

*1st* – Swiss bank UBS revealed a writedown of SFr4bn. (US$3.4bn.) on hedge fund losses and exposure to the sub-prime mortgage market. The group announced plans to shed 1,500 jobs.

*5th* – Investment bank Merrill Lynch revealed a third-quarter writedown of US$5.5bn.

*15th* – Citigroup announced a total of US$6.5bn. in writedowns.

*24th* – Merrill Lynch announced US$8.4bn. of losses and writedowns. A quarterly loss of US$2.24bn. was the largest in its history. Stan O'Neal, chief executive, resigned six days later.

*31st* – The US Federal Reserve reduced interest rates from 4.75% to 4.5%.

**November 2007**

*1st* – Swiss bank Credit Suisse revealed a US$1bn. writedown.

*4th* – Citigroup announced further writedowns of US$8–11bn. Charles Prince resigned as chairman and chief executive.

*7th* – US investment bank Morgan Stanley forecast a loss of US$3.7bn. against fourth-quarter revenues.

*9th* – Wachovia, the USA's fourth largest lender, unveiled losses of US$1.1bn. for Oct. owing to the continued decline in value of its mortgage debt.

*13th* – The Bank of America revealed it would write off US$3bn. of bad debts linked to the US sub-prime crisis during the last quarter of 2007 and would inject a further US$600m. into a structured investment vehicle with high exposure to sub-prime mortgages.

*14th* – HSBC, the world's second largest bank, claimed it was writing off US$38m. of loans a day to struggling Americans and raising its sub-prime bad debt provision to US$3.4bn.

*14th* – The Bank of England forecast a sharp slowdown in UK domestic growth in 2008 together with higher inflation.

*15th* – Barclays, the UK's third largest bank, announced a writedown of US$2.6bn. on securities related to the US sub-prime mortgage market, having lost US$1.64bn. in Oct. alone.

*16th* – Northern Rock's Adam Applegarth resigned as chief executive.

*20th* – Shares in Paragon, the UK's third largest buy-to-let mortgage lender, were suspended after falling in value by 50%. It warned shareholders it could face collapse if it could not raise an extra £250m.

*20th* – Freddie Mac, the USA's second largest provider of mortgage financing, announced its largest quarterly loss so far after unveiling US$4.8bn. of bad debts and writedowns.

*27th* – Citigroup agreed to sell shares in its company worth US$7.5bn. to the Abu Dhabi Investment Authority, making it the largest shareholder with a stake of 4.9%.

**December 2007**

*4th* – The Bank of Canada cut interest rates by a quarter of a percentage point from 4.5% to 4.25%.

*6th* – The Bank of England lowered interest rates, from 5.75% to 5.5%.

*6th* –RBS warned investors it expected to write off £1.25bn. as a result of exposure to the US sub-prime mortgage market.
*6th* – President Bush unveiled plans to freeze rates on sub-prime mortgages for the next five years.
*10th* – UBS revealed it had written off a further SFr11.2bn. (US$10bn.) against its US sub-prime mortgage exposure.
*10th* – France's second largest bank, Société Générale, moved to bailout its structured investment vehicle with a credit line of up to US$4.3bn.
*11th* – The US Federal Reserve cut interest rates for the third time in four months, reducing them from 4.5% to 4.25%.
*12th* – Five central banks from the UK, Europe and USA launched a US$110bn. joint cash injection targeting international interbank borrowing markets.
*14th* – Citigroup brought US$49bn. worth of sub-prime debts to keep afloat seven high-risk structured investment vehicles.
*17th* – The US Federal Reserve made US$20bn. available to major banks to ease interbank lending rates as the first part of a plan agreed by five central banks.
*18th* – The Bank of England released £10bn. of funds to UK banks and financial institutions.
*18th* – The ECB injected €348.7bn. (US$502bn.) into banks to help ease credit fears over the Christmas period.
*19th* – US investment bank Morgan Stanley wrote down US$9.4bn. in sub-prime losses. A cash injection of US$5bn. (equating to 9.9% of the bank) was provided by China Investment Corporation (CIC).

**January 2008**

*9th* – The World Bank forecast a 0.3% slowdown in global economic growth to 3.3% in 2008 but claimed growth in China and India would soften the impact.
*9th* – James Cayne, chief executive of US investment bank Bear Stearns, stepped down.
*11th* – Countrywide Financial, the USA's largest mortgage lender, was bought by the Bank of America for US$4bn.
*15th* – Citigroup reported a US$9.8bn. loss for the fourth quarter, the largest in its history. The bank also announced a capital injection of US$6.9bn. from the Government of Singapore Investment Corporation (GIC). In total Citigroup and Merrill Lynch had received over US$21bn. from foreign investors including Saudi Arabia and Kuwait.
*21st* – Stock markets across the world suffered their biggest losses since 11 Sept. 2001, triggered by fears of a looming recession in the USA.
*22nd* – The US Federal Reserve slashed interest rates by 0.75% to 3.5%, its largest cut in over 25 years.
*28th* – European bank Fortis warned that its losses connected to US sub-prime mortgage debt could be as much as €1bn. (US$1.5bn.).
*30th* – The US Federal Reserve cut interest rates by a further 50 basis points from 3.5% to 3.0%.
*31st* – MBIA, the world's largest bond insurer, revealed a US$2.3bn. loss in the fourth quarter.

**February 2008**

*6th* – Wall Street had its worst share losses in over a year, while the UK's FTSE 100 fell by 2.6%.

*7th* – The Bank of England reduced interest rates from 5.5% to 5.25%.

*10th* – Finance ministers from the G7 group of industrialized nations warned of worldwide losses from the US mortgage crisis of up to US$400bn.

*13th* – The Financial Services Agency, Japan's financial watchdog, said Japanese banks had lost a total of 600bn. yen (US$5.6bn.) from the US sub-prime mortgage crisis in the previous 12 months.

*14th* – UBS confirmed it had made a loss of SFr4.4bn. (US$4bn.) in 2007, following US$18.4bn. of writedowns.

*14th* – Commerzbank, Germany's second largest bank, announced writedowns of €774m. (US$1.1bn.), despite record-year profits.

*17th* – UK Chancellor Alistair Darling confirmed mortgage lender Northern Rock would be brought into temporary public ownership.

**March 2008**

*3rd* – HSBC, the UK's largest bank, unveiled total writedowns of US$17.2bn., despite an annual profit increase of 10%.

*5th* – Credit Agricole, France's largest retail bank, announced a loss of €857m. (US$1.3bn.) in the fourth quarter, following a €3.3bn. charge at its Calyon investment banking arm on losses related to the credit crisis.

*6th* – Peloton Partners, a London-based hedge fund, was forced to liquidate its £1bn. ABS Master Fund after failing to meet interest payments on loans taken out to buy assets.

*7th* – Carlyle Capital Corporation, a US$22bn. credit fund owned by US private equity firm Carlyle Group, collapsed.

*7th* – The former chief executives of Merrill Lynch, Citigroup and Countrywide Financial were questioned before a Congressional committee over their large salary and pay-off packages while their firms experienced heavy losses.

*7th* – The US Federal Reserve made available up to US$200bn. of emergency financing in response to 'rapid deterioration' in the credit markets.

*14th* – US investment bank Bear Stearns received emergency funding from JP Morgan Chase with the US Federal Reserve's backing, following a collapse in confidence from its hedge fund clients.

*16th* – Bear Stearns was bought out by JP Morgan Chase for US$236m or US$2 per share, a fraction of its previous value, backed by US$30bn. in loans from the US Federal Reserve.

*16th* – The US Federal Reserve lowered its lending rate to financial institutions by a quarter of a point to 3.25% and created a new lending facility for large investment banks to secure short-term loans.

*18th* – Wall Street investment banks Goldman Sachs and Lehman Brothers reported a halving of profits in the first quarter of 2008. The results were better than expected, boosting shares in both firms.

*31st* – Henry Paulson, the US Treasury Secretary, announced a package of reforms designed to help the Federal Reserve tackle financial market turmoil and improve regulation of the financial system.

**April 2008**

*1st* – UBS revealed a further US$19bn. of asset writedowns on top of the US$18.4bn. already lost in 2007. Chief executive Marcel Ospel resigned.

*7th* – UK mortgage lender Abbey withdrew 100% mortgage deals available to UK borrowers.

*8th* – The IMF warned potential losses from the global credit crunch could reach US$945bn.

*10th* – The Bank of England cut interest rates by a quarter point to 5%.

*14th* – Wachovia, the fourth largest US bank, revealed a US$4.4bn. writedown for the first quarter following a jump in foreclosures in California and Florida.

*16th* – JP Morgan Chase reported a US$5.1bn. writedown for the first quarter against investments in mortgage-backed securities and its portfolio of homeloans.

*17th* – Merrill Lynch unveiled a loss of US$1.96bn. in the first quarter.

*18th* – Citigroup posted its second consecutive quarterly loss, of US$5.1bn., and announced it would cut 9,000 jobs after writing off US$15.1bn. in toxic assets.

*21st* – The Bank of England unveiled a £50bn. plan to aid the UK banks by allowing lenders to exchange potentially risky mortgage debts for government-backed bonds.

*22nd* – RBS, the UK's second largest bank, revealed pre-tax writedowns of £5.9bn. and requested £12bn. from shareholders to rebuild its capital base.

*24th* – Credit Suisse reported a quarterly loss of SFr2.5bn. (US$2.1bn.), its first loss in nearly five years, following asset writedowns of US$5.2bn.

*30th* – Nationwide Building Society recorded the first annual fall in UK house prices for ten years, with prices 1% lower in April than the previous year.

**May 2008**

*2nd* – The US Federal Reserve, European Central Bank and Swiss National Bank expanded liquidity by injecting an extra US$82bn. into the banking system.

*12th* – HSBC announced it had written off US$3.2bn. in the first quarter as a result of the sub-prime crisis.

*13th* – UK bank Alliance & Leicester disclosed a £391m. writedown in the first quarter.

*14th* – UK mortgage lender Bradford & Bingley launched an emergency £300m. rights issue.

*15th* – Barclays revealed a further £1.7bn. in writedowns.

*22nd* – Swiss bank UBS launched a SFr16bn. (US$15.5bn.) rights issue to cover its US$37bn. writedowns.

**June 2008**

*19th* – Chicago-based firm Hedge Fund Research showed 170 funds had been forced into liquidation during the first quarter, while fewer funds were launched than at any time since 2000.

*19th* – Two former managers of US investment bank Bear Stearns were charged with fraud. It was alleged they had misled investors about the health of their hedge funds.
*25th* – Major new investors in Barclays, including the Qatar Investment Authority, invested £1.7bn. (US$3.3bn.) for a 7.7% share in the business.

**July 2008**

*8th* – A quarterly survey of businesses by the British Chambers of Commerce (BCC) found that the UK faced a serious risk of recession.
*10th* – Share prices in the USA's two largest mortgage finance companies, Fannie Mae and Freddie Mac, plummeted by nearly 50% as investor anxiety grew over government intervention that would leave their stock worthless.
*11th* – The FTSE 100 fell deep into a bear market (a 20% fall from its market peak in June 2007) as blue-chip stocks reached their lowest level since 31 Oct. 2005.
*13th* – US mortgage lender IndyMac Bank, based in California, collapsed, becoming the second largest financial institution to fall in US history.
*14th* – The US government announced emergency measures to expand credit access to mortgage finance companies Fannie Mae and Freddie Mac, and allow the Treasury to buy shares in the companies.
*30th* – UK bank Lloyds TSB revealed £585m. of writedowns as pre-tax profits fell by 70% in the first half of the year.
*31st* – Nationwide recorded an 8.1% fall in the value of houses, the biggest annual fall in UK house prices since their surveys began in 1991.
*31st* – Halifax Bank of Scotland (HBOS) announced that its first-half profits fell by 72% to £848m. while bad debts rose by 36% to £1.31bn.

**August 2008**

*1st* – UK mortgage lender Alliance & Leicester revealed a £209m. hit on risky assets and higher funding costs as pre-tax profits for the first half of the year fell by 99% on the previous year.
*1st* – US mortgage lender IndyMac Bank filed for Chapter 7 bankruptcy protection.
*4th* – HSBC announced a 28% decline in half-year profits to £5.1bn.
*5th* – French bank Société Générale reported a 63% fall in second-quarter profits, after its investment banking division lost €1.2bn. (US$1.9bn.) from sub-prime related investments.
*6th* – US mortgage lender Freddie Mac announced a second quarter loss of US$822m., its fourth successive loss, with credit-related expenses doubling to US$2.8bn. and US$1bn. lost on company writedowns on the value of sub-prime mortgages.
*7th* – Barclays revealed a 33% decline in first-half year profits together with further writedowns of £2.4bn. from bad loans and other credit impairment charges.
*8th* – RBS announced the second largest loss in UK banking history, with a pre-tax loss of £692m. for the first half of the year, resulting from £5.9bn. of writedowns.
*29th* – UK mortgage lender Bradford & Bingley reported a loss of £26.7m for the first six months of the year.
*30th* – Chancellor Alistair Darling warned that the UK economy faced its worst economic crisis in 60 years and claimed that the downturn would be more 'profound and long-lasting' than most people had imagined.

**September 2008**

*5th* – Fears over a global economic slowdown, combined with news that the US economy had shed 84,000 jobs the previous month, led to losses in global stock markets. London's FTSE 100 experienced its biggest weekly decline since July 2002, while markets in Paris, Frankfurt, Japan, Hong Kong, China, Australia and India all fell between 2 and 3%.

*7th* – US mortgage lenders Fannie Mae and Freddie Mac, who together accounted for nearly half of all outstanding mortgages in the USA, were taken into public ownership in one of the largest bail-outs in US history.

*7th* – In the UK, Nationwide Building Society took ownership of smaller rivals Derbyshire and Cheshire Building Societies.

*10th* – The European Commission predicted that the UK, Spain and Germany would fall into recession and eurozone growth would fall to 1.3% in 2008, 0.4% less than previous projections.

*15th* – US investment bank Lehman Brothers filed for Chapter 11 bankruptcy protection after it was unable to find a buyer. It became the first major bank to collapse since the beginning of the credit crisis.

*15th* – The Bank of America bought out US bank Merrill Lynch for US$50bn.

*15th* – Fears over the strength of the global financial system following the collapse of Lehman Brothers caused stock markets across the globe to tumble. The FTSE 100 Index fell by 212.5 points, wiping £50bn. off the top 100 British companies, while the Dow Jones Industrial Average shed 504 points, its biggest fall since the 9/11 attacks.

*16th* – The US Federal Reserve launched an US$85bn. rescue package for AIG, America's largest insurance company, to protect it from bankruptcy in return for an 80% public stake in the business.

*17th* – Lloyds TSB agreed to take over HBOS, Britain's largest mortgage lender, in a deal worth £12bn. following a run on HBOS shares.

*17th* – UK bank Barclays bought Lehman Brothers' North American investment banking and trading unit for US$250m., along with the company's New York HQ and two data centers for a further US$1.5bn.

*18th* – The US Federal Reserve, together with the European Central Bank, the Bank of England, the Bank of Japan, the Bank of Canada and the Swiss National Bank, pumped US$180bn. of extra liquidity into global money markets.

*22nd* – Japan's largest brokerage house Nomura Holdings Ltd acquired the Asian operations of Lehman Brothers, worth around US$230m.

*22nd* – Wall Street banks Morgan Stanley and Goldman Sachs give up their status as investment banks to become lower risk, tightly regulated commercial banks.

*23rd* – Nomura Holdings acquired the European and Middle Eastern equities and investment banking operations of Lehman Brothers.

*25th* – US mortgage lender Washington Mutual collapsed. Its assets were sold to JP Morgan Chase for US$1.9bn.

*25th* – Ireland became the first eurozone economy to fall into recession.

*29th* – European bank Fortis was partially nationalized following talks between the European Central Bank and the Netherlands, Belgium and Luxembourg. Each country agreed to put €11.2bn. (US$16.1bn.) into the bank.

*29th* – UK mortgage lender Bradford & Bingley was taken into public ownership, with the government taking control of the company's £50bn. mortgages and loans, while its savings unit and branches were to be sold to Spain's Santander.
*29th* – US bank Wachovia agreed to a rescue takeover by Citigroup, absorbing US$42bn. of the company's losses.
*29th* – The Icelandic government took a 75% stake in Glitner, Iceland's third largest bank, for €600m. (US$860m.).
*29th* – The German government injected €35bn. (US$50.2bn.) into Hypo Real Estate, the country's second largest commercial property lender.
*29th* – A US$700bn. rescue package was rejected by the US House of Representatives. Wall Street stocks plummeted, with the Dow Jones Index shedding 778 points, its biggest ever one-day fall. The FTSE 100 lost 269 points in one of its worst-ever trading days.
*30th* – European bank Dexia was bailed out, with the Belgian, French and Luxembourg governments injecting €6.4bn. (US$9bn.).
*30th* – The Irish government stepped in with €400bn. (US$562.5bn.) to guarantee all deposits, debts and bonds in six banks until September 2010.
*30th* – Japan's Nikkei 225 stock fell by 4.1% to register its lowest closing point since June 2005, while in Hong Kong the Hang Seng index ended the day down 2.4%.

**October 2008**

*3rd* – The US House of Representatives passed a US$700bn. rescue package. The plan aimed to buy up bad debts of failing banks while guaranteeing deposit accounts up to US$250,000.
*3rd* – US bank Wells Fargo announced a buy-out of Wachovia for US$15.1bn.
*3rd* – The UK government increased guarantees for bank deposits to £50,000, effective from 7 October 2008.
*6th* – Germany's finance ministry, together with private banks, agreed a €50bn. (US$68bn.) deal to save Hypo Real Estate.
*6th* – French bank BNP Paribas announced it had agreed to take control of Fortis' operations in Belgium and Luxembourg, together with its international banking franchises, for €14.5bn. (US$19.7bn.).
*6th* – The Iceland Stock Exchange temporarily suspended trading in six of the economy's largest financial firms. Banks agreed to sell off their foreign assets to help bolster the domestic banking sector.
*7th* – The Icelandic government took control of Landsbanki, the nation's second largest bank. Internet bank Icesave, owned by Landsbanki, suspended all deposits and withdrawals.
*8th* – The UK government announced a £400bn. (US$692bn.) package of reforms, including £50bn. to the top eight financial institutions, an extra £100bn. available in short-term loans from the Bank of England and £250bn. in loan guarantees to encourage banks to lend to each other.
*8th* – Six central banks—the US Federal Reserve, the Bank of England, the European Central Bank, the Bank of Canada, the Swiss National Bank and Sveriges Riksbank—coordinated an emergency interest rate cut of half a percentage point.

*8th* – The UK government announced that it planned to sue Iceland to recover deposits in Icesave, the failed Internet bank that had earlier stopped customers from withdrawing money.
*9th* – The IMF drew up emergency plans to make funds available to governments affected by the financial crisis.
*10th* – Japan's Nikkei stock average shed 881 points, or 9.62%, to fall to its lowest level since May 2003. Yamato Life Insurance became Japan's first major victim of the global financial crisis.
*10th* – Singapore officially fell into recession after the export-dependent economy experienced a fall in demand from US and European markets.
*10th* – The FTSE 100 closed down 8.85%, having lost 381.7 points, its worst fall since the crash of 1987, knocking £89.5bn. off the value of the UK's largest companies.
*11th* – The G7 nations agreed a five-point plan to unfreeze credit markets, including adoption of Britain's proposal to part-nationalize banks.
*13th* – The UK government announced an injection of £37bn. into RBS, Lloyds TSB and HBOS in return for a controlling share of each company.
*13th* – Germany and France led a coordinated plan to restore liquidity into their banking sectors in a move costing up to €2trn. for the EU's 27 states.
*13th* – The Dow Jones Industrial Average gained 936 points or 11%, its highest one-day gain and its largest percentage jump since 1933, following news of plans to increase bank liquidity.
*14th* – The US government revealed a US$250bn. plan to part-nationalize several banks.
*15th* – Retail sales in the US in Sept. recorded their biggest decline in over three years as the Dow Jones index fell by 7.87%, its largest decline since 26 Oct. 1987.
*15th* – JP Morgan Chase announced a quarterly profit fall of 84%, while Wells Fargo suffered a 25% drop in earnings.
*16th* – The Swiss government injected US$60bn. into UBS in return for a 9.3% stake and a boost in capital, while Credit Suisse turned down the offer of state aid but raised capital from private investors and a sovereign wealth fund.
*16th* – Citigroup posted its fourth consecutive quarterly loss with a shortfall of US$2.81bn. for the third quarter, following over US$13bn. of writedowns.
*17th* – French bank Caisse d'Epargne admitted a €600m. (US$807m.) derivatives trading loss triggered by 'extreme market volatility' during the week of 6 October.
*19th* – Dutch savings bank ING received a €10bn. (US$13.4bn.) capital injection from the Netherlands authorities in return for preference shares in the company. The Dutch government established a €20bn. fund to support domestic banks as required.
*19th* – South Korea announced a rescue package worth US$130bn. offering a state guarantee on banks' foreign debts and promising liquidity to firms.
*20th* – Sweden's government offered credit guarantees up to 1.5trn. kroner (US$205bn.), with 15bn. kroner set aside in a bank stabilization fund.
*22nd* – US bank Wachovia reported a US$24bn. loss for the third quarter, the biggest quarterly loss of any bank since the beginning of the credit crunch.

*24th* – Official data showed that the UK economy contracted for the first time in 16 years, with a fall in economic growth of 0.5% for the third quarter.
*24th* – The Danish central bank raised interest rates by a half-point to 5.5%.
*29th* – The US Federal Reserve slashed interest rates by a half-point to 1%, its lowest level since June 2004.
*29th* – The IMF, European Union and World Bank announced a rescue package for Hungary, pledging US$25.1bn. to promote confidence in the country's financial markets and its currency.
*30th* – Deutsche Bank reported a large fall in profits following writedowns of €1.3bn. in the third quarter.
*30th* – Japan unveiled a 27trn. yen (US$270.6bn.) stimulus package for small businesses and to provide emergency cash to families exposed to the credit crunch.
*31st* – The Bank of Japan cut interest rates, from 0.5% to 0.3%, for the first time in seven years in response to the global financial crisis.

**November 2008**

*4th* – HBOS revealed writedowns for the nine months up to Sept. at £5.2bn., up from £2.7bn. for the first half of the year.
*5th* – The Italian government offered up to €30bn. (US$39bn.) to recapitalize banks.
*5th* – Australia's central bank slashed interest rates by a higher-than-expected 75 basis points to 5.25%, the lowest level since March 2005.
*6th* – The IMF approved a US$16.4bn. loan to Ukraine.
*6th* – The Bank of England reduced interest rates by 1.5% to 3%, the lowest level since 1955.
*6th* – The European Central Bank lowered interest rates by a half-point to 3.25%.
*9th* – The Chinese government announced a US$586bn. stimulus package. The plan to relax credit conditions, cut taxes and invest in infrastructure and social projects over a two-year period equated to 7% of the country's GDP.
*11th* – US electronics retailer Circuit City filed for Chapter 11 bankruptcy protection. It became the largest US retailer to fall victim to the credit crisis.
*11th* – Swedish investment bank Carnegie was taken over by the Swedish government after its license was revoked for failures in internal controls.
*14th* – The eurozone officially slipped into recession after figures showed the area shrunk by 0.2% for the second consecutive quarter.
*20th* – The IMF approved a US$2.1bn. loan for Iceland in an attempt to 'restore confidence and stabilize the economy.'
*23rd* – The US government agreed a bailout of Citigroup, injecting US$20bn. of capital in return for preference shares. The move included a guarantee of up to US$306bn. of Citigroup's risky loans and securities.
*24th* – In his pre-Budget report, Chancellor Alistair Darling unveiled a fiscal stimulus plan. VAT was reduced to 15% from 17.5% and an extra £20bn. was to be pumped into the economy, with government borrowing set to increase to record levels.
*25th* – The IMF approved a US$7.6bn. loan to Pakistan.

*25th* – The US Federal Reserve pumped a further US$800bn. into the economy, with US$600bn. to buy up mortgage-backed securities and US$200bn. to unfreeze the consumer credit market.
*26th* – The European Commission unveiled a €200bn. (US$256bn.) economic recovery plan.

**December 2008**

*4th* – French President Nicolas Sarkozy announced a €26bn. (US$33bn.) stimulus plan, including a €1bn. loan to carmakers and €5bn. of new public sector investments. The French government would offer companies €11.5bn. worth of credits and tax breaks on investments for 2009.
*4th* – The Bank of England cut interest rates by 1% to 2% with business surveys suggesting that the downturn had gathered pace.
*4th* – The Reserve Bank of New Zealand reduced interest rates by a record 150 basis points to 5%.
*4th* – The European Central Bank reduced its main interest rate by 75 basis points to 2.5%, its largest ever cut.
*4th* – Sweden's central bank cut interest rates by a record 1.75% to 2%, while Denmark's central bank Nationalbank followed with a 75 basis point reduction to 4.25%.
*9th* – The Bank of Canada lowered its benchmark interest rate by 75 basis points to 1.5%, its lowest rate since 1958.
*11th* – The Bank of Korea reduced interest rates by a record 1% to 3%.
*16th* – The US Federal Reserve slashed interest rates from 1% to a range between zero and 0.25%, its lowest recorded level.
*19th* – Japan's central bank cut interest rates from 0.3% to 0.1%, having projected that the economy would shrink by 0.8% in the current fiscal year and experience zero growth for the year ending March 2010.
*19th* – The US government pledged US$17.4bn. of its US$700bn. originally allocated for the financial sector to help ailing carmakers General Motors, Chrysler and Ford.
*22nd* – China cut interest rates by 27 basis points to 5.31%, its fifth reduction in four months.
*30th* – The US Treasury unveiled a US$6bn. rescue package for GMAC, the car-loan arm of General Motors, aimed at encouraging GMAC to offer funding to potential vehicle buyers.

**January 2009**

*8th* – The Bank of England reduced interest rates by a half-point to 1.5%, the lowest level since the bank was founded in 1694.
*8th* – Commerzbank received €10bn. (US$13.7bn.) of capital from the German government in return for a 25% stake following liquidity problems arising from its decision to purchase Dresdner Bank from insurance company Allianz.
*8th* – South Korea's central bank cut interest rates from 3% to a record low of 2.5%.
*9th* – Official figures showed that more jobs were lost in the USA in 2008 than in any year since the Second World War, with 2.6m. axed. The jobless rate increased to 7.2% in Dec. 2008, its highest level in 16 years.

*13th* – China's exports fell by 2.8% in Dec. compared to the previous year, the largest decline in ten years.
*13th* – German chancellor Angela Merkel unveiled an economic stimulus package worth €50bn. (US$67bn.), including public investments and tax relief.
*14th* – The UK government guaranteed up to £20bn. of loans to small and medium-sized businesses.
*14th* – Shares in Europe and the USA fell sharply following the release of official figures showing a 2.7% fall in US retail sales in Dec. London's FTSE 100 closed down by over 5%, the main markets in France and Germany lost nearly 4.5% and the US Dow Jones index fell by 3%.
*15th* – The European Central Bank slashed interest rates by a half-point to 2%, its lowest level since Dec. 2005.
*16th* – The Irish government moved to nationalize Anglo Irish Bank.
*16th* – Reporting a fourth quarter loss of US$8.29bn., Citigroup announced plans to split into two new firms, Citicorp and Citi Holdings.
*16th* – Bank of America received US$20bn. of fresh US government aid and US$118bn. worth of guarantees following losses incurred in its takeover of Merrill Lynch. Merrill Lynch posted a fourth-quarter loss of US$15.3bn. while Bank of America lost US$1.7bn. in the same period.
*19th* – Spain became the first triple-A rated nation to have its credit rating downgraded since Japan in 2001.
*19th* – Denmark offered up to 100bn. kroner (US$17.6bn.) in loans to help recapitalize its banks.
*20th* – The French government offered its ailing car industry up to €6bn. (US$7.7bn.) in aid.
*23rd* – The UK economy officially entered recession after figures showed a fourth-quarter fall in GDP of 1.5% following a 0.6% drop the previous quarter.
*25th* – The French government provided €5bn. (US$6.5bn.) in credit guarantees to help Airbus.
*26th* – Dutch banking and insurance group ING estimated fourth-quarter losses of €3.3bn. (US$4.3bn.), prompting it to seek state guarantees, replace its chief executive and shed 7,000 jobs.
*28th* – The IMF warned that world economic growth would fall to 0.5% in 2009, its lowest level since the Second World War, and projected the UK economy would shrink by 2.8%, the worst contraction among developed nations.
*28th* – The International Labour Organization claimed 51m. jobs could be lost in 2009, pushing the world unemployment rate to 7.1% compared with 6.0% at the end of 2008.
*28th* – Canada's Conservative government unveiled a $40bn. CDN (US$32bn.) stimulus plan including tax cuts and infrastructure spending.
*29th* – New Zealand's central bank reduced interest rates by 1.5% to 3.5%.

**February 2009**

*3rd* – The Australian government announced a second stimulus package of $A42bn. (US$26.5bn.) to boost long-term growth, including one-off cash payments

to low-income families and investment in infrastructure. The Reserve Bank of Australia reduced interest rates by one percentage point to 3.25%, its lowest level in 45 years.

*5th* – The Bank of England slashed interest rates by a half-point to a record low of 1%.

*5th* – Deutsche Bank unveiled a fourth-quarter loss of €4.8bn. (US$6.1bn.) and a net loss for 2008 of €3.9bn. (US$5bn.)—its first yearly loss since being restructured after the Second World War—citing 'unprecedented' operating conditions and 'weaknesses in our business model.'

*9th* – Barclays announced a pre-tax profit of £6.1bn. (US$9bn.) for 2008, down 14% on profits for the previous year.

*9th* – The French government agreed to provide Renault and Peugeot-Citroën with €3bn. (US$3.9bn.) each in preferential loans in return for maintaining jobs and sites in France. Renault Trucks, owned by Volvo, was offered a loan of €500m. (US$650m.), suppliers €600m. (US$780m.) and the financing arms of the two carmakers loan guarantees of up to €2bn. (US$2.6bn.).

*10th* – Former bosses of RBS and HBOS, two of the UK's largest financial casualties, apologized 'profoundly and unreservedly' for their banks' failure during the UK Treasury Committee's inquiry into the banking crisis.

*10th* – UBS declared a Swiss corporate history record loss of SFr19.7bn. (US$17bn.) for 2008 after suffering a net loss of SFr8.1bn. (US$7bn.) in the fourth quarter, including SFr3.7bn. (US$3.2bn.) in exposure to toxic assets. The bank announced it would axe a further 2,000 jobs at its investment banking arm.

*12th* – The Bank of Korea reduced interest rates by 50 basis points to a record low 2%.

*12th* – The Irish government revised its rescue plans for Allied Irish Bank and the Bank of Ireland. Each bank was to receive €3.5bn. (US$4.5bn.) and would be expected to increase lending and reduce senior executives' pay while remaining in the private sector.

*12th* – The Spanish economy fell into recession for the first time in 15 years, having shrunk by 1% in the fourth quarter of 2008.

*17th* – US President Barack Obama signed his US$787bn. economic stimulus plan after Congress approved the package.

*18th* – Taiwan fell into recession after its economy slumped by 8.4% in the fourth quarter. Taiwan's central bank reduced interest rates by a quarter point to 1.25%.

*19th* – The Bank of Japan bought 1trn. yen (US$10.7bn.) in corporate bonds and maintained a near-zero interest rate.

*26th* – RBS unveiled a loss of £24.1bn. (US$34.2bn.), the largest annual loss in UK corporate history, stemming from a £16.2bn. (US$23bn.) writedown of assets mainly linked to its purchase of ABN Amro. The bank also announced it would put £325bn. of toxic assets into a new government insurance scheme, while the government would inject a further £13bn. to strengthen its balance sheet.

*27th* – The European Bank for Reconstruction and Development (EBRD), the European Investment Bank (EIB) and the World Bank announced a €24.5bn. (US$31bn.) joint rescue package for banking sectors in Central and Eastern Europe.

The two-year initiative would include equity and debt financing and policies to encourage lending, particularly to small and medium-sized firms.

**March 2009**

*2nd* – US insurance company AIG unveiled a US$61.7bn. loss in the fourth quarter of 2008, the largest in US corporate history, and received an additional US$30bn. as part of a revamped rescue package from the US government.

*2nd* – HSBC, Europe's largest bank, confirmed it was looking to raise £12.5bn. (US$17.7bn.) from shareholders through a rights issue after it revealed pre-tax profits for 2008 of US$9.3bn., down 62% on the previous year.

*3rd* – Nationalized UK bank Northern Rock confirmed it made a loss of £1.4bn. (US$2.0bn.) in 2008.

*3rd* – Toyota Motors, the world's largest carmaker by sales, asked for up to US$2bn. in Japanese government-backed aid.

*4th* – The Australian economy shrank by 0.5% in the fourth quarter of 2008.

*4th* – The World Bank signed a US$2bn. contingency facility to Indonesia, the largest ever loan granted to an economy not classified as in crisis. Indonesia's central bank reduced its interest rate by 50 basis points to 7.75%.

*5th* – The Bank of England cut interest rates from 1% to 0.5%. The Bank also announced it was to create £75bn. of new money, called quantitative easing.

*9th* – Iceland nationalized Straumur-Burðarás, the last of the big four banks to be taken into public ownership.

*10th* – Malaysia revealed a 60bn. ringgit (US$16.3bn.) stimulus package over a two year-period, amounting to 9% of GDP. The plan contained increased spending on infrastructure, guaranteed funds for businesses, equity investments to boost the stock market and tax breaks.

*14th* – The G20 group of rich and emerging nations pledged a 'sustained effort' to restore global growth with low interest rates and increase funds to the IMF.

*16th* – Serbia opened talks with the IMF over an emergency loan worth up to €2bn. (US$2.6bn.).

*18th* – The Bank of Japan provided up to 1,000bn. yen (US$10bn.) in subordinated loans to its commercial banks.

*18th* – The US Federal Reserve pledged US$1.2trn. to buy long-term government debt and mortgage-related debt.

*18th* – UniCredito, one of Italy's largest banks, sought €4bn. in aid from Italian and Austrian sources.

*19th* – The US Treasury promised up to US$5bn. to auto parts suppliers, guaranteeing payment for products shipped.

*20th* – The IMF revised its global forecast for 2009, with the world economy set to shrink by between 0.5% and 1%. The world's most developed economies were expected to experience the largest contractions in GDP.

*23rd* – The US announced a 'Public-Private Investment Programme' to buy up to US$1trn. worth of toxic assets. The US Treasury committed between US$75bn. and US$100bn. to the program, in addition to contributions from the private sector.

*25th* – The IMF, along with the World Bank, European Commission and other multilateral organizations, unveiled a €20bn. (US$27.1bn.) financial rescue package for Romania. The agreement stipulated Romania reduce its budget deficit to less than 3% of GDP by 2011.

*25th* – Italian bank Banca Popolare di Milano became the fourth bank in the country to seek funding from the government's €12bn. bank aid scheme. The bank requested €500m.

*26th* – Official statistics revealed that Ireland's economy shrank by 7.5% in the fourth quarter of 2008 compared to the same period the previous year, its largest contraction in decades. For the whole of 2008, the economy contracted by 2.3%, its first fall since 1983.

*26th* – The US economy contracted at an annualized rate of 6.3% in the fourth quarter of 2008, its fastest rate since 1982.

*27th* – The UK economy shrank by 1.6% in the last three months of 2008, its largest fall in GDP since 1980 and higher than the earlier 1.5% estimate.

*29th* – The German government pumped €60m. (US$80m.) into Hypo Real Estate in return for an 8.7% stake.

*30th* – The Spanish government, with the Bank of Spain, launched a €9bn. (US$12bn.) bailout of savings bank Caja Castilla La Mancha, the country's first bank rescue in the financial crisis.

*31st* – The World Bank predicted the global economy would contract by 1.7% in 2009, the first decline since the Second World War. The forecast claimed that the most developed economies would shrink by 3%, while world trade would fall by 6.8%.

**April 2009**

*2nd* – The G20 agreed to tackle the global financial crisis with fresh measures worth up to US$1.1trn. Pledges included US$750bn. made available to the IMF to help troubled economies and US$250bn. to boost global trade.

*6th* – Japan unveiled its latest stimulus package worth 10trn. yen (US$98.5bn.), equivalent to 2% of GDP.

*7th* – The Reserve Bank of Australia reduced its benchmark rate by a quarter point to 3%, its lowest level since 1960.

*7th* – RBS announced it would shed a further 9,000 jobs from its global operations over the next two years.

*14th* – Goldman Sachs reported a higher than expected pre-tax quarterly profit of US$1.8bn. The bank would also place US$5bn. worth of shares on the stock market in order to repay an emergency US$10bn. loan provided by the US government in 2008.

*14th* – Poland's government approached the IMF to secure a US$20.5bn. credit line to increase bank reserves and make Poland 'immune to the virus of the crisis and speculative attacks.'

*14th* – Fortis bank posted a loss of €20.6bn. (US$27.5bn.) for 2008 following write-downs on debt and a separation of the business.

*15th* – UBS unveiled a first quarter loss of SFr2bn. (US$1.75bn.) and announced it would cut 8,700 jobs by 2010 in an effort to reduce costs.

*16th* – China's growth rate slowed to 6.1% in the first quarter of 2009, its slowest pace since quarterly GDP data was first published in 1992. Growth was down from 6.8% in the previous quarter and 9% for the whole of 2008.
*16th* – Consumer prices in the USA fell by 0.4% over the year to March owing to weak energy and food prices, the first year-on-year drop since Aug. 1955.
*16th* – JP Morgan Chase reported a higher than expected first quarter profit of US$2.1bn. compared with net income of US$2.4bn. in the first quarter of 2008.
*18th* – The IMF formally agreed a US$47bn. credit line for Mexico under its new fast track scheme to help developing nations cope with the global financial crisis.
*21st* – UK annual inflation as measured by the Retail Prices Index (RPI) was –0.4% in March (down from zero in Feb.), the first negative figure since 1960.
*21st* – Sweden's central bank reduced its key interest rate by a half point to a record low of 0.5%.
*22nd* – UK chancellor Alistair Darling admitted the economy faced its worst year since the Second World War as he unveiled his latest Budget report. The annual budget deficit would rise sharply to £175bn. over the next two years with total government debt to reach 79% of GDP by 2013.
*22nd* – The IMF said global output would contract by 1.3% in 2009, a 'substantial downward revision' of its Jan. forecasts when it predicted growth of 0.5%. The UK economy was now projected to shrink by 4.1% in 2009, while Germany was set to decline by 5.6% and Japan by 6.2%.
*22nd* – India's central bank slashed interest rates for the sixth time in six months, reducing its key repo lending rate by a quarter point to 4.75%.
*27th* – National Australia Bank, Australia's largest lender, announced a 9.4% fall in cash earnings to A$2bn. (US$1.4bn.) for the Sept.–March period.
*28th* – Fears over a swine flu outbreak continued to have an impact on global shares—the FTSE100 closed down by 1.7%, markets in Paris and Frankfurt ended nearly 2% down, Japan's Nikkei index fell by 1.7% and Hong Kong's Hang Seng shed 1.4%.
*28th* – Lithuania's economy contracted by 12.6% in the first quarter of 2009 compared to the same period in 2008, the largest year-on-year fall in the EU since the start of the recession.
*29th* – US output contracted at an annualized rate of 6.1% in the first quarter of the year, a higher-than-expected result. The contraction was led by a 30% decline in exports, its largest fall in 40 years.

**May 2009**

*1st* – US carmaker Chrysler filed for Chapter 11 bankruptcy protection after a group of hedge and investment funds refused to restructure the company's US$6.9bn. debt.
*1st* – The Reserve Bank of New Zealand reduced interest rates by 50 basis points to a record low of 2.5%. The bank governor, Alan Bollard, said he expected rates to remain at the current (or lower) level until the latter part of 2010.
*4th* – The European Commission forecast that the EU economy would contract by 4% in 2009, more than twice the level predicted at the beginning of the year. It claimed unemployment would now reach 10.9% in 2010.

*5th* – Japan offered US$100bn. of financial assistance to Asian economies affected by the global economic slowdown in a meeting of the finance ministers of the ten-member Association of South East Asian Nations.
*5th* – UBS confirmed it had made a SFr2bn. (US$1.75bn.) loss in the first quarter of 2009.
*6th* –Volkswagen and Porsche agreed to merge, relieving the sports carmaker of its debt burden.
*7th* – Barclays announced a pre-tax profit of £1.37bn. (US$2.07bn.) for the first three months of the year, up 15% from the previous year.
*7th* – Commerzbank agreed to relinquish the core of its commercial property lending business together with Eurohypo's role in public sector finance, in a deal with European competition authorities to compensate for €18.2bn. (US$24.2bn.) of state aid it received.
*7th* – The European Central Bank cut its main interest rate by a quarter point to a record low of 1% and also announced plans to purchase €60bn. (US$80.4bn.) of covered bonds, which are backed by mortgage or public sector loans.
*7th* – The Bank of England announced it would pump a further £50bn. (US$75bn.) into the UK economy in a substantial expansion of its program of government bond purchases.
*8th* – RBS reported a pre-tax loss of £44m. for the first quarter of 2009, compared with a profit of £479m. for the same period the previous year.
*8th* – Several US banks unveiled plans to raise cash a day after the US Treasury said that ten of America's 19 largest banks failed their stress tests and needed to raise a combined total of $74.6bn. Wells Fargo and Morgan Stanley planned to raise US$7.5bn. and US$3.5bn. respectively through share sales, while Bank of America planned to sell assets and raise capital to secure US$33.9bn. it needed.
*13th* – Franco-Belgian bank Dexia, which had been bailed out by three economies the previous year, posted a first quarter profit of €251m. (US$341m.) compared to a loss of €3.3bn. (US$4.5bn.) in 2008.
*13th* – The German cabinet agreed a 'bad bank' scheme, in which banks would be able to swap their toxic debt for government-backed bonds in return for paying an annual fee.
*14th* – Spain suffered a fall in GDP of 1.8% in the first quarter of 2009, its largest contraction in 50 years, according to the National Statistics Institute.
*14th* – Crédit Agricole unveiled a net profit of €202m. (US$275m.) in the first quarter, a 77% fall from the same period the previous year, after more than doubling its loan-loss provisions to €1.1bn.
*15th* – According to Eurostat economies that make up the eurozone contracted by 2.5% in the first quarter of 2009, a higher-than-forecast decline.
*15th* – The EBRD revealed plans to invest a record €7bn. (US$9.4bn.) in 2009 to tackle the slowdown through investments in infrastructure, energy, corporate and finance projects.
*17th* – Carmaker General Motors announced plans to close up to 1,100 dealerships in the USA as it battled to reduce costs and stave off bankruptcy.

*19th* – Inflation in the UK as measured by the Consumer Prices Index (CPI) slowed to 2.3% in April from 2.9% the previous month.
*20th* – Japan's GDP slid by 4% in the first quarter, its largest decline since records began in 1955.
*20th* – Venezuela experienced its slowest rate of growth in five years, with GDP growing by 0.3% in the first quarter of 2009 as the fall in oil prices took effect.
*21st* – The Office for National Statistics said public sector net borrowing in the UK rose to £8.46bn. in April compared to £1.84bn. in the same month the previous year. Concerned about its significant debt burden, Standard & Poor's downgraded the UK's credit rating from 'stable' to 'negative' for the first time since it began analyzing its public finances in 1978.
*22nd* – Private equity firms paid US$900m. to rescue BankUnited, a Florida-based bank worth around US$13bn. It had been closed by federal regulators in what was the biggest US bank failure of 2009 so far.
*22nd* – The US Treasury provided automotive financing group GMAC with a further US$7.5bn. in state aid to help it stay in business and offer loans to potential Chrysler and GM car buyers.
*22nd* – UK output declined by an unrevised 1.9% in the first quarter of 2009, according to figures published by the Office for National Statistics.
*26th* – South Africa fell into recession for the first time since 1992 following an annualized contraction of 1.8% and 6.4% in the previous two quarters.
*27th* – Riksbank announced it was raising foreign currency to boost its US$22bn. currency reserves, causing a sharp fall in the Swedish krona as the central bank warned the worst of the financial crisis may not be over.
*29th* – India's economy grew by 5.8% in the first quarter of 2009, higher than forecast but down from 8.6% in the same quarter the previous year.

**June 2009**

*1st* – US car manufacturer General Motors filed for Chapter 11 bankruptcy protection, the biggest failure of an industrial company in US history.
*2nd* – Switzerland officially entered recession after the economy contracted by 0.8% in the first three months of 2009, following a decline of 0.3% in the final quarter of 2008.
*3rd* – Australia recorded a 0.4% rise in GDP for the first quarter compared to the same period last year, bucking international trends.
*3rd* – Lloyds Banking Group announced plans to cut 530 jobs and close one site in the UK by the end of 2009.
*4th* – Industrial and Commercial Bank of China (ICBC), the world's second largest bank by market value, unveiled plans to buy 70% of Bank of East Asia's Canadian unit as part of a move to expand overseas.
*4th* – The Bank of England kept interest rates unchanged at 0.5% for the third month in a row.
*8th* –The OECD claimed the pace of decline among its 30 member countries was slowing—the composite leading indicators index (CLI) rose 0.5 point in April.
*9th* – Lloyds Banking Group announced it was to shut all 164 Cheltenham & Gloucester branches, putting 1,660 jobs at risk.

*9th* – UK unemployment rose by 244,000 to 2.22m. in the first three months of the year according to the Office for National Statistics (ONS), the largest quarterly rise in the jobless rate since 1981.
*9th* – Official figures showed that exports in Germany were 4.8% lower in April than in March and 28.7% down on the previous year, the biggest annual fall since records began in 1950.
*10th* – The European Central Bank provided an emergency €3bn. to the central bank in Sweden, whose banks dominate the Baltic region's financial sector.
*10th* – BP's annual statistical review indicated that global oil consumption fell by 0.6% in 2008, the first fall since 1993 and the largest drop since 1982.
*10th* – Ten of the largest US banks gained permission from the US Treasury to repay US$68bn. in government bail-out money received through the Troubled Asset Relief Programme (TARP).
*11th* – Figures revealed that Chinese exports fell by a record 26.4% in May from the same month the previous year.
*11th* – Revised GDP growth figures showed Japan contracted by 3.8% in the first quarter of 2009, less than the original estimate of 4%.
*15th* – The Confederation of British Industry (CBI) predicted the UK economy would contract by 3.9% in 2009 before seeing a return to growth of 0.7% in 2010.
*15th* – The IMF revised its growth forecast for 2010 for the USA, claiming that the economy would now grow by 0.75% compared to its forecast of 0% earlier in the year.
*16th* – The Bank of Japan said that the economy was no longer deteriorating, a more positive assessment than the previous month when it had stated that the economy was continuing to worsen. Nonetheless, it maintained interest rates at 0.1%.
*16th* – China introduced an explicit 'Buy Chinese' policy as part of its economic stimulus program, leading to fears of an increase in protectionism across the world.
*17th* – The US government announced a major reform of banking regulation to curb excessive risk-taking among big banks and to prevent future financial crises. President Obama described the reforms as 'the biggest shake-up of the US system of financial regulation since the 1930s.'
*17th* – The OECD revised its growth forecast for Italy, predicted the economy would grow by 0.4% in 2010 compared to a previously estimated contraction of 0.4%. However, it downgraded its forecast for 2009 from a 4.3% decline to 5.3%.
*17th* – The World Bank raised its GDP growth forecast for China to 7.2% in 2009 from a previously estimated 6.5%, citing the impact of a fiscal stimulus package.
*18th* – Official figures showed inflation in India had turned negative for the first time since 1977. Wholesale prices fell 1.61% in the year to 6 June.
*22nd* – The Japanese government looked set to provide up to 100bn. yen (US$1bn.) in state aid to Japan Airlines, the country's biggest airline, on condition that the organization's management improves.
*24th* – The OECD said the world economy was near the bottom of the worst recession in post-war history and predicted that the 30 most industrialized countries would shrink by 4.1% in 2009. UK output was predicted to contract by 4.3% in 2009 and experience zero-growth in 2010.

*24th* – The European Central Bank pumped €442.2bn. (US$628bn.) in one-year loans into the eurozone's weakened banking system in an effort to unlock credit markets and revive the region's economies.
*24th* – Orders for new durable goods in the USA rose unexpectedly by 1.8% in May from the previous month, going against expectations of a drop of 0.9%.
*25th* – The IMF said that Ireland's economy would contract by 8.5% in 2009 and warned it would experience the worst recession in the developed world and struggle to bail out its banks.
*26th* – New Zealand suffered a fifth straight quarterly contraction after official figures showed the economy shrank by 2.7% in the first quarter of 2009.
*26th* – Consumer prices in Japan fell by 1.1% in May compared to the same month the previous year, its biggest fall since records began in 1970, fuelling fears of a new bout of deflation.
*26th* – Spain unveiled a €9bn. (US$12.7bn.) fund aimed at saving banks suffering during the downturn.
*30th* – Eurozone inflation turned negative for the first time since records began in 1991, with consumer prices 0.1% lower in June than twelve months earlier.
*30th* – Malaysia launched economic liberalization measures aimed at attracting foreign investments, including changes to its long-standing policy of giving preferential treatment to the country's ethnic Malay majority.

**July 2009**

*1st* – Japan's Shinsei Bank and Aozora Bank merged to create the country's sixth largest bank with assets of 18trn. yen (US$186bn.).
*1st* – Unemployment in Ireland reached 11.9% in June, its highest level since 1996.
*1st* – India's exports were down 29.2% in May from the same month the previous year, the economy's eighth consecutive fall in exports.
*7th* – Inflation in the Philippines fell to 1.5% in June, its lowest level in 22 years.
*10th* – US carmaker General Motors (GM), 61% owned by the US government, emerged from its bankruptcy protection after creating a 'new GM' made up of four key brands, including Cadillac.
*13th* – The US deficit moved above US$1trn. for the first time in history.
*14th* – Inflation in the UK fell below the Bank of England's target rate of 2% for the first time since 2007. Lower food prices caused the Consumer Prices Index to drop to an annual rate of 1.8% in June, down from 2.2% in May.
*14th* – Singapore grew at an annualized rate of 20.4% in the second quarter, its first quarterly expansion in a year following a revised contraction of 12.7% from January to March.
*14th* – Goldman Sachs reported a net profit of US$3.44bn. for the second quarter of the year, higher than analysts had forecast.
*15th* – UK unemployment increased by a record 281,000 to 2.38m. in the three months to May, its highest level in over ten years.
*15th* – Japan's central bank downgraded its economic forecast to a contraction of 3.4% from 3.1% for the 12 months to end-March 2010, but reiterated that the worst of the recession was over.

*15th* – Russia's economy contracted by 10.1% in the first half of 2009, its sharpest decline since the early 1990s.
*16th* – China's economy grew at an annualized rate of 7.9% in the second quarter, up from 6.1% between January and March, as the government upgraded the growth forecast to 8% for 2009 as a whole.
*16th* – JP Morgan Chase unveiled a second quarter profit of US$2.72bn., an increase of 36% on the same period the previous year.
*17th* – Ghana secured a US$600m. three-year loan from the IMF and was given access to a further US$450m. from the IMF through the special facility set up by the G20 summit to assist poor countries.
*20th* – Iceland announced a 270bn. kr. (US$2.1bn.) recapitalization plan for its banking system, issuing bonds to three new banks set up in 2008 following the collapse of the country's three main banks.
*21st* – UK government debt increased to £799bn., or 56.6% of UK GDP, its highest level since records began in 1974.
*22nd* – The National Institute of Economic and Social Research (NIESR) predicted UK GDP to fall by 4.3% in 2009 and UK GDP per capita to remain below its pre-recession levels until March 2014.
*22nd* – Morgan Stanley reported a loss of US$159bn. in the second quarter of 2009, compared to a US$698m. profit for the same period the previous year.
*23rd* – Credit Suisse unveiled a 29% increase in second quarter net profits of 1.57bn. Swiss francs (US$1.48bn.).
*23rd* – The Asian Development Bank said growth in East Asia, excluding Japan, would double to 6% in 2010, compared to a 3% expansion in 2009.
*23rd* – The rate of decline of Japan's exports slowed in June, a sign that government stimulus spending around the world may be supporting demand. However, exports were still 35.7% lower than the same month the previous year.
*24th* – The IMF approved a 20-month Stand-By Arrangement for Sri Lanka worth US$2.6bn. to support the country's economic reform package.
*24th* – The UK economy contracted by 0.8% in the second quarter of 2009, much lower than the 2.4% decline in the previous quarter but above analysts' 0.3% prediction.
*24th* – The South Korean economy grew by 2.3% from April to June, its fastest expansion in five-and-a-half years.
*28th* – Deutsche Bank unveiled a net profit of €1.09bn. (US$1.56bn.) for the second quarter of 2009, a 67% increase in profits compared to the same period the previous year.
*28th* – BBVA, Spain's second largest bank, reported a net profit of €1.56bn. (US$2.23bn.) for the second quarter thanks to higher income from loans.
*31st* – Mizuho Financial Group revealed a net loss of 4.4bn. yen (US$46m.) for the second quarter, its fourth consecutive quarterly loss.
*31st* – Japan's jobless rate increased by 830,000 in June to 3.48m., its highest level in six years.
*31st* – Eurozone unemployment reached 9.4% (or 14.9m. people) in June, its highest level in ten years.

**August 2009**

*3rd* – Barclays announced a pre-tax profit of £2.98bn. (US$5bn.) for the first six months of the year with an 8% increase in revenue.

*3rd* – HSBC saw pre-tax profits halve to £2.98bn. (US$5bn.) for the first half of 2009 compared to the same period the previous year, following the write-off of US$13.9bn. of bad debt in the USA, Europe and Asia.

*3rd* – World stock markets were boosted by brighter economic data—Standard & Poor's 500 index tipped beyond 1,000 for the first time since Nov. 2008, London's FTSE closed at its highest rate since Oct. 2008, the three major US indexes added over 1.25% by the end of trade after positive manufacturing survey results from July and European indexes also rose.

*4th* – UBS reported a loss of SFr1.4bn. (US$1.32bn.) in the second quarter, an improvement on the SFr2bn. loss made in the previous quarter.

*4th* – UniCredito, Italy's largest bank, unveiled better-than-expected second quarter earnings of €490m. (US$706m.), 9.2% higher than the previous quarter.

*5th* – Société Générale announced a second quarter profit of €309m. (US$445m.), 52% lower than the same period 12 months earlier.

*6th* – The Bank of England injected a further £50bn. into the UK economy as part of its quantitative easing program, bringing its total spending to £175bn.

*6th* – Commerzbank made a €763m. (US$1.1bn.) net loss in the second quarter, a small improvement on the €861m. loss registered in the previous quarter.

*7th* – RBS reported a pre-tax profit of £15m. for the first six months of the year.

*7th* – Italy's economy shrank by 0.5% in the second quarter, its fifth consecutive quarterly contraction but an improvement on the record 2.7% fall in Jan.–March.

*7th* – The IMF and Angola began talks on a loan to help the African country cope with the global economic slowdown.

*12th* – Dutch financial services group ING announced a €71m. (US$100m.) profit in the three months to the end of June, its first profit in three quarters.

*12th* – Commonwealth Bank of Australia, the country's second largest bank by market capitalization, posted net earnings of A$4.72bn. (US$3.89bn.), 1% lower than the previous year owing to higher bad debt charges and reduced wealth management unit income.

*12th* – The UK unemployment rate increased to 7.8% in the second quarter, its highest level since 1995.

*13th* – France and Germany both recorded second quarter growth figures of 0.3%, bringing a year-long recession to an end. However, the Eurozone contracted by 0.1%, its fifth consecutive quarterly fall in output.

*14th* – Colonial BancGroup, a property lender based in Montgomery, Alabama, became the largest bank in the USA to collapse in 2009.

*14th* – The Nigerian Central Bank injected N400bn. (US$2.6bn.) into five banks and sacked their managers, after the regulator claimed the banks were undercapitalized and posed a risk to the entire banking system.

*14th* – Hong Kong posted growth of 3.3% between April and June following four consecutive quarters of contraction. Singapore also announced its emergence from recession, with annualized growth of 20.7% in the second quarter of 2009.

*14th* – South Africa's central bank slashed its lending rate by a half-point to a four-year low of 7%, its sixth cut since Dec. 2008.
*17th* – Japan's economy grew by 0.9% in the second quarter of 2009, ending a run of four consecutive quarters of negative growth.
*18th* – The South African economy contracted for the third quarter in a row as output fell at an annualized rate of 3% between April and June.
*18th* – The CPI measure of inflation in the UK remained at the same level of 1.8% in July, although economists had forecast a decline to 1.5%.
*20th* – The UK's public sector net borrowing totalled £8bn. in July, the first July deficit for 13 years, as the government's overall debt reached its highest level since 1974 at 56.8% of GDP.
*20th* – Mexico's economy contracted by 10.3% in the second quarter owing to a decline in demand for exports and falling levels of tourism resulting from the outbreak of swine flu in April and May.
*24th* – Thailand posted growth of 2.3% in the second quarter of 2009 as it emerged out of recession.
*26th* – The Malaysian economy expanded by 4.8% in the second quarter of 2009 following two straight quarters of contraction.
*27th* – US GDP shrank at an annualized rate of 1% in the second quarter, lower than the 1.5% decline predicted by many economists.
*27th* – Credit Agricole, France's largest retail bank, announced a higher-than-expected second quarter profit of €201m. (US$286m.).
*28th* – The Office for National Statistics (ONS) revised the rate of contraction in the UK economy for the second quarter to 0.7% from the original estimate of 0.8%.
*28th* – Unemployment in Japan hit a record high of 5.7% in July and consumer prices fell by 2.2% compared to a year earlier, its fastest recorded pace.
*31st* – The Eurozone's annual rate of inflation fell by 0.2%, its third consecutive monthly decline.

**September 2009**

*1st* – India's exports fell at an annualized rate of 28% in July, its tenth consecutive monthly contraction.
*2nd* – The *de facto* government of Honduras received US$150m. from the IMF to boost its dollar reserves.
*2nd* – The OECD predicted that the recession in Iceland, marked by a large contraction in domestic demand, would be deeper than in most developed economies.
*3rd* – The OECD forecast the UK to be the only G7 economy to stay in recession at the end of 2009, while the eurozone and the USA would record two quarters of growth.
*4th* – The G20 group of nations agreed to continue fiscal stimulus until the recovery from recession was assured.
*5th* – The IMF sanctioned US$510m. to Zimbabwe, its first loan to the country in a decade, to replenish the economy's dwindling foreign currency reserves.
*8th* – The EBRD announced it would invest a record €8bn. (US$11.6bn.) in central and eastern Europe in the course of 2009.

*8th* – Estonia's GDP shrank at an annualized rate of 16.1% in the second quarter of 2009, its sixth consecutive quarterly contraction. Latvia contracted by 18.7% and Lithuania by 19.5% in the same period.
*8th* – The gold price climbed above $1,000 per ounce for the first time since Feb. on the back of a weakening dollar and lingering concerns over the sustainability of the world economy's recovery.
*9th* – The FTSE 100 broke through the 5,000-point barrier for the first time since Oct. 2008.
*11th* – Brazil emerged from recession after it grew by 1.9% between April and June following two successive quarters of contraction.
*14th* – The European Commission predicted that the eurozone would grow by 0.2% in the third quarter and 0.1% in the fourth quarter, but GDP for the year would fall overall by 4%.
*15th* – Consumer Price Index inflation in the UK measured 1.6% in Aug., its lowest level since Jan. 2005.
*15th* – US Federal Reserve chairman Ben Bernanke claimed recession in the US was 'very likely over' but the economy would remain weak for some time owing to unemployment.
*16th* – Unemployment in the UK rose by 210,000 in the three months to July to take the total to 2.47m., its highest level since 1995.
*17th* – The UK Office for National Statistics reported flat sales volumes in August compared with July, confounding analyst expectations of a 0.2% rise.
*18th* – The UK's public sector net borrowing totalled a record £16.1bn. in Aug., with government's overall debt standing at £804.8bn., or 57.5% of GDP.
*20th* – A further two US banks were closed by the country's federal regulator, taking the total number of US banks failing in 2009 to 94. Irwin Union Bank & Trust and Irwin Union Bank were shut down after their parent firm, Irwin Financial, failed to meet a Federal Deposit Insurance Corporation demand to boost their capital.
*21st* – The pound fell to its lowest level against the euro for five months as concerns continued about the underlying health of the British economy.
*22nd* – The Asian Development Bank made an upward revision of its growth forecast for India and China in 2009, with India expected to grow by 6.0% (up from an earlier forecast of 5.0%) and China by 8.2% (up from 7.0%).
*23rd* – The US dollar fell to a one-year low against the euro with traders switching to other currencies as signs of economic recovery emerged.
*23rd* – The World Bank announced it was to provide India with US$4.3bn. to fund infrastructure projects and support companies needing credit.
*24th* – Loss-making carrier Japan Airlines asked for a government bailout following recently announced plans to cut 6,800 jobs.
*26th* – Speaking at the end of the two-day G20 summit, US President Barack Obama said the world's leading nations had agreed to 'tough new measures' to prevent another global financial crisis, including regulation relating to the amount of money banks hold in reserve and a cap on pay for bankers.
*29th* – The Office for National Statistics revised growth figures for the UK in the second quarter from –0.7% to –0.6%.

*29th* – Core consumer prices in Japan fell 2.4% in Aug. year-on-year, the fourth successive month of contraction.

*30th* – The IMF slashed its forecast for the amount of bad debt likely to be written off globally between 2007 and 2010 from US$4.0trn. to US$3.4trn.

In Oct. 2009 US manufacturers reported that global output was growing at its fastest rate for five years. On 29 Oct. the Department of Commerce announced that the US economy was out of recession, growing by an annualized 3.5% in the third quarter. However, rising unemployment was an ongoing concern, standing at 10.2% in Oct. 2009 (its highest rate since 1983). US president Barack Obama responded to the news of the emergence from recession with caution, commenting: 'We anticipate that we are going to continue to see some job losses in the weeks and months to come.'

By the end of the third quarter of 2009, of the G7 economies only the UK remained in recession, having contracted by 0.4% in the period July–Sept.

This is an edited and updated version of the Credit Crunch Chronology that appears on The Statesman's Yearbook Online: http://www.statesmansyearbook.com/entry.html?entry=chronology_credit

THE STATESMAN'S YEARBOOK TEAM

*See Also* **Banking Crises; Great Depression.**

# credit rating agencies

## History

Bond rating and the establishment of formal CRAs began in 1909 when John Moody began rating US railroad bonds, soon expanding to utility and industrial bonds. Poor's Publishing Company followed in 1916 and Fitch Publishing Company in 1924. The business was characterised by the investor-pays model, where investors bought reports from the CRAs containing their ratings. This changed in 1970, for two reasons. First, with the advent of the photocopier free-riding became commonplace and CRAs found it difficult to sustain their business (White, 2002). Second, in 1970 Penn Central defaulted on its commercial paper obligations, creating vast mistrust among investors and a large demand by issuers for certification. The business thus changed to an issuers-pay model (Cantor and Packer, 1995). In 1975, the Securities and Exchange Commission (SEC) created the Nationally Recognized Statistical Rating Organization (NRSRO) category to designate credit ratings agencies whose ratings were recognised as being valuable for investment decisions. Standard & Poor's, Moody's and Fitch were given this designation immediately, and four other firms attained it in the following 17 years. By 2000, however, mergers returned the number of NRSROs to the big three. The SEC gave out a fourth NRSRO designation in 2003 (Dominion), a fifth in 2005 (A.M. Best), and in response to congressional legislation promoting transparency and entry in 2006 gave out three more designations (White, 2010). All of these new NRSROs, however, remain very small players in the bond and structured finance businesses.

## Important Aspects of Industry Structure

1. Many regulatory agencies use ratings in evaluation, e.g. to determine capital requirements. Moreover, certain entities such as banks, insurance and pension funds are restricted to invest only in *investment grade* securities, i.e. BBB and above (see Cantor and Packer, 1995). This creates an artificial demand for ratings. Kisgen and Strahan (2010) demonstrate that the acquisition of NRSRO status for Dominion Bond Rating Service in 2003 changed the impact of its ratings on bond yields only in situations where this status was important. Coval *et al.* (2009) provide evidence that Collateralized Debt Obligations (CDOs) were inaccurately priced because ratings were overly weighted by investors. Adelino (2009) finds that while initial yields on tranches below AAA for mortgage backed securities predict future credit performance the initial yields on AAA tranches had no predictive power. This is consistent with the hypothesis that investors in AAA tranches had no other information beyond the credit ratings themselves.
2. There are large *barriers to entry* in the credit rating industry: Since Congress, local governments, and regulatory agencies adopted the NRSRO designation and used it

for the determination of investment grade securities (point 1), this created an 'absolute barrier to entry' (White, 2002). Moreover, the need to build a reputation in order to receive business is a natural barrier to entry.

3. The fact that Moody's and S&P rate some corporate bonds which they are not paid for by issuers using public information (*unsolicited ratings*) is controversial. While the firms state that they are providing a service demanded by investors, some parties have raised the point that these ratings may be used to discipline issuers. Poon (2003) demonstrates that unsolicited ratings tend to be lower in general, but correcting for selection does not explain all of the variation.
4. CRAs have been able to avoid *liability* for problems with ratings. Under Section 11 of the Securities Act of 1933 they were immune from misstatements. Moreover, in court they have used the argument that ratings are speech and not recommendations on how to invest (Partnoy, 2002). The Dodd-Frank Financial Reform Bill passed recently exposes CRAs to liability by defining them as experts.
5. The market for corporate bond ratings is different from the market for *structured finance* ratings. Standard & Poor's and Moody's rate all corporate bonds, while the percentage that Fitch rates has been increasing. Most structured finance products receive at least two ratings, but who is rating it depends on the deal (see Ashcraft *et al.*, 2009). The corporate bond market is established and relatively simple, and the models used are well accepted. Structured finance products are fairly new but have grown rapidly; between 1997 and 2003 global structured finance issuance grew from about $280 billion to $800 billion (Committee on the Global Finance System, 2005). These products are very complex and the methods for rating structured products have been imprecise. Errors in the ratings agencies' data, assumptions and modelling have been found. Moreover, agencies are not required to perform due diligence on underlying loans and have difficulties retaining their best employees (Partnoy, 2002).
6. In the structured finance market, ratings *shopping* can occur. This means that if an issuer is unhappy with a rating, it may solicit another one, either from the same CRA or from another CRA. Moreover, 'typically the rating agency is paid only if the credit rating is issued' (US SEC, 2008).

## Evidence on CRAs in the corporate bond market

There has been a large focus on the effect of announcements on the pricing of both bonds and stocks. The main finding is the asymmetry between downgrades and upgrades: downgrades have a significant negative impact on price, but there is virtually no price change following an upgrade. The effect of ratings changes on price is complex, as the impact of ratings changes is different for firms with low ratings than for firms with high ratings. Overall, there is a clear consensus that information provided by CRAs has an effect on price (Hand *et al.*, 1992; Hite and Warga, 1997; Berger *et al.*, 2000; Kliger and Sarig, 2000; Dichev and Piotroski, 2001; Jorion and Zhang, 2007). These findings suggest a role for CRAs in the allocation of capital process.

In terms of accuracy, Cantor and Packer (1995) show that ratings order corresponds to default rankings. Hilscher and Wilson (2009) argue that rating agencies do a poor job at forecasting default probabilities, but capture systematic default risk.

Fitch is generally thought of as having higher ratings than Standard & Poor's and Moody's (Jewell and Livingston, 1999). Becker and Milbourn (2009) finding that increased competition from Fitch's increased market share in the corporate bond market led to more issuer-friendly ratings and also less informative ratings. Bongaerts *et al.* (2009) however, only find a certification role for Fitch in breaking ties between Moody's and Standard & Poor's.

## Structured finance products and the financial crisis of 2007–2009

Much attention has been paid to CRAs as a potential contributor to the financial crisis. The structured finance market collapsed and even 'the highest rated (AAA) mortgage-backed securities (as measured by the corresponding credit default swaps prices) fell by 70 percent between January 2007 and December 2008' (Pagano and Volpin, 2009), implying that ratings were not of high quality. There is debate over whether poor quality ratings were the fault of (i) conflicts of interest, (ii) imprecise modelling, or some mixture of both.

An SEC investigation found that senior analytical managers and supervisors participated in fee discussions with issuers and the analytical staff also discussed ratings decisions and methodology in the context of fees and market share (US SEC, 2008). In addition, CRAs offer related consulting services, such as pre-rating assessments (of what a rating might be).

A few recent theoretical papers study the implications of shopping for ratings. Bolton *et al.* (2010) demonstrate that competition among CRAs may reduce welfare due to shopping by issuers. Faure-Grimaud *et al.* (2009) look at corporate governance ratings in a market with truthful CRAs and rational investors. They show that issuers may prefer to suppress their ratings if they are too noisy. They also find that competition between rating agencies can result in less information disclosure. Skreta and Veldkamp (2009) also assume that CRAs truthfully relay their information and demonstrate how noisier information creates more opportunity for shopping by issuers to take advantage of a naive clientele.

In terms of conflicts of interest, Mathis *et al.* (2009) find that reputation cycles may exist where a CRA builds up its reputation by relaying information accurately only to take advantage of this reputation to later inflate ratings. Bolton *et al.* (2010) show that conflicts of interest for CRAs may be higher when reputation costs are lower and there are more naïve investors. Bar-Isaac and Shapiro (2010) demonstrate that CRAs incentives to produce accurate ratings are likely to be countercyclical, i.e. lower in a boom than in a recession. In Pagano and Volpin (2008), CRAs have no conflicts of interest, but can choose ratings to be more or less opaque depending on what the issuer asks for. They show that opacity can enhance liquidity in the primary market, but may cause a market freeze in the secondary market.

In empirical evidence, Mathis *et al.* (2009) show that, controlling for economic variables, the fraction of structured finance tranches that were rated AAA has increased over the period 2000–2008. Ashcraft *et al.* (2009) examine subprime and Alt-A mortgage backed securities (MBS) during the period leading up to the subprime crisis and find evidence that ratings become less conservative right at the height of the MBS

market peak in 2005–2007. In particular, they demonstrate that ratings quality was worse on low documentation mortgages. Griffin and Tang (2009) look at CRA adjustments to their models' predictions of credit risk in the CDO market and find that the adjustments were overwhelmingly positive, were positively related with future downgrades, and the amount adjusted increased sharply from 2003 to 2007. Benmelech and Dlugosz (2009) find that securities rated by only one agency were 6.1% more likely to be subsequently downgraded and point to shopping as the reason.

JOEL SHAPIRO

*See Also* **barriers to entry; bonds; public debt; reputation.**

*I thank Larry White for helpful comments.*

### *Bibliography*

Adelino, M. 2009. *Do Investors Rely Only on Ratings? The Case of Mortgage-Backed Securities*, Mimeo: MIT.

Ashcraft, A., Goldsmith-Pinkham, P. and Vickery, J. 2009. *MBS ratings and the mortgage credit boom*. Mimeo: Federal Reserve Bank of New York.

Bar-Isaac, H. and Shapiro, J. 2010. *Ratings Quality over the Business Cycle*. Mimeo: NYU.

Becker, B. and Milbourn, T. 2009. *Reputation and competition: evidence from the credit rating industry*. Mimeo: Harvard Business School.

Benmelech, E. and Dlugosz, J. 2009. *The credit rating crisis*. Mimeo: Harvard University.

Berger, A., Davies, S. and Flannery, M. 2000. Comparing market and supervisory assessments of bank performance: who knows what when?*Journal of Money, Credit and Banking* 32, 641–667.

Bolton, P., Freixas, X. and Shapiro, J. 2010. *The credit ratings game*. Mimeo: Columbia University.

Bongaerts, D., Cremer, K. J. M. and Goetzmann, W. 2009. *Tiebreaker: certification and multiple credit ratings*. Mimeo: Yale University.

Cantor, R. and Packer, F. 1995. The credit rating industry. *Journal of Fixed Income*, 1–26.

Committee on the Global Finance System, Bank for International Settlement. 2005. *The Role of Ratings in Structured Finance: Issues and Implications*. http://www.bis.org/publ/cgfs23.pdf?noframes=1.

Coval, J. D., Jurek, J. W. and Stafford, E. 2009. Economic catastrophe bonds. *American Economic Review* 99, 628–666.

Dichev, I. and Piotroski, J. 2001. The long-run stock returns following bond ratings changes. *Journal of Finance* 56, 173–203.

Faure-Grimaud, A., Peyrache, E. and Quesada, L. 2009. The ownership of ratings. *RAND Journal of Economics* 40(2), 234–257.

Griffin, J. M. and Tang, D. Y. 2009. *Did subjectivity play a role in CDO credit ratings?* Mimeo: UT-Austin.

Hand, J., Holthausen, R. and Leftwich, R. 1992. The effect of bond rating agency announcements on bond and stock prices. *Journal of Finance* 47(2), 733–752.

Hilscher, J. and Wilson, M. 2009. *Credit ratings and credit risk*. Mimeo: Oxford University.

Hite, G. and Warga, A. 1997. The effect of bond-rating changes on bond price performance. *Financial Analysts Journal* (May/June), 35–51.

Jewell, J. and Livingston, M. 1999. A comparison of bond ratings from Moody's, S&P and Fitch. *Financial Markets, Institutions and Instruments* 8, 1–45.

Jorion, P. and Zhang, G. 2007. Information effects of bond rating changes: the role of the rating prior to the announcement. *Journal of Fixed Income* 16, 45–59.

Kliger, D. and Sarig, O. 2000. The information value of bond ratings. *Journal of Finance* 55, 2879–2902.

Kisgen, D. J. and Strahan, P. 2010. Do Regulations Based on Credit Ratings Affect a Firm's Cost of Capital? Review of Financial Studies, forthcoming.

Mathis, J., McAndrews, J. and Rochet, J. C. 2009. Rating the raters: are reputation concerns powerful enough to discipline rating agencies?*Journal of Monetary Economics* 56(5), 657–674.

Pagano, M. and Volpin, P. 2008. *Securitization, transparency, and liquidity*. Mimeo: Università di Napoli Federico II and London Business School.

Pagano, M. and Volpin, P. 2009. *Credit rating failures: causes and policy options. Macroeconomic Stability and Financial Regulation: Key Issues for the G20*, CEPR.

Partnoy, F. 2002. The paradox of credit ratings. In *Ratings, Rating Agencies and the Global Financial System* (eds.) R. M. Levich, G. Majnoni and C. Reinhart.Kluwer: Boston.

Poon, W. P. H. 2003. Are unsolicited credit ratings biased downward?. *Journal of Banking and Finance* 27, 593–614.

Skreta, V. and Veldkamp, L. 2009. Ratings shopping and asset complexity: a theory of ratings inflation.*Journal of Monetary Economics* 56(5), 678–695.

United States Securities and Exchange Commission. 2008. *Summary Report of Issues Identified in the Commission Staff's Examinations of Select Credit Rating Agencies*. Washington, SEC.

White, L. 2002. The credit rating industry an industrial organization analysis. In *Ratings, Rating Agencies and the Global Financial System* (eds.) R. M. Levich, G. Majnoni and C. Reinhart. Kluwer: Boston.

White, L. 2010. The credit rating agencies: how did we get here? Where should we go? *Journal of Economic Perspectives* (forthcoming).

## currency crises

A currency crisis occurs when investors flee from a currency en masse out of fear that it might be devalued. Currency crises are episodes characterized by sudden depreciations of the domestic currency, large losses of foreign exchange reserves of the central bank, and (or) sharp hikes in domestic interest rates.

There have been numerous currency crises since 1980. The so-called debt crisis erupted in 1982 following Mexico's default and devaluation in August. This crisis spread rapidly to all Latin American countries, and by the time it was over, most Latin American countries had devalued their currencies and defaulted on their foreign debts. The debt crisis was followed by a decade of negative growth and isolation from international capital markets. The output costs of this crisis were so large that the 1980s became known as the 'lost decade' for Latin America.

Crises are not just emerging-market phenomena. The 1990s opened with crises in industrial Europe – the European Monetary System (EMS) crises of 1992 and 1993. By the end of these crises, in the summer of 1993, the lira and the sterling had been driven from the Exchange Rate Mechanism (ERM); Finland, Norway, and Sweden had abandoned their unofficial peg to the European Currency Unit (ECU); the Spanish peseta, the Portuguese escudo and the Irish punt had devalued; and Europe's central bank governors and finance ministers had widened the ERM's intervention margins to $\pm 15$ per cent from $\pm 2.25$ per cent. Only then did the currency market stabilize.

Crises are hardy perennials. Within one year of the EMS crises, a currency crisis exploded in Mexico, with currency jitters spreading around the Latin American region. In 1997, it was Asia's turn. A new episode of currency turbulences started in July of that year with the depreciation of the Thai baht. Within a few days the crisis had spread to Indonesia, Korea, Malaysia and the Philippines. Turmoil in the foreign exchange market heightened in 1998 with the Russian default and devaluation in August. The Russian crisis spread around the world with speculative attacks in economies as far apart as South Africa, Brazil and Hong Kong. Currency crises have continued to erupt in the new millennium, with Argentina's crisis in December 2001 including the largest foreign-debt default in history.

The numerous financial crises that have ravaged emerging markets as well as mature economies have fuelled a continuous interest in developing models to explain why speculative attacks occur. Models are even catalogued into three generations. The first-generation models focus on the fiscal and monetary causes of crises. These models were mostly developed to explain the crises in Latin America in the 1960s and 1970s. In these models, unsustainable money-financed fiscal deficits lead to a persistent loss of international reserves and ultimately to a currency crash (see, for example, Krugman, 1979).

The second-generation models aim at explaining the EMS crises of the early 1990s. These models focus on explaining why currency crises tend to happen in the midst of unemployment and loss of competitiveness. To explain these links, governments are

modelled facing two targets: reducing inflation and keeping economic activity close to a given target. Fixed exchange rates may help in achieving the first goal but at the cost of a loss of competitiveness and a recession. With sticky prices, devaluations restore competitiveness and help in the elimination of unemployment, thus prompting the authorities to abandon the peg during recessions. Importantly, in this setting of counter-cyclical policies, the possibility of self-fulfilling crises becomes important, with even sustainable pegs being attacked and frequently broken (see, for example, Obstfeld, 1994).

The next wave of currency crises, the Mexican crisis in 1994 and the Asian crisis in 1997, fuelled a new variety of models – also known as third-generation models – which focus on moral hazard and imperfect information. The emphasis here has been on 'excessive' booms and busts in international lending and asset price bubbles. These models also link currency and banking crises, sometimes known as the 'twin crises' (Kaminsky and Reinhart, 1999). For example, Diaz-Alejandro (1985) and Velasco (1987) model difficulties in the banking sector as giving rise to a balance of payments crisis, arguing that, if central banks finance the bail-out of troubled financial institutions by printing money, we have the classical story of a currency crash prompted by excessive money creation. Within the same theme, McKinnon and Pill (1995) examines the role of capital flows in an economy with an unregulated banking sector with deposit insurance and moral hazard problems of the banks. Capital inflows in such an environment can lead to over-lending cycles with consumption booms, real exchange rate appreciations, exaggerated current account deficits, and booms (and later busts) in stocks and property markets. Importantly, the excess lending during the boom makes banks more prone to a crisis when a recession unfolds. In turn, the fragile banking sector makes the task of defending the peg by hiking domestic interest rates more difficult and may lead to the eventual collapse of the domestic currency. Following the crisis in Argentina in 2001, the links between debt sustainability, sovereign defaults, and currency crises again attracted the attention of the economics profession. Finally, currency crises have also been linked to the erratic behaviour of international capital markets. For example, Calvo (1998) has brought to general attention the possibility of liquidity crises in emerging markets due to sudden reversals in capital flows, in large part triggered by developments in the world financial centres.

To summarize, all models suggest that currency crises erupt in fragile economies. Importantly, the three generations of models conclude that vulnerabilities come in different varieties. Still, the first attempts to study the vulnerabilities that precede crises have adopted 'the one size fits all' approach (see, for example, Frankel and Rose, 1996; and Kaminsky, 1998). That is, the regressions estimated to predict crises include all possible indicators of vulnerability. These indicators include those related to sovereign defaults, such as high foreign debt levels, or indicators related to fiscal crises, such as government deficits, or even indicators related to crises of financial excesses, such as stock and real estate market booms and busts. In all cases, researchers impose the same functional form on all observations. When some indicators are not robustly linked to all crises, they tend to be discarded even when they may be of key importance for a subgroup of crises. Naturally, these methods leave many crises unpredicted and, furthermore, cannot capture the evolving nature of currency crises.

The next step in the empirical analysis of crises should be centred on whether crises are of different varieties. The first attempt in this direction is in Kaminsky (2006). In this article, a different methodology is used to allow for *ex ante* unknown varieties of currency crises. To identify the possible multiple varieties of crises, regression tree analysis is applied. This technique allows us to search for an unknown number of varieties of crises and of tranquil times using multiple indicators. This technique was also applied to growth by Durlauf and Johnson (1995).

Interestingly, this method catalogues crises into six classes:

1. *Crises with current account problems*. This variety is characterized by just one type of vulnerability, that of loss of competitiveness, that is, real exchange rate appreciations.
2. *Crises of financial excesses*. The fragilities are associated with booms in financial markets. In particular, they are identified as crises that are preceded by the acceleration in the growth rate of domestic credit and other monetary aggregates.
3. *Crises of sovereign debt problems*. These crises are characterized by fragilities associated with 'unsustainable' foreign debt.
4. *Crises with fiscal deficits*. This variety is just related to expansionary fiscal policy.
5. *Sudden-stop crises*. This type of crisis is only associated with reversals in capital flows triggered by sharp hikes in world interest rates, with no domestic vulnerabilities.
6. *Self-fulfilling crises*. This class of crises is not associated with any evident vulnerability, domestic or external.

These estimations allow us to answer four important questions about crises.

1. *Do crises occur in countries with sound fundamentals?* Even though this estimation allows for the identification of self-fulfilling crises (crises in economies with sound fundamentals), the results indicate that basically all crises are preceded by domestic or external vulnerabilities. Only four per cent of the crises are unrelated to economic fragilities.
2. *How important are sudden reversals in capital flows in triggering crises?* While many have stressed that the erratic behaviour of international capital markets is the main culprit in emerging market currency crises, only two per cent of the crises in developing countries are just triggered by sudden-stop problems. While sudden-stop problems do occur, the reversals in capital flows mostly occur in the midst of multiple domestic vulnerabilities (see, Calvo, Izquierdo and Talvi, 2004).
3. *Are crises different in emerging economies?* Crises in emerging markets are preceded by far more domestic vulnerabilities than those in industrial countries. Overall, 86 per cent of the crises in emerging economies are crises with multiple domestic vulnerabilities, while economic fragility characterizes only 50 per cent of the crises in mature markets.
4. *Are some crises more costly than others?* It is a well-established fact that financial crises impose substantial costs on society. Many economists have emphasized the output losses associated with crises. But these are not the only costs of crises. In the aftermath of crises, most countries lose access to international capital markets, losing the ability to reduce the effect of adverse income shocks by borrowing in

international capital markets. In most cases, countries have to run current account surpluses to pay back their debt. Finally, the magnitude of the speculative attack is itself important. For example, large depreciations may cause adverse balance sheet effects on firms and governments when their liabilities are denominated in foreign currencies. *Crises of financial excesses*, those also associated with banking crises – twin crisis episodes – are the costliest. Not only does the domestic currency depreciate the most, but also output losses are higher and the reversal of the current account deficit is attained via a dramatic fall in imports. In the aftermath of these crises, exports fail to grow even though the depreciations in this type of crises are massive. This evidence suggests that countries are even unable to attract trade credits to finance exports when their economies are mired in financial problems. In contrast, *self-fulfilling crises* and *sudden-stop crises* (but with no domestic vulnerabilities) have no adverse effects on the economies. Output (relative to trend) is unchanged or continues to grow in the aftermath of crises with no observed domestic fragility. In these crises, booming exports are at the heart of the recovery of the current account.

GRACIELA LAURA KAMINSKY

*See Also* **currency crises models.**

### *Bibliography*

Calvo, G. 1998. Capital flows and capital-market crises: the simple economics of sudden stops. *Journal of Applied Economics* 1, 35–54.

Calvo, G., Izquierdo, A. and Mejia, L. 2004. *On the empirics of sudden stops: the relevance of balance-sheet effects*. Working paper No. 10520. Cambridge, MA: NBER.

Diaz-Alejandro, C. 1985. Good-bye financial repression, hello financial crash. *Journal of Development Economics* 19, 1–24.

Durlauf, S. and Johnson, P. 1995. Multiple regimes and cross-country growth behavior. *Journal of Applied Econometrics* 10, 365–84.

Frankel, J. and Rose, A. 1996. Currency crises in emerging markets: an empirical treatment. *Journal of International Economics* 41, 351–66.

Kaminsky, G.L. 1998. Currency and banking crises: the early warnings of distress. International Finance Discussion Papers No. 629. Board of Governors of the Federal Reserve System.

Kaminsky, G.L. and Reinhart, C. 1999. The twin crises: the causes of banking and balance-of-payments problems. *American Economic Review* 89, 473–500.

Kaminsky, G.L. 2006. Currency crises: are they all the same? *Journal of International Money and Finance* 25, 503–27.

Krugman, P. 1979. A model of balance-of-payments crises. *Journal of Money, Credit, and Banking* 11, 311–25.

McKinnon, R.I. and Pill, H. 1995. Credible liberalizations and international capital flows: the 'overborrowing syndrome'. In *Financial Deregulation and Integration in East Asia*, ed. T. Ito and A.O. Krueger. Chicago: University of Chicago Press.

Obstfeld, M. 1994. The logic of currency crises. *Cahiers Economiques et Monétaires* 43, 189–213.

Obstfeld, M. 1996. Models of currency crises with self-fulfilling features. *European Economic Review* 40, 1037–47.

Velasco, A. 1987. Financial and balance-of-payments crises. *Journal of Development Economics* 27, 263–83.

## currency crises models

There have been many currency crises during the post-war era (see Kaminsky and Reinhart, 1999). A currency crisis is an episode in which the exchange rate depreciates substantially during a short period of time. There is an extensive literature on the causes and consequences of a currency crisis in a country with a fixed or heavily managed exchange rate. The models in this literature are often categorized as first-, second- or third-generation.

In first-generation models the collapse of a fixed exchange rate regime is caused by unsustainable fiscal policy. The classic first-generation models are those of Krugman (1979) and Flood and Garber (1984). These models are related to earlier work by Henderson and Salant (1978) on speculative attacks in the gold market. Important extensions of these early models incorporate consumer optimization and the government's intertemporal budget constraint into the analysis (see Obstfeld, 1986; Calvo, 1987; Drazen and Helpman, 1987; Wijnbergen, 1991). Flood and Marion (1999) provide a detailed review of first-generation models.

In a fixed exchange rate regime a government must fix the money supply in accordance with the fixed exchange rate. This requirement severely limits the government's ability to raise seigniorage revenue. A hallmark of first-generation models is that the government runs a persistent primary deficit. This deficit implies that the government must either deplete assets, such as foreign reserves, or borrow to finance the deficit. It is infeasible for the government to borrow or deplete reserves indefinitely. Therefore, in the absence of fiscal reforms, the government must eventually finance the deficit by printing money to raise seigniorage revenue. Since printing money is inconsistent with keeping the exchange rate fixed, first-generation models predict that the regime must collapse. The precise timing of its collapse depends on the details of the model.

The key ingredients of a first-generation model are its assumptions regarding purchasing power parity (PPP), the government budget constraint, the timing of deficits, the money demand function, the government's rule for abandoning the fixed exchange rate, and the post-crisis monetary policy. In the simplest first-generation models there is a single good whose domestic currency price is $P_t$ and whose foreign currency price is 1. Let $S_t$ denote the nominal exchange rate. PPP implies $P_t = S_t$. Suppose for simplicity that the government has a constant ongoing primary deficit, $\delta$. It finances this deficit by reducing its stock of foreign reserves, $f_t$, which can either evolve as a smooth function of time or jump discontinuously. In the former case, $f_t$ evolves according to $\dot{f}_t = rf_t - \delta + \dot{M}_t/S_t$, where $r$ is the real interest rate, $M_t$ is the monetary base, and a dot over a variable denotes its derivative with respect to time. When foreign reserves change discontinuously, $\Delta f_t = \Delta(M_t/S_t)$. When $\delta > rf_0$ interest income from foreign assets will not be sufficient to finance the deficit.

To illustrate the key properties of first-generation models, we make three simplifying assumptions. First, money demand takes the Cagan (1956) form,

$M_t = \theta P_t exp[-\eta(r + \pi_t)]$, where $\theta > 0$ and $\pi_t = \dot{P}_t/P_t$ is the inflation rate. Second, the government abandons the fixed exchange rate regime when its foreign reserves are exhausted. Third, as soon as foreign reserves are exhausted, the government prints money at a constant rate $\mu$ to fully finance its deficit.

These assumptions imply that after the crisis the level of real balances, $m_t = M_t/P_t$, is constant and equal to $\overline{m} = \theta exp[-\eta(r + \mu)]$. The post-crisis government budget constraint reduces to $\delta = \mu\overline{m}$. This equation determines $\mu$. Let $t^*$ denote the date at which foreign reserves are exhausted and the government abandons the fixed exchange rate regime. PPP implies $S_{t^*} = P_{t^*} = \overline{M}/\overline{m}$, where $\overline{M}$ is the monetary base the instant after date $t^*$. Under perfect foresight the exchange rate cannot jump discontinuously at $t^*$ since such a jump would imply the presence of arbitrage opportunities. Given that the exchange rate must be a continuous function of time at $t^*$, $S_{t^*} = S$ and $\overline{M} = \overline{m}S$.

Prior to the crisis real balances are given by $m = \theta exp(-\eta r)$. Therefore, at date $t^*$ there is a sudden drop in real money demand from $m$ to $\overline{m}$ implying that reserves drop discontinuously to zero at time $t^*$: $\Delta f_{t^*} = \overline{m} - m$. This is why the literature refers to $t^*$ as the date of the speculative attack. Prior to the crisis the government's reserves fall at the rate $\dot{f}_t = rf_t - \delta$. The budget constraint implies that $t^* = ln\{[\delta - r(m - \overline{m})]/(\delta - rf_0)\}/r$. While the collapse of the fixed exchange rate regime is inevitable, it does not generally occur at time zero unless $m - \overline{m} > f_0$.

A shortcoming of this type of first-generation model is that the timing of the speculative attack is deterministic and the exchange rate does not depreciate at the time of the attack. These shortcomings can be remedied by introducing shocks into the model, as in Flood and Garber (1984).

Early first-generation models predict that ongoing fiscal deficits, rising debt levels, or falling reserves precede the collapse of a fixed exchange rate regime. This prediction is inconsistent with the 1997 Asian currency crisis. This inconsistency led many observers to dismiss fiscal explanations of this crisis. However, Corsetti, Pesenti and Roubini (1999), Burnside, Eichenbaum and Rebelo (2001a), and Lahiri and Végh (2003) show that bad news about prospective deficits can trigger a currency crisis. Under these circumstances a currency crisis will not be preceded by persistent fiscal deficits, rising debt levels, or falling reserves. These models assume that agents receive news that the banking sector is failing and that banks will be bailed out by the government. The government plans to finance, at least in part, the bank bailout by printing money beginning at some time in future. Burnside, Eichenbaum and Rebelo (2001a) show that a currency crisis will occur before the government actually starts to print money. Therefore, in their model, a currency crisis is not preceded by movements in standard macroeconomic fundamentals, such as fiscal deficits and money growth. Burnside, Eichenbaum and Rebelo argue that their model accounts for the main characteristics of the Asian currency crisis.

This explanation of the Asian currency crisis stresses the link between future deficits and current movements in the exchange rate. This link is also stressed by Corsetti and Mackowiak (2006), Daniel (2001), and Dupor (2000), who use the fiscal theory of the price level to argue that prices and exchange rates jump in response to news about future deficits.

In first-generation models the government follows an exogenous rule to decide when to abandon the fixed exchange rate regime. In second-generation models the government maximizes an explicit objective function (see, for example, Obstfeld, 1994; 1996). This maximization problem dictates if and when the government will abandon the fixed exchange rate regime. Second-generation models generally exhibit multiple equilibria so that speculative attacks can occur because of self-fulfilling expectations. In Obstfeld's models (1994; 1996) the central bank minimizes a quadratic loss function that depends on inflation and on the deviation of output from its natural rate (see Barro and Gordon, 1983, for a discussion of this type of loss function). The level of output is determined by an expectations-augmented Phillips curve. The government decides whether to keep the exchange rate fixed or not. Suppose agents expect the currency to devalue and inflation to ensue. If the government does not devalue then inflation will be unexpectedly low. As a consequence output will be below its natural rate. Therefore the government pays a high price, in terms of lost output, in order to defend the currency. If the costs associated with devaluing (lost reputation or inflation volatility) are sufficiently low, the government will rationalize agents' expectations. In contrast, if agents expect the exchange rate to remain fixed, it can be optimal for the government to validate agents' expectations if the output gains from an unexpected devaluation are not too large. Depending on the costs and benefits of the government's actions, and on agents' expectations, there can be more than one equilibrium. See Jeanne (2000) for a detailed survey of second-generation models.

Morris and Shin (1998) provide an important critique of models with self-fulfilling speculative attacks. They emphasize that standard second-generation models assume that fundamentals are common knowledge. Morris and Shin demonstrate that introducing a small amount of noise into agents' signals about fundamentals will lead to a unique equilibrium.

Many currency crises coincide with crises in the financial sector (Diaz-Alejandro, 1985; Kaminsky and Reinhart, 1999). This observation has motivated a literature that emphasizes the role of the financial sector in causing currency crises and propagating their effects. These third-generation models emphasize the balance-sheet effects associated with devaluations. The basic idea is that banks and firms in emerging market countries have explicit currency mismatches on their balance sheets because they borrow in foreign currency and lend in local currency. Banks and firms face credit risk because their income is related to the production of non-traded goods whose price, evaluated in foreign currency, falls after devaluations. Banks and firms are also exposed to liquidity shocks because they finance long-term projects with short-term borrowing. Eichengreen and Hausmann (1999) argue that currency mismatches are an inherent feature of emerging markets. In contrast, authors such as McKinnon and Pill (1996) and Burnside, Eichenbaum and Rebelo (2001b) argue that, in the presence of government guarantees, it is optimal for banks and firms to expose themselves to currency risk.

Different third-generation models explore various mechanisms through which balance-sheet exposures may lead to a currency and banking crisis. In Burnside, Eichenbaum and Rebelo (2004) government guarantees lead to the possibility of self-fulfilling speculative attacks. In Chang and Velasco (2001) liquidity exposure leads to the possibility of a Diamond and Dybvig (1983) style bank run. In Caballero and

Krishnamurthy (2001) firms face a liquidity problem because they finance risky long-term projects with foreign loans but have access to limited amounts of internationally accepted collateral.

An important policy question is: what is the optimal nature of interest rate policy during and after a currency crisis? There has been relatively little formal work on this topic. Christiano, Braggion and Roldos (2006) take an important first step in this direction. They argue that it is optimal to raise interest rates during a currency crisis and to lower them immediately thereafter. Studying optimal monetary policy in different models of currency crises remains an important area for future research.

CRAIG BURNSIDE, MARTIN EICHENBAUM AND SERGIO REBELO

*See Also* **currency crises; fiscal theory of the price level.**

### *Bibliography*

Barro, R. and Gordon, D. 1983. A positive theory of monetary policy in a natural rate model. *Journal of Political Economy* 91, 589–610.

Burnside, C., Eichenbaum, M. and Rebelo, S. 2001a. Prospective deficits and the Asian currency crisis. *Journal of Political Economy* 109, 1155–98.

Burnside, C., Eichenbaum, M. and Rebelo, S. 2001b. Hedging and financial fragility in fixed exchange rate regimes. *European Economic Review* 45, 1151–93.

Burnside, C., Eichenbaum, M. and Rebelo, S. 2004. Government guarantees and self-fulfilling speculative attacks. *Journal of Economic Theory* 119, 31–63.

Caballero, R. and Krishnamurthy, A. 2001. International and domestic collateral constraints in a model of emerging market crises. *Journal of Monetary Economics* 48, 513–48.

Cagan, P. 1956. Monetary dynamics of hyperinflation. In *Studies in the Quantity Theory of Money*, ed. M. Friedman. Chicago: University of Chicago Press.

Calvo, G. 1987. Balance of payments crises in a cash-in-advance economy. *Journal of Money Credit and Banking* 19, 19–32.

Chang, R. and Velasco, A. 2001. A model of financial crises in emerging markets. *Quarterly Journal of Economics* 116, 489–517.

Christiano, L., Braggion, F. and Roldos, J. 2006. The optimal monetary response to a financial crisis. Mimeo, Northwestern University.

Corsetti, G. and Mackowiak, B. 2006. Fiscal imbalances and the dynamics of currency crises. *European Economic Review* 50, 1317–38.

Corsetti, G., Pesenti, P. and Roubini, N. 1999. What caused the Asian currency and financial crisis? *Japan and the World Economy* 11, 305–73.

Daniel, B. 2001. The fiscal theory of the price level in an open economy. *Journal of Monetary Economics* 48, 293–308.

Diamond, D. and Dybvig, P. 1983. Bank runs, deposit insurance and liquidity. *Journal of Political Economy* 91, 401–19.

Diaz-Alejandro, C. 1985. Good-bye financial repression, hello financial crash. *Journal of Development Economics* 19, 1–24.

Drazen, A. and Helpman, E. 1987. Stabilization with exchange rate management. *Quarterly Journal of Economics* 102, 835–55.

Dupor, W. 2000. Exchange rates and the fiscal theory of the price level. *Journal of Monetary Economics* 45, 613–30.

Eichengreen, B. and Hausmann, R. 1999. Exchange rates and financial fragility. In *New Challenges for Monetary Policy: A Symposium Sponsored by the Federal Reserve Bank of Kansas City*. Kansas City: Federal Reserve Bank of Kansas City.

Flood, R. and Garber, P. 1984. Collapsing exchange rate regimes: some linear examples. *Journal of International Economics* 17, 1–13.
Flood, R. and Marion, N. 1999. Perspectives on the recent currency crisis literature. *International Journal of Finance and Economics* 4, 1–26.
Henderson, D. and Salant, S. 1978. Market anticipations of government policies and the price of gold. *Journal of Political Economy* 86, 627–48.
Jeanne, O. 2000. Currency crises: a perspective on recent theoretical developments. Special Papers in International Economics, No. 20, International Finance Section, Princeton University.
Kaminsky, G. and Reinhart, C. 1999. The twin crises: the causes of banking and balance-of-payments problems. *American Economic Review* 89, 473–500.
Krugman, P. 1979. A model of balance of payments crises. *Journal of Money, Credit and Banking* 11, 311–25.
Lahiri, A. and Végh, C. 2003. Delaying the inevitable: interest rate defense and BOP crises. *Journal of Political Economy* 111, 404–24.
McKinnon, R. and Pill, H. 1996. Credible liberalizations and international capital flows: the overborrowing syndrome. In *Financial Deregulation and Integration in East Asia*, ed. T. Ito and A. Krueger. Chicago: University of Chicago Press.
Morris, S. and Shin, H. 1998. Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review* 88, 587–97.
Obstfeld, M. 1986. Speculative attack and the external constraint in a maximizing model of the balance of payments. *Canadian Journal of Economics* 29, 1–20.
Obstfeld, M. 1994. The logic of currency crises. *Cahiers Economiques et Monétaires* 43, 189–213.
Obstfeld, M. 1996. Models of currency crises with self-fulfilling features. *European Economic Review* 40, 1037–47.
Wijnbergen, S. van. 1991. Fiscal deficits, exchange rate crises and inflation. *Review of Economic Studies* 58, 81–92.

# euro zone crisis 2010

## Introduction

The euro zone crisis started in early 2010 when it emerged that the Greek government had for years doctored the official data on its deficits and debt. The figures for the deficit and debt level presented by the new government were so much higher than the previous ones that rating agencies and many market participants downgraded their assessment of Greece's ability to service its debt fully. As a result, the cost of refinancing the Greek debt increased sharply and the government could not secure the resources needed to fund its current deficit and roll over the portion of the debt coming due. By the end of April 2010 it had to be bailed out with a €110 billion programme.

The second stage of the crisis came about six months later when it emerged that the Irish government had been 'misled' about the scale of the losses in its banks. As the Irish government had guaranteed all the liabilities of its banks it was now itself on the brink of insolvency. Moreover (although this was not made public at the time), the ECB had become uncomfortable with the huge exposure it had to Irish banks, which had become totally dependent on central bank financing. The ECB therefore pushed the Irish government to recapitalize its banks, but this could be done only with outside help. The Irish government had thus little choice but to apply for external financial support.

With the Greek and Irish bailouts, the euro zone has shown the world two pure specimens of financial crisis: one originated by the mismanagement of fiscal policy (Greece), the other by mismanagement of a credit bubble and banking supervision (Ireland). The Portuguese crisis, which emerged in early 2011, seems to represent a hybrid specimen: a combination of a fiscal crisis (like Greece) and a private debt crisis (like Ireland).

## A brief chronology

Although Greece accounts for a small portion (less than 3%) of the euro area GDP (and even less of its banking assets), in early 2010 financial markets reacted strongly to the prospect of a sovereign insolvency. A first consequence of the realization that Greece would not be solvent without external financial support was that investors started to price more widely government solvency in the bond market. As a result, the risk premia on the debt of other countries with weak fundamentals also rose. But more important was a generalized increase in risk aversion, which led to a fall in the prices of all risky assets in a similar vein (but of course a much less severe magnitude) as after the collapse of Lehman Brothers in late 2008.

The European banking sector was particularly affected because it was widely believed that a number of banks would not survive a default by Greece. However, which banks held how much of Greek debt was not known. In an environment of widespread risk

aversion and many highly leveraged banks this resulted in a drying up of parts of the interbank market, which performs a vital role in the financial system.

The German government reiterated on several occasions its aversion to a bailout, stressing that this must be only an *ultima ratio* mechanism. But when faced with the spectre of a 'second Lehman crisis' and the prospect of large losses in the weak German banks heavily exposed to Greece and other peripheral countries, it had no choice but agree to a rescue package of about €110 billion. This is an EU/IMF rescue package according to which the IMF provides support under a three-year €30 billion standby arrangement (the IMF's standard lending instrument) while euro area members pledge a total of €80 billion in bilateral loans against the implementation of strict austerity measures monitored by the IMF. The sum agreed is supposed to fully finance Greece's remaining deficits (and rollover obligations) during the following three years. It was assumed then (on the basis of experience with 'normal' IMF programs) that Greece would be able to access private capital markets at reasonable rates towards the end of this period. However, in early 2011 it became clear that the hypothesis was far too optimistic. In March the terms and the conditions of the loans to Greece were reviewed to include an extension of the maturity and lower interest rates.

In the spring of 2010, Europe's leaders also thought that Greece was a unique and special case and that no other country would ever need financial support. However, only a few days after the Greek rescue, financial markets went into such a tailspin (risk premia rose, some markets ceased to function) that a new and much larger financing mechanism had to be hastily created.

During the dramatic weekend of 9 May 2010, two financing mechanisms were set up in order to allow the authorities to react to future financial crises in a more coordinated and organized manner. The headline figure of the total potential funding was €750 billion, to be provided by three different entities: €60 billion, guaranteed by the EU budget, coming through a newly created European Financial Stabilization Mechanism (EFSM); €440 billion, guaranteed on a pro rata basis by euro area member states, coming through the also newly created European Financial Stability Facility (EFSF); and up to €250 billion from the IMF.

Together with the ECB interventions in the euro area public and private debt securities markets (Securities Markets Programme) aiming at ensuring liquidity in those market segments judged to be 'dysfunctional', this package did restore stability in the financial markets for a few weeks.

In early June 2010, since tensions in the interbank market persisted, member states and the European Institutions (Commission and Committee of European Banking Supervisors, CEBS) agreed to make public for the first time the results of ongoing stress tests for major European banks.[1]

The rationale for the tests was to disclose information about the state of the European banking system in order to dissipate doubts about their resilience. The Spanish supervisory authorities were particularly keen on this move because they hoped that by showing that their banks were 'safe and sound', it would be easier for Spanish

[1] http://stress-test.c-ebs.org/documents/Summaryreport.pdf

banks to regain access to the interbank market. More generally, the publication of the stress tests was supposed to prove that the most important banks had sufficient capital to withstand even a so-called 'adverse' scenario. This should have improved confidence in the banking system in general.

Yet the objective of the exercise was achieved only temporarily.[2] During the summer of 2010 risk premia on the government bonds of the four 'fiscally challenged' countries (Portugal, Ireland, Greece and Spain) started to increase again. This accelerated after a Franco-German agreement in Deauville on economic governance and the decision by the European Council of 28 October to establish a permanent crisis mechanism to safeguard the financial stability of the euro area. This decision proved to be a watershed because it suggested a change in the ground rules of peripheral euro area debt markets: on that occasion all 27 Member States agreed on the proposal (then submitted to the European Council and implemented in early 2011) for a limited, technical Treaty amendment to provide a legal basis for establishing a permanent crisis mechanism. In March 2011, the European Council adopted the basic features of the new device: the European Stability Mechanism (ESM). The ESM, which will be operational as of mid-2013, is based on the existing EFSF but, unlike the EFSF, the provision of liquidity is conditional to a debt sustainability assessment (conducted by the European Commission and the IMF, in liaison with the ECB). In the event that the analysis reveals that a member state is insolvent, the country is expected to negotiate a comprehensive plan with its private creditors. Moreover collective action clauses (CACs) will be included in the terms and conditions of all new euro area sovereign bonds, starting in June 2013. These clauses should provide the legal basis for the negotiation process with creditors and enable them to pass by qualified majority a decision agreeing a legally binding change to the terms of payment. This could take different forms (standstill, extension of maturity, interest-rate cut and/or haircut) depending on the specific case, but clearly implies that if losses materialize they will be borne, at least partially, by the private sector.

Financial markets did not welcome this approach, and Ireland became the first victim of deteriorating market conditions. Indeed, market pressures on Ireland had started mounting in October 2010, when the Irish government decided to rescue some of its banks that had published losses that were considerably higher than estimated a few months earlier. The high costs of this bank bailout program resulted in a deficit of 32% of GDP, and the risk premia on Irish government (and bank) bonds shot up. As a consequence the Irish government quickly had to ask for external support. On 28 November, an €85bn financial assistance package was agreed and Ireland committed to a sweeping restructuring of its banking system and even more sweeping budget cuts. According to the rescue plan, the EU provides financial assistance for €45bn, through the European Financial Stability Mechanism and the European Financial Stability Fund, together with bilateral loans from the UK (€3.8bn), Sweden (€0.6bn) and Denmark (€0.4bn). The IMF provides €22.5bn and the Irish sovereign €17.5bn through the Treasury cash buffer and investments of the National Pension Reserve Fund. It was

[2]See, among others, Veron (2010) and Blundell-Wignall and Slovik (2010).

also agreed that more than one third of the total package (35bn) was to be destined to recapitalization measures in support of the banking system.

After some hesitation, the Irish parliament did ratify the bailout agreement, but the government fell, new elections were set for 25 February 2011 and resulted in the victory of the opposition who had promised to renegotiate the agreement.

The Irish bailout (as that of Greece) did not have an immediate impact on risk premia and interest rates did not fall (nor for other countries). If anything, the Irish crisis had two major consequences. First it discredited completely the results of the banks' stress tests, as in July 2010 only six small banks had not passed the test and Allied Irish Bank and Bank of Ireland, the two largest Irish banks, both passed the test (Anglo Irish Bank was not included in the tests). Second, it did not allay concerns about the sustainability of Irish debt because the interest charged (close to 6%) on the EFSF loans is much higher than the growth rate Ireland could hope to achieve.

The brief review of the chronology of the crisis shows that Greece was just a trigger and the euro zone crisis is in fact a complex tangle of sovereign debt and banking crises.

The Irish experience has shown that even a government with a strong fiscal position (budget surplus during boom and low initial debt level) can become insolvent in the attempt to save insolvent banks. The sequence of events in Ireland is archetypal: a property bubble ending with a bust leaves a massive housing overhang. This leads to huge losses in banks which had fuelled the bubble with excessive lending. As often happens, the local regulators pretend that there is no problem; but as the losses mount investors pull the plug and the risk of collapse of the entire system increases. This is what happened during the late summer of 2010: as banks were shut out of the interbank market and depositors started to withdraw their funds, the Irish government decided to stand behind the banks and put the entire nation at risk, transforming a banking crisis into the second sovereign debt crisis in the euro zone. A third case of crisis has emerged in early 2011. Portugal has not experienced a bubble as Ireland, neither its fiscal stance is as bad as the Greek one, but the overall financial position of the country is extremely weak. Both private and public sectors have been accumulating excessive levels of foreign debt, which international investors are not willing to finance at sustainable rates and hence increasing dramatically the probability of another bail-out.

**The sequence of the European Council and Euro Group statements in response to the crisis**

- **16 February 2010:** The Council focuses on the situation regarding government deficit and debt in Greece, adopting:
  - an opinion on an update by Greece of its stability programme, which sets out plans for reducing its government deficit below 3% of gross domestic product by 2012;
  - a decision giving notice to Greece to correct its excessive deficit by 2012, setting out budgetary consolidation measures according to a specific timetable, including deadlines for reporting on measures taken;

- a recommendation to Greece to bring its economic policies into line with the EU's broad economic policy guidelines.

(http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/112912.pdf)

- **2 May:** Eurozone finance ministers agreed upon a rescue package for Greece amounting to €110 billion: €80 billion in bilateral loans over three years and €30 billion coming from the International Monetary Fund.
- **9/10 May:** The Council and the member states decide on a comprehensive package of measures to preserve financial stability in Europe, including a European Financial Stabilization Mechanism, with a total volume of up to €500 billion from euro area countries and European institutions and the IMF commitment to provide funding up to EUR 250 billion.

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/114324.pdf

- **29 October 29:** The European Council endorses the report of the Task Force on economic governance. The report also sets out the guiding principles for a robust framework for crisis management and stronger institutions; this includes the involvement of the private sector in the crisis mechanism.

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/117496.pdf

- **28/29 October:** The European Council agrees on the need to set up a permanent crisis mechanism to safeguard the financial stability of the euro area as a whole. Eurogroup Ministers agree that the European Stability Mechanism (ESM) will be based on the European Financial Stability Facility, capable of providing financial assistance packages to euro area Member States under strict conditionality functioning according to the rules of the current EFSF. Two further elements are key here:
  - First, support will be available only on the basis of 'a rigorous debt sustainability analysis conducted by the European Commission and the IMF'.
  - Second, 'an ESM loan will enjoy preferred creditor status'.

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/118050.pdf

- **28 November: (Euro-group statement on the Irish rescue package)** Ministers unanimously agreed to grant financial assistance in response to the Irish authorities' request on 22 November 2010. Ministers concur with the Commission and the ECB that providing a loan to Ireland is warranted to safeguard financial stability in the euro area and the EU as a whole. The total size of the package is €85 billion, one-third of it coming from the IMF.

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ecofin/118051.pdf

- **16 December:** The Council agreed on the text of a limited amendment to the Treaty on the establishment of a future permanent mechanism to safeguard the financial stability of the euro area. This amendment should enter into force on 1 January 2013. Heads of state reiterated their commitment to reach agreement on the legislative proposals on economic governance by the end of June 2011, with the aim of strengthening the economic pillar of the EMU.
- **24/25 March:** The Council endorses the features of the EMS decided by the euro area Heads of State or Government and takes necessary steps to ensure that the effective lending capacity of the EMS is of EUR440bn.

http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/120296.pdf

## The root causes of the crisis: leverage and bubbles

The euro zone crisis is certainly not the result of a single cause but the outcome of a combination of several factors and dynamics of macroeconomic, regulatory and institutional nature. These include irresponsible behaviour by several euro zone governments, the steady deterioration in peripheral EMU Member States of macroeconomic fundamentals to levels inconsistent with long-term EMU participation, failures in financial market regulation at global level, shift in markets' expectations induced by the financial crisis of 2007–08 and finally, also, defects in the institutional organization of the European Monetary Union.

All these factors are likely to have played a role in originating the crisis, but even together they are still insufficient to account for its systemic nature. This feature can only emerge from the vulnerability of the highly integrated European financial system. Had the Greek and Irish crises occurred when euro zone banks were strong and/or not very interconnected, the euro zone crisis would not have happened. But the European financial system was (and still is) fragile because of the high level of leverage accumulated over the credit boom.[3]

Excessive leverage is an essential ingredient in any major financial crisis and this case is no exception. In financial markets, leverage is defined as the ratio of debt to equity financing; when this ratio increases in general the capacity of a firm to absorb losses declines and hence its fragility is boosted. In macroeconomic terms, leverage is better defined as the ratio of debt to GDP and the concept can be applied to all the sectors of the economy. Leverage defined this way increases when credit expands without a consistent adjustment in GDP. Since regular cash flows are proportional to GDP, this implies that many agents have issued promises to pay a certain nominal amount but do not necessarily have the 'expected' regular cash flow to honour these promises (see Minsky (2008) for the classical description of leverage schemes leading systems towards instability). It is not possible to establish an absolute benchmark for leverage, as

[3]We leave aside the question of why the build-up of the credit boom was ignored. Inflation targeting by central banks was probably one key reason. According to Borio and Lowe (2002), a low-inflation environment increases the likelihood that excess demand pressures show up in the form of credit growth and asset prices bubble rather than in goods price inflation. If this is the case, inflation-targeting central banks with a 'myopic behaviour' could contribute to financial instability (Grauwe, 2009; de Grauwe and Gros, 2009).

different financial systems can support quite different ratios of credit to GDP. However, rapid and persistent increases in this ratio constitute alarm signals which have been identified as reliable predictors of financial crisis. These signals were clearly blinking before 2007, but they were ignored. Table 1 shows that over the last decade euro zone private debt relative to GDP increased by about 100 percentage points, more than it did in the USA. In addition, and unlike the USA, the increase took place in the financial system, whose fragility became apparent first in 2008 and then again in May 2010.
The question is why and how this could actually have happened.
Excess leverage in the banking sector was probably encouraged by scant financial regulation, but it would be too easy to blame car accidents for the absence of speed limits (despite speed limits helping to reduce accidents) or police control. The main driver of growing leverage was of an economic nature and tightly linked to large capital flows flying from core euro zone countries into the periphery after the creation of the euro. The peripheral euro zone economies (Greece, Ireland and Spain) in their catching-up phase appeared to core European Member States with large savings and little domestic investment prospects as a great investment opportunity: they seemed to offer the opportunities of emerging economies, but without the exchange rate risk.

The capital inflows generated their own fundamentals: high growth rates driven by strong demand for consumption and construction investment, supported by easy credit fed from abroad. In all this the financial system, banks in particular, played a crucial role. They made the capital flows possible and magnified the availability of credit through leverage by generating a tight network of intra-sector exposures.

Table 2 shows the level of leverage and the break down by sector in the euro zone countries embedding the most extreme conditions. Data suggest that while leverage barely changed in Germany over the prior decade, in the peripheral euro zone countries, and in particular in Spain and Ireland, the increase was dramatic.

However, it turned out that growth was unsustainable because it was driven by a bubble, and when the bubble burst, banks, not only in the periphery but also in core countries, who were at the origin of the credit flows, found themselves weak (because of high leverage) and very exposed to large potential losses.

**Table 1** *Leverage: euro zone versus USA (*source: *Federal Reserve, Flow of Funds Z1 (outstanding debt), Eurostat and authors' calculations).*

| | Non-financial corporations | Financial corporations | General government | Households |
|---|---|---|---|---|
| **Euro area** | | | | |
| **1999** | 67 | 66 | 74 | 49 |
| **2007** | 93 | 111 | 69 | 62 |
| **2010** | 102 | 127 | 87 | 65 |
| **US** | | | | |
| **1999** | 63 | 76 | 51 | 67 |
| **2007** | 74 | 113 | 51 | 96 |
| **2010** | 75 | 101 | 76 | 92 |

*Note:* For the euro area debt is computed as sum of loans and securities other than shares, excluding financial derivatives (only loans in the case of HH). This definition broadly corresponds to the definition of the outstanding debt used in the US flow of funds.

**Table 2** *Leverage for euro zone selected countries and sector break-down (*source*: Eurostat and authors' calculations).*

| Debt-to-GDP | Financial corporations | | Non-financial sector | | Households and non-financial corporations | |
|---|---|---|---|---|---|---|
| | 2000 | 2007 | 2000 | 2007 | 2000 | 2007 |
| **Greece** | 132 | 162 | 175 | 219 | 55 | 105 |
| **Ireland*** | 450 | 1142 | 181 | 294 | 151 | 210 |
| **Spain** | 164 | 310 | 187 | 255 | 122 | 214 |
| **Germany** | 273 | 293 | 200 | 196 | 139 | 130 |

*Note:* Debt is computed as sum of loans and securities other than shares, excluding financial derivatives, only loans in the case of households and including also deposits in the case of financial corporations. Non financial sector includes households, non financial corporations and government.
*Data for 2000 are not available, those shown refer to 2001.

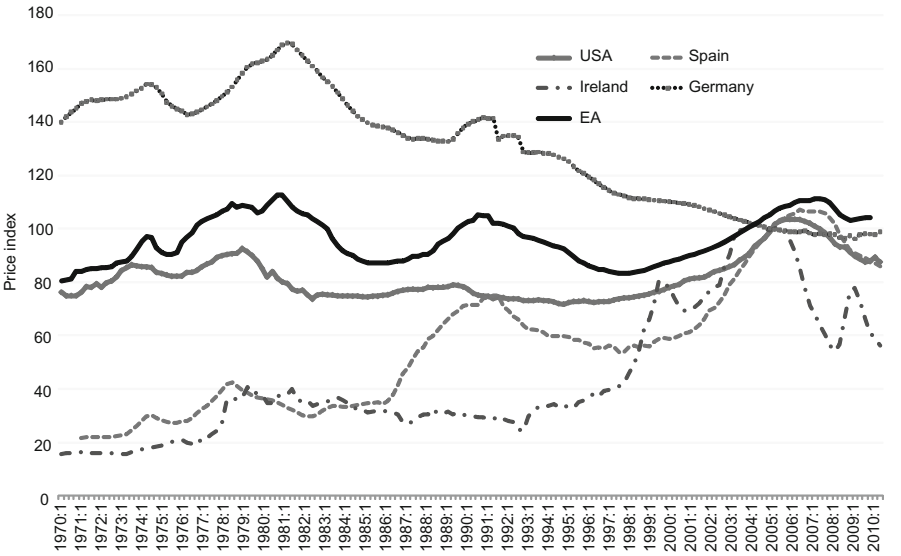



**Figure 1** House prices: price-to-rent ratios. *Source*: OECD, March 2011, and author's computations. *Note*: Euro area index is defined as the weighted average (by GDP) of Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands and Spain.

The magnitude of the losses was, and still is, potentially very large because some euro zone member countries (notably Ireland and Spain) experienced a real estate price bubble of the magnitude of the USA. Figure 1 provides evidence of this by showing the house price-to-rent ratios. This ratio, similar to the price/earnings ratio for stocks, should be stable over long periods. From the chart it is apparent that since the mid-1990s house prices have increased by almost exactly the same relative amount, reaching an unprecedented level on both sides of the Atlantic. The main difference between the USA and the euro area is that since 2006–07 house prices have declined more in the USA than in the euro zone.

As shown in the Figure 1, the euro area average hides important differences between countries: Between 1995 and 2006, while house prices have been declining or stable in Germany, they increased by over 80% and more than 140% (more than in the USA) in Spain and Ireland respectively. Furthermore, as shown in Figure 2, in these two countries the average investment in construction relative to GDP reached 18% and 21% of GDP respectively against an EU average of about 11% (see Figure 2). This seems to suggest that those countries are destined to suffer for years the consequences of housing and debt overhangs, and dealing with the legacy of national real estate bubbles and busts will remain a challenge for monetary union for some time to come.

This argument is of course related to the so-called 'Walters critique', which holds that a monetary union can amplify shocks because in a country subject to an inflationary pressures the real interest rate will be lower than in the rest of the union. This will fuel domestic demand, which in turn drives inflation even higher, thus lowering real rates even further. This feedback loop is self amplifying and could even be explosive.

However, it seems that in reality the importance of lower real interest rates, defined as nominal interest rates deflated by consumer price inflation, has been overrated. In the case of Spain, consumer price inflation was about 1.6% higher than in Germany over the first 8 years of the euro, but mortgage interest rates were actually over 1% point lower than Germany because they were indexed on short term rates and, even more importantly, house price inflation was 10% points higher than in Germany. This suggests that difference in the characteristics of national financial markets (e.g. the availability of mortgages indexed on short-term rates, different loan-to-value ratios etc.) meant that the easing of financial conditions after the creation of the euro had quite

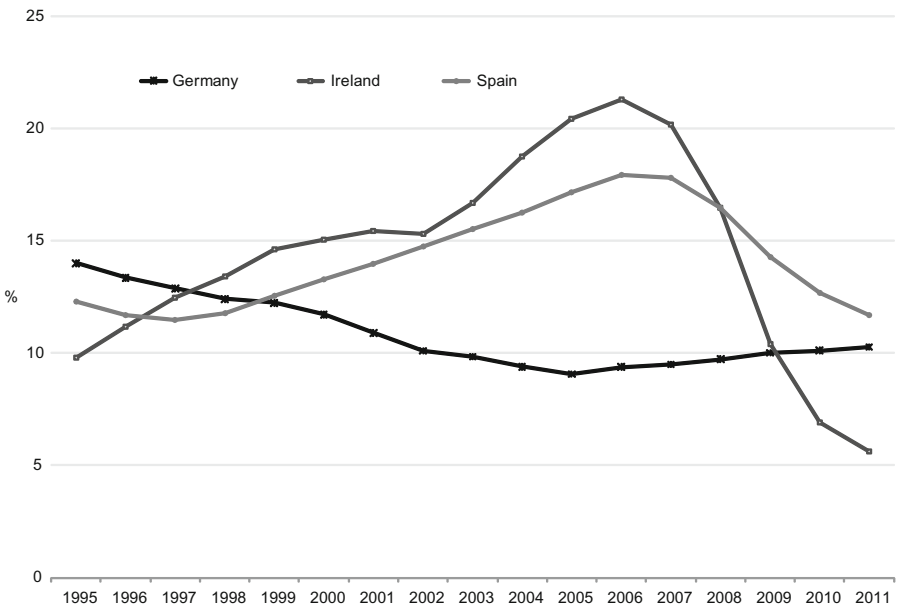



**Figure 2** Investment in construction % of GDP. *Source*: European Commission Services (Ameco database, Gross fixed capital formation at current prices: construction).

differentiated impacts on different member countries (Gros, 2009; Baldwin et al., 2010; Calza et al., 2009) with the housing markets playing a key transmission mechanism in Spain and Ireland.

DANIEL GROS AND CINZIA ALCIDI

*See Also* **banking crises; euro; European Central Bank; European Monetary Union; sovereign debt.**

### *Bibliography*

Alcidi, C. and Gros, D. 2009. Why Europe will suffer more. *Intereconomics* July/August.

Agnello, L. and Schuknecht, L. 2009. Booms and busts in housing markets: determinants and implications. *ECB Working Paper No. 1071*, European Central Bank, Frankfurt, July.

Baldwin, R., Gros, D. and Laeven, L. (eds.) 2010. *Completing the Euro Zone Rescue: What More Needs to be Done?* e-book, VoxEU, 17 June.

Blundell-Wignall, A. and Slovik, P. 2010. The EU stress test and sovereign debt exposures. *OECD Working Paper on Finance, Insurance and Private Pensions 4*, OECD Financial Affairs Division.

Borio, C. and Lowe, P. 2002. Asset prices, financial and monetary stability: exploring the nexus. *Working Paper 114*, Bank for International Settlements.

Calza, A., Monacelli, T. and Stracca, L. 2009. Housing Finance and Monetary Policy. *ECB Working Paper 1069*.

De Grauwe, P. 2009. Keynes' savings paradox, Fisher's debt deflation and the banking crisis, *CEPS Working Document No. 319*.

De Grauwe, P. and Gros, D. 2009. A new two-pillar strategy for the ECB. *CEPS Policy Brief No. 191*.

Eichengreen, B. 2010a. The euro: love it or leave it? *VoxEU*, 4 May. http://www.voxeu.org/index.php?q=node/729.

Eichengreen, B. 2010b. Ireland's rescue package: Disaster for Ireland, bad omen for the Eurozone. *VoxEU*, 3 December. http://www.voxeu.org/index.php?q=node/5887.

Gros, D. 2009. Comments on Charles Wyplosz, "Ten Years of EMU: Successes and Puzzles". In *Spain and the Euro, the First 10 Years* (ed. J. F. Jimeno). Banco de España.

Gros, D. 2010. Adjustment difficulties in the Gipsy Club. *CEPS working document No. 326*, Brussels.

Minsky, H. 2008. S*tabilizing an Unstable Economy*, McGraw-Hill, London.

Veron, N. 2010. The EU has not yet solved its banking problem. *VoxEU*, 26 October. http://www.voxeu.org/index.php?q=node/5714.

# Federal Reserve System

The Federal Reserve System of the United States was established on 23 December 1913, when President Woodrow Wilson signed the Federal Reserve Act. The need for a new federal banking institution became clear when a severe crisis occurred in 1907. In May 1908 the Aldrich–Vreeland Act established a bipartisan National Monetary Commission that proposed establishing a National Reserve Association with 15 locally controlled branches that would 'provide an elastic note issue based on gold and commercial paper' (Warburg, 1930, p. 59). The proposal was not enacted, nor was a subsequent proposal for a central bank with about 20 branches that would be controlled by a centralized Federal Reserve Board, consisting largely of commercial bankers. In the debate preceding the Federal Reserve Act, banking industry domination was rejected in favour of a board that had five members appointed by the President and two ex officio members, the Secretary of the Treasury and the Comptroller of the Currency. The appointed members had staggered terms and were to represent different commercial, industrial, and geographic constituencies. A sixth appointed member representing agriculture was added in 1923. The composition of the Board and its relation to Federal Reserve banks were drastically changed in 1935. Partly because of continuing disagreements about public versus commercial bank control, the new Board's powers were left ambiguous in the act.

The act mandated that all national banks become members of the new system and stockholders of Federal Reserve banks. Because reserves were to be concentrated in 12 Federal Reserve banks, the act substantially reduced reserve requirements at national banks. State chartered banks could join if they chose to and were judged to be financially strong. The first Board was sworn in on 10 August 1914 and the system opened for business on 16 November 1914. Federal Reserve notes that were backed 100 per cent by 'eligible paper' and, additionally, 40 per cent by gold began to circulate. Eligible paper was self-liquidating, short-term paper that arose in commerce and industry. The rationalization for eligible paper was the real bills doctrine, which held that credit extended for financing only the production and distribution of goods would not lead to inflation. The doctrine is invalid because of fungibility; there is no relation between paper acquired by Federal Reserve banks and loans the commercial banks are extending. In addition, all deposits at Federal Reserve banks had to be backed at least 35 per cent by gold. Subsequent amendments to the act effectively eliminated the supra-100 per cent collateralization of notes. A June 1917 amendment to the act forced all member banks to pool required reserves at Federal Reserve banks and further reduced reserve requirements to decrease the burden of membership on national banks and attract more state-chartered banks to the system.

### The early years

The early years of the Federal Reserve System were marked by struggles to define the distribution of power between Federal Reserve banks and the Board, in the context of

growing US involvement in the First World War. The Board gradually assumed more powers, but was unsuccessful in controlling open-market trading, which inevitably was concentrated in New York. Benjamin Strong, the New York bank governor, managed system trading. (Until 1935 the chief executives of Federal Reserve banks were called 'governors'. After 1935 their title was changed to 'president' and members of the Board were called 'governors'.) The Federal Reserve System was made fiscal agent for the Treasury in 1920, but the Treasury dealt directly with Federal Reserve banks, not the Board. Until 1922 the Board's statistical research office was located in New York, and arguably the Board was less informed than the New York bank about money market conditions.

Federal Reserve banks immediately sought earning assets in order to pay expenses and the six per cent required dividends on member bank capital subscriptions. As they expanded their portfolios of bills, US securities, discounted commercial paper, and acceptances, the breadth and liquidity of these markets increased. In early 1915 the New York bank was buying and selling for other Federal Reserve banks. Discount rates charged by reserve banks varied across Federal Reserve districts.

In anticipation of the US declaration of war on Germany in 1917, Federal Reserve banks became responsible for issuing and redeeming short-term Treasury debt certificates before and during Liberty Loan drives. There would be four large Liberty Loans and a Victory Loan in 1919 that required extensive Federal Reserve involvement. US bonds were sold to the public on an instalment plan by member banks; the interest rate banks charged on the unpaid balance on a bond was equal to the coupon rate on the bond. Member banks, in turn, discounted short-term US debt at Federal Reserve banks at an interest rate below the yield on the debt, which allowed them to recover their costs of instalment lending.

US government interest-bearing debt rose from $1.0 billion at the end of 1916 to $25.5 billion at the end of 1919, and would never again fall below $15 billion. This huge increase, and the fact that Federal Reserve banks offered preferentially low interest rates when member banks discounted government debt, had important lasting consequences on the money market. Before the war, Federal Reserve banks had schedules of discount rates that varied across the quality and maturity of discounted paper and the amount of borrowing by a member bank. Because of the low discount rate on government debt, member banks almost exclusively offered it as collateral when borrowing. The discount rate effectively became the rate charged on government debt. By 1922 each reserve bank effectively had a single discount rate, but rates still varied across Federal Reserve districts.

The November 1918 armistice brought new challenges. Continuing shortages of food and other goods in Europe and large increases in the stock of money led to inflation in the United States. The rate of inflation peaked in May 1920 and was followed by a sharp deflation in the following year of about 45 per cent in wholesale prices. In that year industrial production fell by about 30 per cent and unemployment soared. Until October 1919 Federal Reserve banks were obliged to keep the low wartime discount rates in order to allow banks and the public to absorb the 1919 Victory Loan. In November, Federal Reserve banks began raising their discount rates in an effort to combat inflation. In June 1920 four banks raised the rate to seven per cent. Amplifying

the effects of the interest rate increases was an outflow of gold to Europe and a sharp reduction in discount window borrowing as Federal Reserve banks cut back on subsidizing the public's instalment purchases of US bonds.

The Boston bank lowered its rate from seven per cent to six per cent in April 1921, and was gradually followed by other reserve banks in an effort to respond to the slowdown. Deposits at all member banks reached a local maximum of $26.1 billion in the December 1919 call report and then fell to $22.8 billion in the April 1921 report. Discount window borrowings reached a year end high of $2.7 billion in December 1920 and then fell to $0.6 billion at the end of 1922 as gold flows turned positive. As gold flowed in, reserve banks lowered their discount rates to 4.5 per cent in 1923 and early 1924.

While gold inflows slackened after 1923, it became apparent that new operating guidelines were needed. Governor Strong understood that the real bills doctrine was invalid and that many countries were not acting according to the old gold-standard rules. As interest rates fell, most reserve banks were again acquiring securities to augment their income. Strong, on the other hand, had begun to sterilize the New York bank's holdings of gold by selling its securities in the open market. The Treasury was concerned that reserve bank trading was upsetting securities markets when it was buying or selling debt. In May 1922 the reserve banks established the Governors Executive Committee consisting of the governors of the Boston, Chicago, Cleveland, New York, and Philadelphia banks to manage transactions for all 12 banks. The committee executed orders on behalf of the banks in the light of Treasury plans and made recommendations, but acted only as agents and had no executive power. In April 1923 it was renamed the Open Market Investment Committee (OMIC), which had the same membership as its predecessor but was required

> to come under the general supervision of the Federal Reserve Board; and that it be the duty of this committee to devise and recommend plans for the purchase, sale and distribution of open-market purchases of the Federal Reserve Banks in accordance with ... principles and such regulations as may from time to time be laid down by the Federal Reserve Board. (Chandler, 1958: 227–8)

Strong dominated the OMIC and began to understand the way open-market operations worked. He noted in particular that the sum of reserve bank open-market purchases and gold inflows almost equalled negative changes in member bank borrowing. He developed a case for active monetary policy and argued that restrictive monetary policy should be initiated with open-market sales and followed by increases in the discount rate. This was the likely origin of member bank borrowings and nominal interest rates as indicators of monetary policy. Policy instruments were open-market operations and the discount rate. While proposals to change discount rates originated with Federal Reserve banks, they required Board approval, which may explain why Strong preferred to lead with open-market operations. Strong was sensitive to the effects of monetary policy on prices, but objected to any legislated targeting of prices. His analysis was seriously incomplete when banks were not net borrowers from the Federal Reserve, and in such circumstances so were his policy tactics. Tragically, beginning in 1916 Strong suffered from recurrent attacks of tuberculosis and would die in October 1928, before such circumstances arose.

The 1923 Board Annual Report advocated an activist policy, but continued to support the real bills doctrine. In response to pressure from the Treasury and the Board, Federal Reserve banks sold most of their government securities in 1923; yearend holdings fell from $436 million to $134 million between 1922 and 1923. Federal Reserve notes and member bank reserves backed by such assets were unjustifiable under the doctrine, and the Treasury objected to Federal Reserve banks profiting from such assets. However, at the end of 1924 the banks held $540 million, and the banks' portfolio of government securities fluctuated considerably in the following years in response to changes in the volume of discounted bills and gold flows. Discount rates at Federal Reserve banks were lowered in the latter half of 1924 and 1925 before converging on four per cent at the beginning of 1926, largely following short-term interest rates in New York. Short-term market rates fell because of a sharp recession; the Federal Reserve index of industrial production (1997 = 100) fell from 7.84 in May 1923 to 6.43 in July 1924. Clearly policy was active, but not because of the real bills doctrine!

The discount rate was four per cent in June, when Federal Reserve banks began to cut the rate to 3.5 per cent and to make open-market purchases. At the beginning of 1928 discount rates were increased because of developing speculation in the stock market and continued to rise to as much as six per cent in October 1929, when the stock market crashed. In part, Federal Reserve discount rates were again responding to changes in industrial production, which had been quite sluggish until the end of 1927 and then began to grow rapidly until July 1929. In part, the 1927 rate cut reflected Federal Reserve efforts to help the United Kingdom maintain sales of gold at the pre-war sterling price, which had been restored in 1925. Governor Strong and Montagu Norman, the Governor of the Bank of England, were working to re-establish a gold standard that could restore order to international finance. To help the United Kingdom in 1925, the New York bank extended the Bank of England a $200 million gold credit and attempted to keep interest rates low in New York relative to those in London. By reopening gold sales at the pre-war price, Britain had effectively revalued the pound upward in 1925 by about ten per cent, with devastating consequences for its economy.

As Strong's health failed in 1928, a leadership vacuum developed. In an attempt to coordinate policy among all 12 reserve banks and the Board, the Board proposed in August 1928 that the five member OMIC be replaced by a new Open Market Policy Committee (OMPC) that included all 12 reserve bank governors and was chaired by the Governor of the Federal Reserve Board. This proposal was rejected by bank governors, but a modified form was adopted in January 1930. Strong had been aware of growing stock market speculation and did not object to Federal Reserve open-market sales and the increase in the discount rate. These actions were reinforced by outflows of gold. In mid-1928 gold flows reversed, apparently attracted by high and rising short-term interest rates. Federal Reserve banks continued to sell bills and government debt, forcing member banks into the discount window to the extent of about $1 billion in the second half of 1928 and in the middle of 1929. At the end, Strong was aware of the danger of restrictive monetary policy actions over an extended period on the real economy, but remained reasonably optimistic that the situation could be controlled (Chandler, 1958: 460–3). After his death the struggle for control continued between his successor at the New York bank, George L. Harrison, and the Board; the latter argued

that the real bills doctrine was not dead and that reserve banks should take direct action to penalize member banks making loans that supported security speculation. The Federal Reserve index of industrial production peaked in July 1929, Bureau of Labor Statistics (BLS) wholesale and consumer price indices had been slowly falling since 1926, and in October the stock market collapsed.

## The Great Depression

Led by the New York bank, the Federal Reserve flooded the money market with cash by aggressively buying government securities. Discount window borrowing by member banks fell from $1,037 million in June 1929 to $632 million in December and to $271 million in June 1930. Further, discount rates at reserve banks were rapidly reduced; at the New York bank the rate was lowered from six per cent in October to 2.5 per cent in June 1930. The monthly average Standard and Poor common stock index (1935–1939 = 100) began to stabilize; it was 195.6 in January 1929, 237.8 in September, 159.6 in November, and 191.1 in April 1930. However, the index of industrial production continued to fall after the open-market purchases, and the BLS index of wholesale prices was ten per cent lower in 1930 than in 1929.

In mid-1930 reserve banks sharply reduced their purchases of government securities in the belief that monetary policy was adequately expansionary. The OMPC seems to have been guided by what Meltzer (2003: 164) calls the Riefler–Burgess Doctrine: 'If [discount window] borrowing and interest rates were low, policy was easy; if the two were high policy was tight.' An interpretation is that if member banks wanted to lend they could have inexpensive and relatively easy access to funds; if not, there was little more that the Federal Reserve could do. While total member bank discount window borrowing was positive, many banks were holding excess reserves. Conventional wisdom has it that the reserve banks should have continued buying securities. However, it is unclear even today whether continued large open-market purchases by the Federal Reserve would have had much of an impact on real economic activity in late 1930; the experiment was never tried. Rapid expansion of reserves and member bank deposits did occur in the late 1930s, with little effect on real economic activity.

On average about 600 bank failures a year occurred between 1920 and 1930; most failing banks were small and not members of the Federal Reserve System. The number of failing banks doubled in 1930 and increased by another 70 per cent in 1931. The total deposits of failing banks between 1920 and 1930 averaged less than $200 million a year, but more than quadrupled in 1930 and doubled again in 1931. Total deposits and currency had begun to fall after December 1928 and continued to fall after the stock market crash. Currency in circulation began to rise in November 1930, as bank failures increased. Industrial production and wholesale prices were falling at an accelerating rate. The directors of the New York bank counselled Governor Harrison to continue open-market purchases in 1930, but he encountered opposition in the OMPC and little was done. Net gold inflows were offset by open-market sales because the OMPC collectively believed monetary policy was expansionary. Reserve bank discount rates and money market interest rates trended down until 21 September 1931, when the United Kingdom suspended gold payments.

The British abandonment of gold led to very large withdrawals of gold and currency from the United States that were initially partially offset by open-market purchases of bills and increased discount window borrowing, which occurred at sharply higher interest rates as recommended by Bagehot (1873). However, Federal Reserve bank credit fell from $2.2 billion in October 1931 to $1.6 billion in March 1932. During this period of rising bank failures, rapidly declining economic activity, and falling prices, Harrison argued against open-market purchases for a number of reasons, but primarily because of the possibility of a shortage of 'free gold', that is, gold that was not required as collateral for Federal Reserve notes and reserves. The Glass–Steagall Act of 1932 authorized the Federal Reserve banks temporarily to use US government securities as collateral for Federal Reserve notes and thus largely solved the problem of a lack of free gold. In February 1932 Federal Reserve banks began aggressive open-market purchases of government securities that more than offset continuing gold losses and allowed member bank borrowings to fall about 50 per cent by August 1932. Discount rates at the New York and Chicago banks were lowered to 2.5 per cent in June 1932, but all other banks kept their rates at 3.5 per cent until the national banking 'holiday' that began on 5 March 1933 when President Roosevelt closed all US banks. Net free reserves (excess reserves minus discount window borrowing) had turned positive in September and thus signalled excessive ease to some individuals on the OMPC.

## Restructuring the Federal Reserve System

It was obvious that the Federal Reserve had been ineffective in combating the collapse of the banking system and responding to the Great Depression. The banking system and the Federal Reserve needed to be restructured and strengthened. The Emergency Banking Act of 9 March 1933 authorized the Treasury to license and reopen national banks that were judged to be sound; state chartered banks that were sound would receive licences from state banking commissioners. Many reopening banks received capital injections by selling preferred stock to the Reconstruction Finance Corporation. At year end 1929 there were 24,026 commercial banks of which 8,522 were members of the Federal Reserve System; at year end 1933 there were 14,440 commercial banks of which 6,011 were member banks. For a period of one year all banks, whether members or not, could borrow on acceptable collateral from Federal Reserve banks.

Many of the reforms that were adopted would survive at least until late in the 20th century. Because of a belief that the collapse lay in undisciplined stock market trading, the Glass–Steagall Act of 1933 required that commercial banks divest themselves of investment banking activities. This act introduced deposit insurance that became effective in January 1934. It also banned interest payments on demand deposits and allowed the Board to impose ceilings on interest rates that banks could pay on time and savings deposits. Finally, the act renamed the OMPC the 'Federal Open Market Committee' (FOMC), but as in earlier incarnations its executive committee remained the same. The Securities Exchange Act of 1934 authorized the Board to impose margin requirements on stock market trades. Federal Reserve banks were authorized to make commercial and industrial loans to non-financial firms.

Having failed to expand reserve bank credit between July 1932 and February 1933, the Board found itself under extraordinary political pressure to expand resources to the banking system. As Meltzer (2003: 435–41) explains, President Roosevelt threatened to have the Treasury issue currency in the form of greenbacks if the FOMC failed to expand sufficiently. Net free reserves turned positive in May 1933 and rose to more than $3.0 billion by January 1936. The revaluation of gold in February 1934 together with subsequent large gold inflows from Europe and hesitancy to lend by member banks contributed to this surge in excess reserves.

The reconstruction of the Federal Reserve System continued with Roosevelt's nomination of Marriner Eccles to become Governor of the Federal Reserve Board in November 1934. Eccles had argued that system power should be concentrated in the Board and that reserve banks be prevented from undertaking open-market operations on their own accounts. Eccles's initiatives were opposed by Senator Carter Glass, many reserve bank governors, and the banking industry, but he largely succeeded in achieving his goals. The reforms were in the Banking Act of 1935, which restructured the Board to consist of seven appointed governors, each with a staggered 14-year term. The FOMC was restructured to consist of the seven governors and five reserve bank presidents. Two of the governors were to be appointed for four-year terms as chairman and vice-chairman of the Board by the president, with the advice and consent of the Senate. Eligible paper was no longer restricted to being short-term paper that originated in commerce and industry. The Board was empowered to vary reserve requirements; the upper limit was twice the percentages that were specified in the 1917 amendments to the Federal Reserve Act.

Members of the renamed Board of Governors of the Federal Reserve System took office in February 1936, with Eccles as chairman. For some time the FOMC had expressed concern about the inflationary potential of large excess reserves. In particular, because excess reserves exceeded reserve bank credit, the FOMC would not be able to absorb them without an increase in reserve requirements. Employing its new policy instrument, on 14 July 1936 the Board announced an increase in reserve requirements on August 15 of 50 per cent on all deposits at member banks. The increase was expected to absorb less than half of system excess reserves and was not expected to impinge on member bank lending or the economic recovery. In part because of continuing gold inflows, excess reserves were $3.0 billion at the end of July 1936, and averaged about $2.0 billion through the end of February 1937. Because excess reserves continued to be large, the Treasury began to sterilize gold inflows in December 1936, but not to the extent desired by the Board. At the end of January the Board announced a further two-step increase in reserve requirements of one-third to take place in March and May 1937. These actions took reserve requirements to their legal maxima and reduced excess reserves to below $800 million in summer months. In August and September reserve banks reduced their discount rates to one per cent or 1.5 per cent, levels that would last until December 1941. Coinciding with the May increase, the industrial production index (1997 = 100) reached a high of 10.4 and then decreased to 7.0 in May 1938. Continuing gold inflows and the Treasury's February 1938 abandonment of gold sterilization allowed excess reserves to increase to $1.5 billion in March 1938. Beginning after the Board's reduction in reserve requirements of more than ten

per cent in April 1938, excess reserves began a rise to nearly $7 billion in late 1940; however, industrial production did not pass its 1937 peak until October 1939, after the Second World War had begun in Europe.

## Second World War and recovery

As the war approached gold flowed into the United States, and the FOMC allowed its security holdings to fall and their maturity to lengthen. In response to inflationary pressures, the Board introduced consumer credit controls in September 1941 and again raised reserve requirements to their legal maxima in November. After the United States declared war, monetary policy was constrained to facilitate war finance. In April 1942 the FOMC set interest rate ceilings on treasury bills at 0.375 per cent and on long-term bonds at 2.5 per cent. The yield curve was upward-sloping and effectively 'pegged' by these two boundary conditions into the post-war period. Because capital gains could be earned by buying high coupon securities and selling as they approached maturity, the cost of intermediate term debt was higher than rates shown on the yield curve. Discount rates were lowered to one per cent by all reserve banks and were not raised again until 1948. A preferential discount rate of 0.5 per cent was charged for loans collateralized by short-term US debt. Reserve requirements for central reserve city member banks were lowered in 1942, causing interest-free reserves to disappear into interest-bearing US securities. Finally, a variety of selective credit controls were imposed during and after the war, which ended in August 1945.

Yearend deposits and government securities of member banks had risen from $61.7 billion and $19.5 billion in 1941 to $129.7 billion and $78.3 billion respectively in 1945. Because of the pegging of the yield curve, Federal Reserve bank yearend ownership of US securities rose from $2.3 billion in 1941 to $24.3 billion in 1945; treasury bills were $10 million in 1941 and $14.4 billion in 1946.

The preferential discount rate was eliminated in the spring of 1946. In July 1947 the FOMC relaxed the rate ceiling on treasury bills and the rate rose to about one per cent by yearend. Reserve banks raised the discount rate to 1.25 per cent in early 1948. Eccles's long term as chairman ended in February 1948, but he continued as a member of the Board. Reserve requirements were increased in 1948 as the Board sought to control inflation, although prices were actually falling at yearend when a recession occurred. Indeed, the reserve requirement policy instrument was used many times between April 1948 and February 1951 because it was perceived not to have a direct effect on treasury interest rates. A continuing struggle between the Board and the Treasury for an independent monetary policy would not be resolved until a spurt of inflation after the start of the Korean War led to an accord signed on 4 March 1951. It effectively freed the Board from pegging interest rates. Partly because of frictions leading to the accord, a new chairman, William McChesney Martin, Jr., was appointed in April.

## Resumption of discretionary monetary policy

In the Martin era of discretionary monetary policy, new operating techniques were needed. In 1953 the FOMC settled on a policy of 'bills only', which meant that

open-market operations would be largely confined to the market for treasury bills, because it was recognized that large policy actions in thin markets could impair market efficiency. Indicators of monetary policy continued to be net free reserves and market interest rates. Because evidence was lacking that interest rates had much effect on private sector investment, a new paradigm, the 'availability of credit' doctrine, was used to rationalize the transmission of policy actions to the real economy. It argued that banks rationed credit to marginal borrowers when restrictive policy led to rising interest rates or indebtedness at the discount window. With these adjustments the FOMC vigorously and unsuccessfully pursued goals of lowering inflation and combating unemployment in the turbulent decade of the 1950s. In that decade there were three business cycles, which were marked by successively rising peaks of interest rates, inflation, and unemployment. The reason for this failure was thought to be inflation-induced rising marginal rates of taxation, which were addressed by large tax cuts in the following decade.

As interest rates rose, the opportunity cost of holding excess reserves rose, which led to the reappearance of a federal funds market in which banks traded reserves. Because banks paid no interest on demand deposits, there was also rapid expansion of the market for commercial paper in which large firms with good credit ratings traded idle funds without the direct intervention of banks. Both markets had atrophied after the 1920s because of low interest rates, and served to change the relation between open-market operations and real economic activity. They were precursors of a wave of innovations that would have similar effects in the coming decade. These included large-denomination negotiable certificates of deposit, one-bank holding companies, offshore 'shell' branches, the Eurodollar market, and bank-related commercial paper.

Beginning in 1961, the Kennedy administration attempted to coordinate fiscal and monetary policy by proposing large tax cuts to encourage investment and economic expansion. A new problem was that the United States was experiencing large gold outflows as the world continued to recover from the world war. To cope with this new approach and problem, the FOMC was encouraged to abandon its bills-only policy and to attempt to twist the yield curve by buying long-term bonds and selling bills. As short-term rates rose the Board repeatedly raised the ceiling on interest rates that banks could pay on time and savings deposits. It was argued that lower long-term interest rates would encourage capital formation and that higher short rates would discourage foreign interests from converting dollars into gold, as they were entitled to under the Bretton Woods agreements. These efforts were not successful in discouraging gold outflows, but investment and the economy expanded strongly. In 1965 the Board introduced a Voluntary Foreign Credit Restraint programme, which discouraged banks from overseas lending that was not financing US exports. Nevertheless, gold continued to flow out and the requirement that Federal Reserve notes and reserves be backed by gold was cancelled in 1968. Large open-market purchases had been needed to offset gold losses.

Policy coordination between the Board and the new Johnson administration effectively ended in December 1965, when the Board approved an increase in the discount rate because of inflation arising from mobilizing for the Vietnamese War. Net free reserves had turned negative in 1965 and were increasingly so until late 1966.

Short-term interest rates rose until October. Higher rates increased the cost of the mobilization and had devastating effects on residential construction and the savings and loan associations and mutual savings banks (hereafter thrifts) that financed it, because in September Congress passed legislation limiting interest rates that thrifts could pay on time and savings accounts. These limits meant thrifts would experience withdrawals of funds or 'disintermediation' because depositors switched funds to government securities, which had no limits. This policy transmission channel would soon disappear because Congress and the administration could not withstand the resulting political pressures. In 1968 the Federal National Mortgage Association was privatized and in 1970 the Federal Home Loan Mortgage Corporation was created. Both bypassed depository institutions by securitizing mortgage loans. Banks also responded to Board policies and restrictions on innovations by opening overseas offices that were not subject to them. A ten per cent income tax surcharge in 1967 was insufficient to stop inflation, and short-term interest rates rose to new highs in January 1970, when Chairman Martin's term ended. Net free reserves averaged about a negative $1 billion between May 1969 and July 1970. A decrease in short-term interest rates followed the then largest-ever US bankruptcy of the Penn Central Transportation Company in June 1970, but led to large new capital outflows in 1971 that pressured the dollar. The FOMC responded by forcing short-term rates and net borrowed reserves up again.

## Towards flexible exchange rates

The amplitude of changes in interest rates increased between 1965 and 1971, and the United States experienced a recession in 1970. As in the 1950s the Federal Reserve was unable simultaneously to achieve satisfactory unemployment, inflation, and exchange rate outcomes. Many of the Board's policy instruments, such as the discount rate, reserve requirement changes, and many regulations had effectively been disabled by innovations, so that only open-market operations were available to achieve multiple targets. For example, an increase in reserve requirements induced banks to resign from the system or to conduct more of their business overseas. One exception to this loss of powers was the 1970 amendments to the Bank Holding Company Act, which finally gave the Board regulatory authority over one-bank holding companies. In August 1971 the Nixon administration, with new Board Chairman Arthur F. Burns as an advisor, announced a 90-day freeze on prices and wages, suspension of gold sales, and several other major changes in the United States. The suspension of gold sales led to a floating exchange rate system, devaluation of the dollar, and sharp rises in dollar-denominated prices in international markets. The shift from a fixed to a floating exchange rate system is likely to have increased the potency of monetary policy, as was predicted by Mundell (1961). The FOMC responded to consequent high inflation by driving nominal short-term interest rates to very high levels in 1973 and 1974, which helped to induce a severe recession beginning in August 1973, but were inadequate because on average the real federal funds interest rate (calculated with the GDP deflator) was negative between the end of 1973 and 1978. Real estate and other durable goods prices rose relative to the GDP deflator, and the international value of the dollar fell. After the resignation of President Nixon in 1974, Congress required the Chairman to explain

policy in semi-annual public hearings and report the FOMC's targets for two money stock measures: M1, a measure of transactions balances, and M2, a measure of liquid assets. Friedman and Schwartz (1963) had recommended using money as an indicator of monetary policy instead of interest rates or net free reserves.

Part of the explanation for the policy failure was continuing financial market innovation. Foreign banks operating in the United States grew rapidly and were unregulated until the 1978 International Banking Act, which placed them under Board supervision. The introductions of money market mutual funds (MMMFs) and negotiable order of withdrawal (NOW) accounts in 1972, the Chicago Board Options Exchange in 1973, and financial futures markets in 1975 again began changing the relation between financial and real markets. A more important change was the rapid expansion of repurchase agreements after 1970. In a repurchase agreement, a client's deposits are borrowed to finance a bank's or dealer's inventory of government securities, often only overnight. Large bank holdings of government securities often represented transactions balances of large corporations and state governments that could not easily be controlled.

The real federal funds rate turned distinctly positive in the third quarter of 1979 when Paul A. Volcker became chairman. In early October he announced that the FOMC would no longer limit fluctuations in short-term interest rates and would use open-market operations to control bank reserves. This was a major policy change from practices dating from the 1951 accord. Further, he imposed eight per cent marginal reserve requirements on non-deposit liabilities, that is, Eurodollar borrowing, federal funds purchased from non-member banks, and funds acquired through repurchase agreements. These vigorous actions together with large income tax cuts by the Reagan administration between 1981 and 1983 drove real short-term interest rates to levels not seen since the early 1930s and caused MMMFs to grow rapidly. In only two quarters between 1979 and 1986 was the average real federal funds less than five per cent. These high rates caused the trade-weighted value of the US dollar to appreciate by 87 per cent between July 1980 and February 1985, which savaged US exports and attracted imports with adverse consequences for US manufacturing.

## Financial deregulation

The landmark Depository Institutions Deregulation and Monetary Control Act was signed by President Carter at the end of March 1980. It radically changed the Federal Reserve System by eliminating the significance of membership in the system. After an eight-year phase-in period, all depository institutions would be subject to uniform reserve requirements on demand and time deposits, although the requirement on the first $25 million of transactions deposits was less than that on other transactions deposits. The Board could vary reserve requirements. All depository institutions had access to reserve bank discount windows. This strengthened the system because banks could no longer threaten to leave it in order to get the lower requirements that many states imposed. Further, Federal Reserve banks were required to charge banks for the cost of services they provided. Before this act they had been giving away services as an inducement for banks to stay in the system. This pricing requirement in turn forced depository institutions to begin to charge their clients for services, which changed the way

banking services were used. The act mandated that interest rate ceilings on time and savings accounts be eliminated after six years, increased deposit insurance, and had other important provisions that are beyond the scope of this discussion.

In late 1980 the Board announced that transfers from overseas branches to the United States could be treated as collected funds on the day they were transferred. Before then, transfers in a day were not 'good funds' until the following day. The expansionary effects of this change, rapidly growing repurchase agreements, and other innovations are evident in demand deposit turnover statistics that the Board reported from 1919 until August 1996. Turnover is the annualized value of all withdrawals from deposit accounts divided by aggregate deposit balances.

High interest rates were savaging thrift institutions, which had negative gaps (more fixed-rate assets than fixed-rate liabilities on most future dates), and allowed MMMFs to expand rapidly. Congress intervened in September 1982 by passing the Garn–St Germain Act, which provided temporary emergency assistance and among other changes introduced money market deposit accounts and super NOW accounts, which paid market interest rates. MMMF growth was slowed by this act, but the weakening condition of banks and thrift institutions would result in large numbers of failures as the decade wore on. Large banks also experienced large losses because the appreciating dollar had resulted in failures of sovereign states, especially in Latin America, to meet their loan obligations. Chairman Volcker was heavily involved in negotiating solutions for these defaults.

The restrictive monetary policy resulted in the deepest recession since the Depression; the unemployment rate was 10.8 per cent at the end of 1982. At the end of Volcker's term in August 1987 the unemployment rate had fallen to six per cent and the consumer inflation rate was less than two per cent. Real interest rates had fallen from 10.5 per cent in mid-1981 to four per cent, and the trade-weighted value of the dollar fell correspondingly. Volcker's February 1987 statement of monetary policy objectives to the Congress reported that M1 was not a reliable indicator of monetary policy and would be de-emphasized.

While his successor, Alan Greenspan, inherited a much improved economy, many problems remained from a rising wave of bank failures and the collapse of thrift institutions. Real estate markets were especially disorderly when the thrift crisis was resolved beginning in 1989 and were further distorted by provisions in the Tax Reform Act of 1986, which disallowed many interest tax deductions. After 1990 interest on home loans was effectively the only deductible interest on individual income tax returns. In addition, a collapse of stock prices in October 1987, strong foreign demand for US currency associated with the collapse of the Soviet Union, and a recession at the end of 1990 presented further challenges. The FOMC responded to these challenges by varying the real federal funds rate, defined using the contemporaneous GDP price deflator inflation rate. This rate fell sharply for two quarters after the stock market crash, rose before falling for two quarters after a second stock market dip in October 1989, and then began to fall in the fourth quarter of 1990. In July 1993 testimony before Congress, Greenspan disclosed that the FOMC was downgrading M2 as an indicator of monetary policy and, as could have been surmised from its actions, that an important guidepost was now real interest rates. The real federal funds rate averaged

less than one per cent in 1993. In early 1995 it had risen to four per cent and held that value as an average until the collapse of a large hedge fund in September 1998. After the fallout from the hedge fund collapse had been resolved, the real federal funds rate was restored to an average of about four per cent in 2000. When a new recession appeared in 2001 together with a sustained large collapse in stock market prices, the real federal funds rate was lowered to near zero in the fourth quarter; the rate had averaged zero for 13 consecutive quarters as of March 2005.

Between December 1990 and April 1992 reserve requirements on time and demand deposits were reduced, which helped banks to increase net income. In January 1994 'retail sweep programmes' were introduced. In these programmes, a bank shifts funds from a depositor's transactions account to a synthetic time deposit account in the depositor's name in order avoid reserve requirements, usually without the depositor's knowledge. The Board does not measure the amount of funds swept, except at the time the programme was established. The Board estimated that as of August 1997 required reserves fell by one-third because of these programmes.

In November 1999 President Clinton signed the Financial Services Modernization (Gramm–Leach–Bliley) Act, which reversed the 1933 Glass–Steagall Act's ban on combining commercial and investment banking. The ban had been eroding since 1987, when some large bank holding companies were authorized by the Board to establish subsidiaries that could underwrite state and local government revenue bonds. The new act authorized the establishment of financial holding companies, which were to be regulated by the Board and could engage in an approved list of activities that included commercial banking, insurance, securities underwriting, merchant banking, and complementary financial undertakings. In 2003 there were more than 600 financial holding companies, which resemble the universal banks that exist in other countries.

In December 2002 the Federal Reserve discarded the discount rate as a policy instrument by replacing it with an interest rate on primary credit extended by the discount window that is one per cent above the FOMC target federal funds rate. Primary credits are collateralized loans to banks in sound financial condition.

As the foregoing dramatic institutional changes suggest, the Federal Reserve System is a work in progress. Its set of policy instruments and its dimensions have radically changed. Because of offshore banking facilities and retail sweep accounts, reserve requirement changes are no longer an effective policy instrument. As noted in the preceding paragraph, the discount rate has been discarded as an instrument; it is simply a penalty rate that is related to a bank rate, as is often the practice in other countries. Regulations on the interest rates banks pay on time and savings deposits have been discarded. Open-market operations are almost the sole policy instrument that can be used to achieve the Board's target nominal and real federal funds interest rates. While the FOMC has been able to control the overnight federal funds rate, the linkage between it and real economic activity is changing. First, the combined holdings of US government securities by foreign central banks have recently exceeded those of Federal Reserve banks. Foreign central bank holdings are partly a result of their efforts to manipulate exchange rates; their holdings are likely to change when FOMC policies change. Second, repurchase agreements and offshore transactions vary considerably over time and their volumes appear to be sensitive to US economic activity. Third, the

outstanding stock of securitized mortgage and other debt has been growing rapidly; such debt is a close substitute for US government debt and its amount has real economic effects. Fourth, because of decreasing required reserves and growing offshore holdings of US currency, 89 per cent of Federal Reserve liabilities were in the form of Federal Reserve notes in December 2003; the corresponding share was 34 per cent in 1941, 57 per cent in 1970, and 79 per cent in 1989. In part, the Federal Reserve recently has become an institution for collecting seigniorage from the rest of the world. Finally, over the decade ending in 2003, the share of all credit market assets held by depository institutions in the Federal Reserve's flow of funds accounts fell. In the context of the most recent 13 quarters of a zero real federal funds interest rate, more changes could be expected.

DONALD D. HESTER

*See Also* **Great Depression; monetary and fiscal policy overview.**

### *Bibliography*

Bagehot, W. 1873. *Lombard Street: A Description of the Money Market*. Homewood, IL: Richard D. Irwin, 1962.

Chandler, L. 1958. *Benjamin Strong: Central Banker*. Washington, DC: The Brookings Institution.

Chandler, L. 1970. *America's Greatest Depression 1929–1941*. New York: Harper and Row.

Crabbe, L. 1989. The international gold standard and US monetary policy from World War I to the New Deal. *Federal Reserve Bulletin* 75, 423–40.

Dykes, S. 1989. The establishment and evolution of the Federal Reserve Board: 1913–23. *Federal Reserve Bulletin* 75, 227–43.

Friedman, M. and Schwartz, A. 1963. *A Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.

Goldenweiser, E. 1951. *American Monetary Policy: A Research Study for the Committee for Economic Development*. New York: McGraw-Hill.

Greider, W. 1987. *Secrets of the Temple: How the Federal Reserve Runs the Country*. New York: Simon and Schuster.

Meltzer, A. 2003. *A History of theFederal Reserve, Volume I (1913–1951)*. Chicago: University of Chicago Press.

Mundell, R. 1961. Flexible exchange rates and employment policy. *Canadian Journal of Economics* 27, 509–17.

Scott, I., Jr. 1957. The availability of credit doctrine: theoretical underpinnings. *Review of Economic Studies* 25(4), 41–8.

Warburg, P. 1930. *The Federal Reserve System: Its Origin and Growth*, vol. 1. New York: Macmillan.

# gold standard

The classical gold standard (which ended in 1914) and the interwar gold standard are examined within the same framework, but their experiences are vastly different.

## Types of gold standard

All gold standards involve (*a*) a fixed gold content of the domestic monetary unit, and (*b*) the monetary authority both buying and selling gold at the mint price (the inverse of the gold content of the monetary unit), whereupon the mint price governs in the marketplace. A 'coin' standard has gold coin circulating as money. Privately owned bullion (gold in form other than domestic coin) is convertible into gold coin, at (approximately) the mint price, at the government mint or central bank. Private parties may melt domestic coin into bullion – the effect is as if coin were sold to the monetary authority for bullion. The authority could sell gold bars directly for coin, saving the cost of coining.

Under a pure coin standard, gold is the only money. Under a mixed standard, there are also notes issued by the government, central bank, or commercial banks, and possibly demand deposits. Government or central-bank notes (and central-bank deposit liabilities) are directly convertible into gold coin at the fixed price on demand. Commercial-bank notes and demand deposits are convertible into gold or into gold-convertible government or central-bank currency. Gold coin is always exchangeable for paper currency or deposits at the mint price. Two-way transactions again fix the currency price of gold at the mint price.

The coin standard, naturally 'domestic', becomes 'international' with freedom of international gold flows and of foreign-exchange transactions. Then the fixed mint prices of countries on the gold standard imply a fixed exchange rate (mint parity) between their currencies.

A 'bullion' standard is purely international. Gold coin is not money; the monetary authority buys or sells gold bars for its notes. Similarly, a 'gold-exchange' standard involves the monetary authority buying and selling not gold but rather gold-convertible foreign exchange (the currency of a country on a gold coin or bullion standard).

For countries on an international gold standard, costs of importing and exporting gold give rise to 'gold points', and therefore a 'gold-point spread', around the mint parity. If the exchange rate, number of units of domestic per unit of foreign currency, is greater (less) than the gold export (import) point, arbitrageurs sell (purchase) foreign currency at the exchange rate and also obtain (relinquish) foreign currency by exporting (importing) gold. The domestic-currency cost of the transaction per unit of foreign currency is the gold export (import) point; so the 'gold-point arbitrageurs' receive a profit proportional to the exchange-rate/gold-point divergence. However, the arbitrageurs' supply of (demand for) foreign currency returns the exchange rate to below (above) the gold export (import) point. Therefore perfect arbitrage would keep the exchange rate within the gold-point spread. What induces

gold-point arbitrage is the profit motive and *the credibility of the monetary-authorities' commitment* to (*a*) the fixed gold price and (*b*) freedom of gold and foreign-exchange transactions.

A country can be effectively on a gold standard even though its legal standard is bimetallism. This happens if the gold–silver mint-price ratio is greater than the world price ratio. In contrast, even though a country is legally on a gold standard, its government and banks could 'suspend specie payments', that is, refuse to convert their notes into gold; so that the country is in fact on a 'paper standard'.

## Countries on the classical gold standard

Britain, France, Germany and the United States were the 'core countries' of the gold standard. Britain was the 'centre country', indispensable to the spread and functioning of the standard. Legally bimetallic from the mid-13th century, Britain switched to an effective gold standard early in the 18th century. The gold standard was formally adopted in 1816, ironically during a paper-standard regime (Bank Restriction Period). The United States was legally bimetallic from 1786 and on an effective gold standard from 1834, with a legal gold standard established in 1873–4 – also during a paper standard (the greenback period). In 1879 the United States went back to gold, and by that year not only the core countries but also some British dominions and non-core western European countries were on the gold standard. As time went on, a large number of other countries throughout the globe adopted gold; but they (along with the dominions) were in 'the periphery' – acted on rather than actors – and generally (except for the dominions) not as committed to the gold standard.

Almost all countries were on a mixed coin standard. Some periphery countries were on a gold-exchange standard, usually because they were colonies or territories of a country on a coin standard.

In 1913, the only countries not on gold were traditional silver-standard countries (Abyssinia, China, French Indochina, Hong Kong, Honduras, Morocco, Persia, Salvador), some Latin American paper-standard countries (Chile, Colombia, Guatemala, Haiti, Paraguay), and Portugal and Italy (which had left gold but 'shadowed' the gold standard, pursuing policies as if they were gold-standard countries, keeping the exchange rate relatively stable).

## Elements of instability in classical gold standard

Three factors made for instability of the classical gold standard. First, the use of foreign exchange as official reserves increased as the gold standard progressed. While by 1913 only Germany among the core countries held any measurable amount of foreign exchange, the percentage for the rest of the world was double that for Germany. If there were a rush to cash in foreign exchange for gold, reduction of the gold of reserve-currency countries would place the gold standard in jeopardy.

Second, Britain was in a particularly sensitive situation. In 1913, almost half of world foreign-exchange reserves was in sterling, but the Bank of England had only three per cent of gold reserves. The Bank of England's 'reserve ratio' (ratio of 'official reserves' to 'liabilities to foreign monetary authorities held in London financial institutions') was

only 31 per cent, far lower than those of the monetary authorities of the other core countries. An official run on sterling could force Britain off the gold standard. Private foreigners also held considerable liquid assets in London, and could themselves initiate a run on sterling.

Third, the United States was a source of instability to the gold standard. Its Treasury held a high percentage of world gold reserves (in 1913, more than that of the three other core countries combined). With no central bank and a decentralized banking system, financial crises were more frequent and more severe than in the other core countries. Far from the United States assisting Britain, gold often flowed from the Bank of England to the United States, to satisfy increases in US demand for money. In many years the United States was a net importer rather than exporter of capital to the rest of the world – the opposite of the other core countries. The political power of silver interests and recurrent financial panics led to imperfect credibility in the US commitment to the gold standard. Indeed, runs on banks and on the Treasury gold reserve placed the US gold standard near collapse in the 1890s. The credibility of the Treasury's commitment to the gold standard was shaken; twice the US gold standard was saved only by cooperative action of the Treasury and a bankers' syndicate, which stemmed gold exports.

## Automatic force for stability: price specie-flow mechanism

The money supply is the product of the money multiplier and the monetary base. The monetary authority alters the monetary base by changing its gold holdings and domestic assets (loans, discounts, and securities). However, the level of its domestic assets is dependent on its gold reserves, because the authority generates demand liabilities (notes and deposits) by increasing its assets, and convertibility of these liabilities must be supported by a gold reserve. Therefore the gold standard provides a constraint on the level (or growth) of the money supply.

Further, balance-of-payments surpluses (deficits) are settled by gold imports (exports) at the gold import (export) point. The change in the money supply is the product of the money multiplier and the gold flow, providing the monetary authority does not change its domestic assets. For a country on a gold-exchange standard, holdings of foreign exchange (a reserve currency) take the place of gold.

A country experiencing a balance-of-payments deficit loses gold and its money supply decreases *automatically*. Money income contracts and the price level falls, thereby increasing exports and decreasing imports. Similarly, a surplus country gains gold, exports decrease, and imports increase. In each case, balance-of-payments equilibrium is restored via the current account, the 'price specie-flow mechanism'. To the extent that wages and prices are inflexible, movements of real income in the same direction as money income occur; the deficit country suffers unemployment, while the payments imbalance is corrected.

The capital account also acts to restore balance, via interest-rate increases in the deficit country inducing a net inflow of capital. The interest-rate increases also reduce real investment and thence real income and imports. The opposite occurs in the surplus country.

## Rules of the game

Central banks were supposed to reinforce (rather than 'sterilize') the effect of gold flows on the monetary base, thereby enhancing the price specie-flow mechanism. A gold outflow decreases the international assets of the central bank and the money supply. The central-bank's 'proper' response is: (1) decrease lending and sell securities, thereby decreasing domestic assets and the monetary base; (2) raise its 'discount rate', which induces commercial banks to adopt a higher reserves–deposit ratio, thereby reducing the money multiplier. On both counts, the money supply is further decreased. Should the central bank increase its domestic assets when it loses gold, it engages in sterilization of the gold flow, violating the 'rules of the game'. The argument also holds for gold inflow, with sterilization involving the central bank decreasing its domestic assets when it gains gold.

Monetarist theory suggests the 'rules' were inconsequential. Under fixed exchange rates, gold flows adjust money supply to money demand; the money supply is not determined by policy. Also, prices, interest rates, and incomes are determined worldwide. Even core countries can influence these variables domestically only to the extent that they help determine them in the global marketplace. Therefore the price-specie flow and like mechanisms cannot occur. Historical data support this conclusion: gold flows were too small to be suggestive of these processes; and, at least among the core countries, prices, incomes, and interest rates moved closely in correspondence, contradicting the specie-flow mechanism and rules of the game.

Rather than rule (1), central-bank domestic and international assets moving in the same direction, the opposite behaviour – sterilization – was dominant, both in core and non-core European countries. The Bank of England followed the rule more than any other central bank, but even so violated it more often than not!

The Bank of England did, in effect, manage its discount rate ('Bank Rate') in accordance with rule (2). The Bank's primary objective was to maintain convertibility of its notes into gold, and its principal tool was Bank Rate. When the Bank's 'liquidity ratio' (ratio of gold reserves to outstanding note liabilities) decreased, it usually increased Bank Rate. The increase in Bank Rate carried with it market short-term interest rates, inducing a short-term capital inflow and thereby moving the exchange rate away from the gold-export point. The converse also held, with a rise in the liquidity ratio generating a Bank Rate decrease. The Bank was constantly monitoring its liquidity ratio, and in response altered Bank Rate almost 200 times over 1880–1913.

While the Reichsbank also generally moved its discount rate inversely to its liquidity ratio, other central banks often violated rule (2). Discount-rate changes were of inappropriate direction, or of insufficient magnitude or frequency. The Bank of France kept its discount rate stable, choosing to have large gold reserves, with payments imbalances accommodated by fluctuations in its gold rather than financed by short-term capital flows. The United States, lacking a central bank, had no discount rate to use as a policy instrument.

## Reason for stability: credible commitment to convertibility

From the late 1870s onward, there was absolute private-sector credibility in the commitment to the fixed domestic-currency price of gold on the part of Britain,

France, Germany, and other important European countries. For the United States, this absolute credibility applied from about 1900. That commitment had a contingency aspect: convertibility could be suspended in the event of dire emergency; but, after normal conditions were restored, convertibility and honouring of gold contracts would be re-established at the pre-existing mint price – even if substantial deflation was required to do so. The Bank Restriction and greenback periods were applications of the contingency. From 1879, the 'contingency clause' was exercised by none of these countries.

The absolute credibility in countries' commitment to convertibility at the existing mint price implied that there was zero 'convertibility risk' (Treasury or central-bank notes non-redeemable in gold at the established mint price) and zero 'exchange risk' (alteration of mint parity, institution of exchange control, or prohibition of gold export).

Why was the commitment to credibility so credible?

1. Contracts were expressed in gold; abandonment of convertibility meant violation of contracts – anathema to monetary authorities.
2. Shocks to economies were infrequent and generally mild.
3. The London capital market was the largest, most open, most diversified in the world, and its gold market was also dominant. A high proportion of world trade was financed in sterling, London was the most important reserve-currency centre, and payments imbalances were often settled by transferring sterling assets rather than gold. Sterling was an international currency – a boon to other countries, because sterling involved positive interest return, and its transfer costs were much less than those of gold. Advantages to Britain were the charges for services as an international banker, differential interest return on its financial intermediation, and the practice of countries on a sterling (gold-exchange) standard of financing payments surpluses with Britain by piling up short-term sterling assets rather than demanding Bank gold.
4. 'Orthodox metallism' – authorities' commitment to an anti-inflation, balanced-budget, stable-money policy – reigned. This ideology implied low government spending, low taxes, and limited monetization of government debt. Therefore, it was not expected that a country's price level would get out of line with that of other countries.
5. Politically, gold had won over paper and silver, and stable-money interests (bankers, manufacturers, merchants, professionals, creditors, urban groups) over inflationary interests (farmers, landowners, miners, debtors, rural groups).
6. There was a competitive environment and freedom from government regulation. Prices and wages were flexible. The core countries had virtually no capital controls, Britain had adopted free trade, and the other core countries had only moderate tariffs. Balance-of-payments financing and adjustment were without serious impediments.
7. With internal balance an unimportant goal of policy, preservation of convertibility of paper currency into gold was the primary policy objective. Sterilization of gold flows, though frequent, was more 'meeting the needs of trade' (passive monetary policy) than fighting unemployment (active monetary policy).

8. The gradual establishment of mint prices over time ensured that mint parities were in line with relative price levels; so countries joined the gold standard with exchange rates in equilibrium.
9. Current-account and capital-account imbalances tended to be offsetting for the core countries. A trade deficit induced a gold loss and a higher interest rate, attracting a capital inflow and reducing capital outflow. The capital-exporting core countries could stop a gold loss simply by reducing lending abroad.

### Implications of credible commitment

Private parties reduced the need for balance-of-payments adjustment, via both gold-point arbitrage and stabilizing speculation. When the exchange rate was outside the spread, gold-point arbitrage quickly returned it to the spread. Within the spread, as the exchange value of a currency weakened, the exchange rate approaching the gold-export point, speculators had an ever greater incentive to purchase domestic with foreign currency (a capital inflow). They believed that the exchange rate would move in the opposite direction, enabling reversal of their transaction at a profit. Similarly, a strengthened currency involved a capital outflow. The further the exchange rate moved toward a gold point, the greater the potential profit opportunity in betting on a reversal of direction; for there was a decreased distance to that gold point and an increased distance from the other point. This 'stabilizing speculation' increased the exchange value of depreciating currencies, and thus gold loss could be prevented. Absence of controls meant such private capital flows were highly responsive to exchange-rate changes.

### Government policies that enhanced stability

Specific government policies enhanced gold-standard stability. First, by the turn of the 20th century, South Africa – the main world gold producer – was selling all its gold output in London, either to private parties or to the Bank of England. Thus the Bank had the means to replenish its gold reserves. Second, the orthodox-metallism ideology and the leadership of the Bank of England kept countries' monetary policies disciplined and in harmony. Third, the US Treasury and the central banks of the other core countries manipulated gold points, to stem gold outflow. The cost of exporting gold was artificially increased (for example, by increasing selling prices for bars and foreign coin) and/or the cost of importing gold artificially decreased (for example, by providing interest-free loans to gold importers).

Fourth, central-bank cooperation was forthcoming during financial crises. The precarious liquidity position of the Bank of England meant that it was more often the recipient than the provider of financial assistance. In crises, the Bank would obtain loans from other central banks, and the Bank of France would sometimes purchase sterling to support that currency. When needed, assistance went from the Bank of England to other central banks. Also, private bankers unhesitatingly made loans to central banks in difficulty.

Thus, 'virtuous' interactions were responsible for the stability of the gold standard. The credible commitment to convertibility of paper money at the established mint price, and therefore to fixed mint parities, were both a cause and an effect of the

stable environment in which the gold standard operated, the stabilizing behaviour of arbitrageurs and speculators, and the responsible policies of the authorities – and these three elements interacted positively among themselves.

## Experience of periphery

An important reason for periphery countries to join and maintain the gold standard was the fostering of access to core-countries' capital markets. Adherence to the gold standard connoted that the peripheral country would follow responsible macroeconomic policies and repay debt. This 'seal of approval', by reducing the risk premium, involved a lower interest rate on the country's bonds sold abroad, and very likely a higher volume of borrowing, thereby enhancing economic development.

However, periphery countries bore the brunt of the burden of adjustment of payments imbalances with the core (and other western European) countries. First, when the gold-exchange-standard periphery countries ran a surplus (deficit), they increased (decreased) their liquid balances in the United Kingdom (or other reserve-currency country) rather than withdraw gold from (lose gold to) the reserve-currency country. The monetary base of the periphery country increased (decreased), but that of the reserve-currency country remained unchanged. Therefore, changes in domestic variables – prices, incomes, interest rates, portfolios – that occurred to correct the imbalance were primarily in the periphery.

Second, when Bank Rate increased, London drew funds from France and Germany, which attracted funds from other European countries, which drew capital from the periphery. Also, it was easy for a core country to correct a deficit by reducing lending to, or bringing capital home from, the periphery. While the periphery was better off with access to capital, its welfare gain was reduced by the instability of capital import. Third, periphery-countries' exports were largely primary products, sensitive to world market conditions. This feature made adjustment in the periphery take the form more of real than financial correction.

The experience of adherence to the gold standard differed among periphery groups. The important British dominions and colonies successfully maintained the gold standard. They paid the price of serving as an economic cushion to the Bank of England's financial situation; but, compared with the rest of the periphery, gained a stable long-term capital inflow. In southern Europe and Latin America, adherence to the gold standard was fragile. The commitment to convertibility lacked credibility, and resort to a paper standard occurred. Many of the reasons for credible commitment that applied to the core countries were absent. There were powerful inflationary interests, strong balance-of-payments shocks, and rudimentary banking sectors. The cost of adhering to the gold standard was apparent: loss of the ability to depreciate the currency to counter reductions in exports. Yet the gain, in terms of a steady capital inflow from the core countries, was not as stable or reliable as for the British dominions and colonies.

## Breakdown of classical gold standard

The classical gold standard was at its height at the end of 1913, ironically just before it came to an end. The proximate cause of the breakdown of the classical gold standard

was the First World War. However, it was the gold-exchange standard and the Bank of England's precarious liquidity position that were the underlying cause. With the outbreak of war, a run on sterling led Britain to impose extreme exchange control – a postponement of both domestic and international payments – making the international gold standard inoperative. Convertibility was not suspended legally; but moral suasion, legalistic action, and regulation had the same effect. The Bank of England commandeered gold imports and applied moral suasion to bankers and bullion brokers to restrict gold exports.

The other gold-standard countries undertook similar policies – the United States not until 1917, when it adopted extra-legal restrictions on convertibility and restricted gold exports. Commercial banks converted their notes and deposits only into currency. Currency convertibility made mint parities ineffective; floating exchange rates resulted.

## Return to the gold standard

After the First World War, a general return to gold occurred; but the interwar gold standard differed institutionally from the classical gold standard. First, the new gold standard was led by the United States, not Britain. The US embargo on gold exports was removed in 1919, and currency convertibility at the pre-war mint price was restored in 1922. The gold value of the dollar rather than pound sterling was the typical reference point around which other currencies were aligned and stabilized. The core now had two central countries, the United Kingdom (which restored gold in 1925) and the United States.

Second, for many countries there was a time lag between stabilizing the currency in the foreign-exchange market (fixing the exchange rate or mint parity) and resuming currency convertibility. The interwar gold standard was at its height at the end of 1928, after all core countries were fully on the standard and before the Great Depression began. The only countries that never joined the interwar gold standard were the USSR, silver–standard countries (China, Hong Kong, Indochina, Persia, Eritrea), and some minor Asian and African countries.

Third, the 'contingency clause' of convertibility conversion, that required restoration of convertibility at the mint price that existed prior to the emergency (the First World War), was *broken* by various countries, and even core countries. While some countries (including the United States and United Kingdom) stabilized their currencies at the pre-war mint price, others (including France) established a gold content of their currency that was a fraction of the pre-war level: the currency was devalued in terms of gold, the mint price was higher than pre-war. Still others (including Germany) stabilized new currencies adopted after hyperinflation.

Fourth, the gold coin standard, dominant in the classical period, was far less prevalent in the interwar period. All four core countries had been on coin in the classical gold standard; but only the United States was on coin interwar. The gold-bullion standard, non-existent pre-war, was adopted by the United Kingdom and France. Germany and most non-core countries were on a gold-exchange standard.

## Instability of interwar gold standard

The interwar gold standard was replete with forces making for *instability*.

1. The process of establishing fixed exchange rates was piecemeal and haphazard, resulting in disequilibrium exchange rates. Among core countries, the United Kingdom restored convertibility at the pre-war mint price without sufficient deflation, and had an overvalued currency of about ten per cent. France and Germany had undervalued currencies.
2. Wages and prices were less flexible than in the pre-war period.
3. Higher trade barriers than pre-war also restrained adjustment.
4. The gold-exchange standard economized on total world gold via the gold of the United Kingdom and United States in their reserves role for countries on the gold-exchange standard and also for countries on a coin or bullion standard that elected to hold part of their reserves in London or New York. However, the gold-exchange standard was unstable, with a conflict between (*a*) the expansion of sterling and dollar liabilities to foreign central banks, to expand world liquidity, and (*b*) the resulting deterioration in the reserve ratio of US and UK authorities.This instability was particularly severe, for several reasons. First, France was now a large official holder of sterling, and France was resentful of the United Kingdom. Second, many more countries were on the gold-exchange standard than pre-war. Third, the gold-exchange standard, associated with colonies in the classical period, was considered a system inferior to a coin standard.
5. In the classical period, London was the one dominant financial centre; in the interwar period it was joined by New York and, in the late 1920s, Paris. Private and official holdings of foreign currency could shift among the two or three centres, as interest-rate differentials and confidence levels changed.
6. There was maldistribution of gold. In 1928, official reserve-currency liabilities were much more concentrated than in 1913, British pounds accounting for 77 per cent of world foreign-exchange reserves and French francs less than two per cent (versus 47 and 30 per cent in 1913). Yet the United Kingdom held only seven per cent of world official gold and France 13 per cent. France also possessed 39 per cent of world official foreign exchange. The United States held 37 per cent of world official gold.
7. Britain's financial position was even more precarious than in the classical period. In 1928, the gold and dollar reserves of the Bank of England covered only one-third of London's liquid liabilities to official foreigners, a ratio hardly greater than in 1913. UK liquid liabilities were concentrated on stronger countries (France, United States), whereas UK liquid assets were predominantly in weaker countries (Germany). There was ongoing tension with France, which resented the sterling-dominated gold-exchange standard and desired to cash in its sterling holding for gold, to aid its objective of achieving first-class financial status for Paris.
8. Internal balance was an important goal of policy, which hindered balance-of-payments adjustment, and monetary policy was influenced by domestic politics rather than geared to preservation of currency convertibility.
9. Credibility in authorities' commitment to the gold standard was not absolute. Convertibility risk and exchange risk could be high, and currency speculation could

be destabilizing rather than stabilizing. When a country's currency approached or reached its gold-export point, speculators might anticipate that currency convertibility would not be maintained and that the currency would be devalued.

10. The 'rules of the game' were violated even more often than in the classical gold standard. Sterilization of gold inflows by the Bank of England can be viewed as an attempt to correct the overvalued pound by means of deflation. However, the US and French sterilization of their persistent gold inflows reflected exclusive concern for the domestic economy and placed the burden of adjustment (deflation) on other countries.
11. The Bank of England did not provide a leadership role in any important way, and central-bank cooperation was insufficient to establish credibility in the commitment to currency convertibility. The Federal Reserve had three targets for its discount-rate policy: strengthen the pound, combat speculation in the New York stock market, and achieve internal balance – and the first target was of lowest priority. Although, for the sake of external balance, the Bank of England kept Bank Rate higher than internal considerations would dictate, it was understandably reluctant to abdicate Bank Rate policy entirely to the balance of payments, with little help from the Federal Reserve. To keep the pound strong, substantial international cooperation was required, but was not forthcoming.

## Breakdown of interwar gold standard

The Great Depression triggered the unravelling of the gold standard. The depression began in the periphery. Low export prices and debt-service requirements created insurmountable balance-of-payments difficulties for gold-standard commodity producers. However, US monetary policy was an important catalyst. In 1927 the Federal Reserve favoured easy money, which supported foreign currencies but also fed the New York stock-market boom. Reversing policy to tame the boom, higher interest rates attracted monies to New York, weakening sterling in particular. The crash of October 1929, while helping sterling, was followed by the US depression. This spread worldwide, with declines in US trade and lending. In 1929 and 1930 a number of periphery countries – both dominions and Latin American countries – either formally suspended currency convertibility or restricted it so that currencies violated the gold-export point.

It was destabilizing speculation, emanating from lack of confidence in authorities' commitment to currency convertibility, which ended the interwar gold standard. In May 1931 there was a run on Austria's largest commercial bank, and the bank failed. The run spread to other eastern European countries and to Germany, where an important bank also collapsed. The countries' central banks lost substantial reserves; international financial assistance was too late; and in July 1931 Germany adopted exchange control, followed by Austria in October. These countries were definitively off the gold standard.

The Austrian and German experiences, as well as British budgetary and political difficulties, were among the factors that destroyed confidence in sterling, which occurred in mid-July 1931. Runs on sterling ensued, and the Bank of England lost much of its reserves. Loans from abroad were insufficient, and in any event taken as a sign of weakness. The gold standard was abandoned in September, and the pound quickly and

sharply depreciated on the foreign-exchange market, as overvaluation of the pound would imply.

Following the UK abandonment of the gold standard, many countries followed, some to maintain their competitiveness via currency devaluation, others in response to destabilizing capital flows. The United States held on until 1933, when both domestic and foreign demands for gold, manifested in runs on US commercial banks, became intolerable. 'Gold bloc' countries (France, Belgium, Netherlands, Switzerland, Italy, Poland), with their currencies now overvalued and susceptible to destabilizing speculation, succumbed to the inevitable by the end of 1936.

The Great Depression was worsened by the gold standard: gold-standard countries hesitated to inflate their economies, for fear of suffering loss of gold and foreign-exchange reserves, and being forced to abandon convertibility or the gold parity. The gold standard involved 'golden fetters', which inhibited monetary and fiscal policy to fight the Depression. As countries left the gold standard, removal of monetary and fiscal policy from their 'gold fetters' enabled their use in expanding real output, providing the political will existed.

In contrast to the interwar gold standard, the classical gold standard functioned well because of a confluence of 'virtuous' interactions, involving government policies, credible commitment to the standard, private arbitrage and speculation, and fostering economic and political environment. We will not see its like again.

LAWRENCE H. OFFICER

*See Also* **banking crises; Bank of England; bimetallism; Bretton Woods system; commodity money; silver standard; specie-flow mechanism.**

### *Bibliography*

Bayoumi, T., Eichengreen, B. and Taylor, M.P., eds. 1996. *Modern Perspectives on the Gold Standard*. Cambridge: Cambridge University Press.

Bordo, M.D. and Kydland, F.E. 1995. The gold standard as a rule: an essay in exploration. *Explorations in Economic History* 32, 423–64.

Bordo, M.D. and Rockoff, H. 1996. The gold standard as a 'Good Housekeeping Seal of Approval'. *Journal of Economic History* 56, 389–428.

Bordo, M.D. and Schwartz, A.J., eds. 1984. *A Retrospective on the Classical Gold Standard, 1821–1931*. Chicago: University of Chicago Press.

De Macedo, J.B., Eichengreen, B. and Reis, J., eds. 1996. *Currency Convertibility: The Gold Standard and Beyond*. London: Routledge.

Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939*. New York: Oxford University Press.

Eichengreen, B. and Flandreau, M. 1997. *The Gold Standard in Theory and History*, 2nd edn. London: Routledge.

Gallarotti, G.M. 1995. *The Anatomy of an International Monetary Regime: The Classical Gold Standard, 1880–1914*. New York: Oxford University Press.

Officer, L.H. 1996. *Between the Dollar–Sterling Gold Points*. Cambridge: Cambridge University Press.

Officer, L.H. 2001. Gold standard. *EH. Net Encyclopedia*, ed. R. Whaples. Online. Available at http://eh.net/encyclopedia/article/officer.gold.standard, accessed 20 October 2006.

# Greek crisis in perspective: Origins, effects and ways-out

## 1. Introduction

In the aftermath of the global financial crisis of 2008, a number of Eurozone countries were engulfed in a spiral of rising public deficits and explosive borrowing costs that eventually drove them out of markets and into bail-out agreements jointly undertaken by the International Monetary Fund (IMF), the European Union (EU) and the European Central Bank (ECB). Greece was by far the most perilous case with a double-digit fiscal deficit, an accelerating public debt which in GDP terms was twice as much the Eurozone average and an external deficit near 5,000 US Dollars per capita in 2008, one of the largest worldwide. No wonder that Greece was the first to seek the bail-out assistance and the last expected to exit its ever-changing conditionality terms.

Two and a half years after the bail-out Memorandum was signed, the situation remains highly uncertain. The economy faces an unprecedented recession, unemployment is rocketing, social unrest undermines the implementation of reforms and the fiscal front is not yet under control, despite extensive cuts in wages, salaries and pensions. In the summer of 2011 uncertainties multiplied at such a rate that the possibility of Greece exiting the Eurozone was widely discussed either as a punishment mechanism from abroad for not accepting the pains of adjustment or as a quick fix from within to avoid them for good.

In two subsequent EU summits, held respectively in July and October 2011, the Memorandum agreement was substantially broadened to include a radical debt reduction, a second round of bail-out loans by IMF and the EU and a generous release of European structural funds to assist the real economy. The agreement was conditional on being approved by the national Parliaments of the lender states as well as by the European Parliament. Finally, the conditionalities of the Memorandum were approved by the Greek Parliament in February 2012 and the debt-cutting process was concluded in May. However, most of the envisaged measures were delayed for the third quarter of the year, as two round of elections took place to provide new legitimacy for carrying on the program and implementing reforms. The prolonged electoral uncertainty meant that most of the adjustment measures were weakened or postponed, leading to new tensions over Greece's determination. A coalition Government was finally formed in June by parties vowing to apply all policies deemed necessary for the country to remain in the Eurozone, though at the same time seeking some relaxation of the time frame from European authorities.

[a]AUEB, Athens University of Economics and Business, Department of International and European Economic Studies (DIEES), Patision Str 76, Athens 10434, Greece. Fax: ++30-210-8221011. E-mail address: nchris@aueb.gr

It is tempting to note that the economic capacity of the country to adjust and the social endurance are diminished exactly when the European environment is becoming more helpful for stressed countries, This makes the Greek problem an unusually interesting case for analysis, not only for understanding its origins and causes but also for devising a realistic strategy to solve it.

The purpose of the present article is twofold: First to provide a historical account of debt accumulation, identify the main difficulties of fiscal stabilization and explain the factors that led to the present crisis and the failure to prepare for it. Second, to assess the main reasons for missing the targets set by the Memorandum agreement and the need for encompassing a growth strategy in order to make reforms acceptable and more effective to achieve debt sustainability in the longer run.

Section 2 describes the main episodes of debt escalation in the 1980s, Section 3 the stabilization effort on the way to EMU and Section 4 the toxic combination of fiscal irresponsibility, external deficits and political indecision during the more recent period that led to the present crisis. Section 5 describes some recurrent facts on fiscal policies that repeatedly hinder stabilization and growth. Section 6 attempts an ex post assessment of the policies conditioned by the Memorandum agreement to correct the economy while Section 7 argues why exiting the Eurozone should not be an option for Greece. An alternative scenario based on higher growth is shown to be more credible in achieving fiscal consolidation and stabilizing the debt over the medium term. Section 8 concludes with the need to fight the current recession as the only way for Greece to regain social coherence and debt sustainability in the new landscape of the Eurozone.

## 2. The period of debt escalation: 1980–1993

In 1980, Greece became a full-fledged member of the European Union and this marked a wholly new period for the economic and political developments in the country. Greece was one of the first non-founding countries to start accession talks with the Common Market as early as 1961, but the process was abruptly suspended with the advent of the military dictatorship that lasted until 1974. Membership in the European Union was rightly viewed as an anchor of political and institutional stability for the newly restored democracy, but nonetheless it also fed and multiplied uncertainties over the economy.

After a long period of growth, Greece entered a period of recession in late 1970s, not only as a consequence of worldwide stagflation, but also because – on its way to integration with the common market – it had to dismantle its preferential system of subsidies, tariffs and state procurement by which several companies were kept profitable without being competitive. Soon after accession, many of these companies went out of business and unemployment rose for the first time in many decades.

The Government opted for a massive fiscal expansion that included demand–push policies to boost activity and the public underwriting of several ailing companies to maintain employment. The effect was quite predictable: private debts turned into a chronic hemorrhage of budget deficits without any supply-side improvements. Similarly, the expansion of demand simply led to more imports and higher prices.

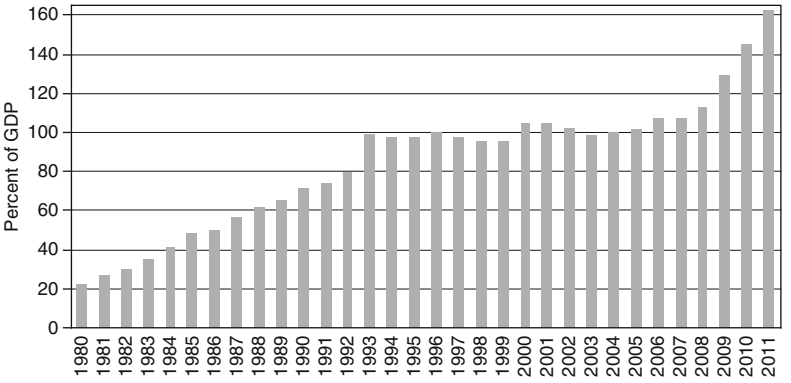



**Figure 1** Greek public Debt as %GDP for the period 1980–2011. Source: Debt of General Government, ESA95 definition, Ameco Eurostat 2011. GDP at market prices, IMF WEO Database 2010.

Activity got stuck and Greece ended up in a typical stagflation, perhaps the quickest assimilation to European practices of the time.

As a result, accession to the Promised Land strangely coincided with the unleashing of a nightmare thought to be in dormant thus far: *public debt*.

Looking at Fig. 1, there are three distinguishable phases for the dynamics of debt: The first covers the period 1980–1993 during which public debt rose from slightly above 20% of GDP toward 100% in 1993. The second phase spans the period 1994–2005 in which public debt ends up again at around 100% of GDP after two mild reductions in between. The third phase covers the period 2006–2011 when public debt surpasses the 100% threshold, accelerates after 2008 and ends up exceeding 160% of GDP in 2011.

The above periodicity broadly coincides with substantial shifts in the context of economic policies, as suggested by developments in the fiscal patterns shown in Fig. 2 and in the Current Account depicted in Fig. 3 and briefly discussed below.

Regarding fiscal developments, the main characteristic of the first period was the substantial expansion of public spending and the concomitant rise in budget deficits and government debt. Revenues increased as a proportion of GDP, but were outpaced by the steadily growing expenditure. Both fiscal components appear to be volatile in the election years 1981, 1985 and 1989, suggesting the presence of a strong political cycle in public finances, as will be discussed later in more detail.

To maintain competitiveness, authorities had adopted, since the mid 1970s, a real exchange-rate target with a crawling peg. After the Government adopted an automatic wage indexation scheme in 1982, the only effect of the exchange rate policy was to fuel price increases and aggravate trade deficits. To break the vicious cycle of depreciation and inflation, a discrete devaluation combined with a temporary wage freeze was implemented in 1983, but it was superseded by a new phase of expansion as elections were approaching leaving public debt at even higher levels.

The external deficit approached 8% of GDP in 1985, an alarming threshold as several Latin American economies with similar imbalances were serially collapsing at

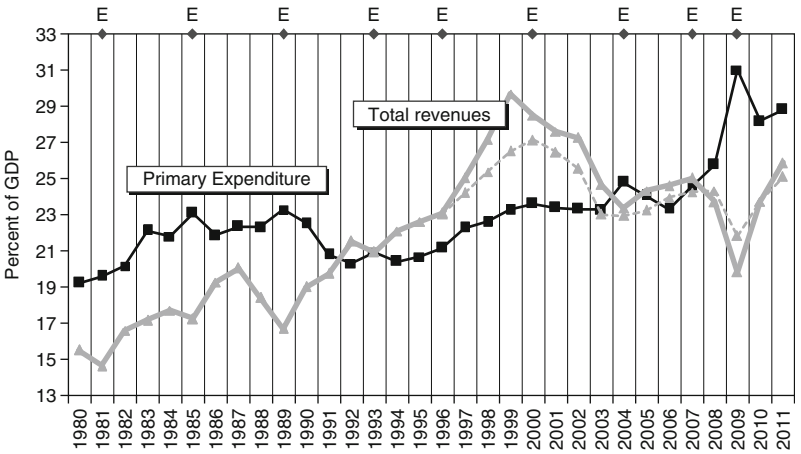



**Figure 2** Primary public expenditure and total revenues (incl. privatisation proceeds) as %GDP in Greece, 1980–2011. Election years denoted by (E). Dotted line denotes public receipts net of privatizations. Source: Budget Reports. GDP at market prices, AMECO Database 2012.

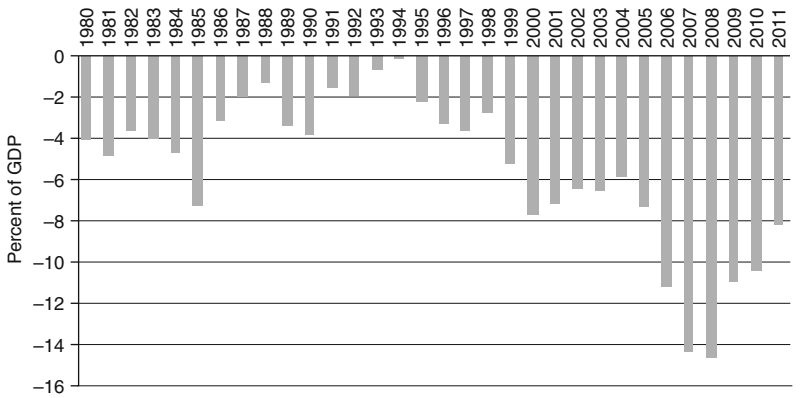



**Figure 3** Current Account in Greece as % of GDP, 1980–2011. Source: IMF WEO Database 2011.

that time. A coherent stabilization program was called for in October 1985 enforcing a discrete devaluation by 15%, a tough incomes policy and extensive cuts in public spending. The program achieved a rise in revenues by beating several tax-evasion practices and replacing previous indirect taxes with the more effective VAT system adopted by the European Union. Public debt was stabilized, but only until the program was finally abandoned in 1988, after being fiercely opposed from within the Government and the ruling party.

*The first fiscal crisis*

Two general elections in 1989 failed to secure majority, thus leading to the formation of coalition Governments, an event that was hailed as a confirmation of political maturing and an opportunity to overcome partisan differences on major issues. But

self-indulging admiration was short-lived, as stabilization policies are notoriously difficult to implement through party coalitions because each party tries to avoid the cost falling on its own constituency. Greece was no exception to the rule and the economy suffered a major setback in 1989, far more serious than previous fiscal failures.

Two episodes are characteristic of how a rhetoric designed to please everybody in combination with naïve policies can lead to disaster: Despite looming deficits, in 1989 the coalition Government decided to abolish prison terms for major tax arrears hoping to induce offenders to repent and reconsider their strategy. Expectedly, the move was interpreted the other way around as a signal of relaxed monitoring in the future, thus encouraging further evasion.

Another bizarre policy was to cut import duties for car purchases by repatriates returning to Greece after the collapse of the Soviet Union. The measure was viewed as a gesture to facilitate mobility back in the motherland, but it was quickly turned into a black-market scheme. At a small bribe, immigrants were purchasing luxury cars only to immediately resell them to rich clients who could thereby avoid the duty tax. The Budget was deprived of badly needed revenues and evaders had yet another reason for celebration.

As a result, revenues collapsed and the country suffered a major fiscal crisis, until a majority Government was elected in 1990 and enacted a new stabilization program. Despite substantial cuts in spending and a rise in revenues, public debt as a ratio to GDP continued to rise because of the higher cost of borrowing worldwide and a stagnant output. The sharp rise in 1993 in particular, is due to the inclusion of extensive debts initially contracted by public companies under state guarantees but finally underwritten by the Budget. Except for the electoral years 1989–90, fiscal consolidation significantly improved the Current Account and such a rarity as a balanced external position was reached in 1994.

## 3. Debt stabilization and EMU membership

Although Greece was a signatory of the Maastricht Treaty in 1991, it was far from obvious whether, how and when the country could comply with the nominal convergence criteria required to join the Economic and Monetary Union. Public deficits and inflation were galloping at two-digit levels and there was great uncertainty about the viability of the exchange rate system; for a detailed analysis of the period see Christodoulakis (1994).

In May 1994, capital controls were lifted in compliance with European guidelines and this prompted fierce speculation in the forex market. Interest rates reached particularly high levels and the Central Bank of Greece exhausted most of its reserves to stave off the attack; for an account of the successful defense see Flood and Kramer (1996). This episode proved to be a turning point for the determination of Greece to pursue accession to EMU in order to be shielded by the common currency and avoid similar attacks in the future . Soon afterwards the "Convergence Program" was adopted that set time limits to satisfy the Maastricht criteria and included a battery of reforms in the banking and the public sectors.

International markets were not impressed and continued to be unconvinced about exchange rate viability. With the advent of the Asian crisis in 1997 spreads rose again

dramatically and – after months of credit shortages – Greece finally decided to devalue by 12.5% in March 1998 and subsequently enter the Exchange Rate Mechanism wherein it had to stay for two years. The country was not yet ready to join the first round of Eurozone countries in 1998,and Greece was granted a transition period to comply with the convergence criteria by the end of 1999.

After depreciation, credibility was further enhanced by structural reforms and reduced state borrowing so that when the Russian crisis erupted in August 1998, the currency came under very little pressure. Public expenditure was kept below the peaks it had reached in the previous decade and was increasingly outpaced by the rising revenues and various one-off receipts. Tax collection was enhanced by the introduction of a scheme of minimum turnover on SMEs, the elimination of a vast number of tax allowances, the imposition of a new levy on large property and a re-reorganization of the auditing system. Proceeds were further augmented by privatization of public companies and, as result, public debt fell to 93% of GDP in 1999. Although still higher than the 60% threshold required by the European Treaty, Greece benefited from the convenient interpretation that it suffices "*to lean toward that level*", as previously used by other countries – such as Italy and Belgium – in their own way to enter EMU.

*The implementation of market reforms*

In the 1980s, structural reforms were hardly on the agenda of Greek economic policy. In fact for most of the period the term was a misnomer used to describe further state intervention in economic activity, rather than market-oriented policies as practised in other European countries. Market reforms were introduced for the first time in 1986 aiming at the modernization of the outmoded banking and financial system in compliance with European directives. A major reform in social security took place for the first time in 1992, curbing early retirement and excessively generous terms on the pension/income ratios.

Throughout the 1990s, various reform programs were aimed at the restructuring of public companies whose chronic deficits had contributed to the fiscal crisis in 1989. Privatization was attempted through direct sales of state-owned utilities as the quick way to reduce deficits. Despite some initial success, the program was fiercely opposed by the trade unions of public companies and eventually led to the demise of the Government. Privatizations were conveniently brandished as sell-outs, and it took a few more years for the concept to reappear on the political agenda.

A new wave of reforms was launched after 1996 in the course of the "Convergence Program". State banks were privatised or merged, dozens of outmoded organizations were closed down, and a series of IPOs – taking advantage of the stock market bonanza - provided capital and restructuring finance to several public utilities. Other structural changes included the lifting of closed-shop practices in shipping, the entry of more players into the mobile telephony market and a series of efforts to make the economic environment more conducive to entrepreneurship and employment.

*Post-EMU fatigue*

After 2000, Greece emulated some other euro area members in exhibiting a '*post-EMU fatigue*' and the reform process gradually slowed down. As shown in Fig. 9, proceeds

from privatization peaked in 1999, but subsequently remained low as a result of the contraction in capital markets after the dot.com bubble and the global recession in 2003; for an extensive discussion of reforms in Greece over the period 1990–2008 see Christodoulakis (2012).

An attempt in 2001 to deeply reform the pension system led to serious social confrontations and was finally abandoned. Though replaced by a watered-down version one year later, the failure left a mark of reform timidity for many years. Two other mild reforms followed in 2006 and 2010, but the social security system is still characterised by inequalities, inefficiencies and structural deficits that exert a substantial burden on the General Government finances.

The fatigue spread more widely after the Olympic Games in 2004. With the exemption of the sale of Greek Telecom to the German state company and the privatization of the national air carrier after a decade of failed attempts, most other reforms were consisting of small IPOs with no structural spillovers to the rest of the economy.

*Why debt reduction was insufficient*

Despite having achieved substantial primary surpluses throughout 1994–2002 – and around 1999 in particular – public debt over the same period fell only slightly. There are three reasons to explain this outcome. First, during this period the Government had to issue bonds to accumulate a sufficient stock of assets for the Bank of Greece as a prerequisite for its inclusion in the Euro-system, and this capital injection led to a substantial increase in public debt without affecting the deficit.

Second, after a military stand-off in the Aegean in 1996, Greece increased defence procurement to well above 4% of GDP per year. In line with Eurostat rules, the burden was fully recorded in the debt statistics at the time of ordering but only gradually in the current expenditure following the pattern of actual delivery of equipment. This practice created a considerable lag in the debt-deficit adjustment and, in 2004, the Government enforced a massive revision of the deficit figures by retroactively augmenting public spending on the date of ordering, prompting a major dispute over the quality and integrity of the statistics of public finances in Greece. Though a decision by Eurostat in 2006 made the delivery-based rule obligatory for all countries, Greece did not withdraw the self-inflicted revision. As a consequence, deficits were statistically augmented for 2000–2004 and scaled-back for 2005–2006 relative to what they should have been otherwise, in an awkward demonstration of political interference.

The third reason was the strong appreciation of the Yen/Euro exchange rate by more than 50% between 1999 and 2001. This significantly augmented Greek public debt as a proportion of output due to the fact that substantial loans were contracted in the Japanese currency during the 1994 crisis. To alleviate this exogenous deterioration, Greece entered a large currency swap in 2001 by which the debt to GDP ratio was reduced by 1.4% in exchange for a rise in deficits by 0.15% of GDP in subsequent years, so that the overall fiscal position remained unchanged in present value terms. Although the transaction had no bearing on the statistics for 1999 on which EMU entry was assessed, Greece suffered extensively from criticisms that mistook the swap as a ploy to circumvent a proper evaluation. Values shown in Fig. 1 are net of swap effects, and this partly explains the peak in 2001.

*The Current Account*

After the Eurozone became operational, hardly any attention was paid to Current Account imbalances, regarding Greece or any other deficit country. Even after they reached huge proportions, external disparities in the euro area continued to remain surprisingly unnoticed from a policy point of view. It was only in the aftermath of the 2008 crisis that policy bodies in the European Union started emphasising the adverse effects that external imbalances may have on the sustainability of the common currency (see for example EC, 2009).

The reason for this complacency was not merely that devaluations were ruled out by the common currency. A widespread – and unwisely comfortable – view held that external imbalances were mostly demand-driven effects and, as such, they would sooner or later dissipate as a result of ongoing fiscal adjustment in member-states. When, for example, Blanchard and Giavazzi (2002) asked whether countries such as Portugal or Greece should worry about and take measures to reduce their Current Account deficits they "*... conclude(d), to a first order, that they should not*". A few years later this proved to be just another misguided assessment; Blanchard (2006) – overturning his previous optimism – remarked that Current Account deficits were steadily increasing within the euro area and urged immediate action otherwise "*...implications can be bad*". And indeed they were.

Although improved for a while after the country joined the common currency, the subsequent vast deterioration in the Greek Current Account played a crucial role in inviting the global crisis home. The reason behind the initial containment was that factor income flows from abroad increased as a result of extensive Greek Foreign Direct Investment in neighbouring countries while labour immigration kept domestic wage increases at bay. The deficit started to deteriorate after 2004 as domestic demand peaked in the post Olympics euphoria, inflation differentials with other Eurozone countries widened and the Euro appreciated further. Unit labour costs increased and as shown in Fig. 4 the relevant index rose by 10% in the period 1999–2010. As an *ex post* wisdom, it is worth noticing from the same figure that a similar erosion of competitiveness took place in *all* other Eurozone countries that are currently in bailout agreements (Ireland by 12% and Portugal 8%) or considered to be at the risk of seeking one (Spain by 9% and Italy by 8%).

Compared to Germany, Greek unit labour costs increased by 27% causing significant bilateral imbalances. However, this erosion was gradual and cannot have been the single reason for the rapid deterioration experienced after 2006. Other factors affecting the investment environment, such as the quality of the regulatory framework, elimination of corruption practices and overall Government effectiveness might as well have been crucial in shaping productivity and competitiveness. Using the Worldwide Governance Indicators published by the World Bank as proxies for how the above factors evolved during the period from 1996 to 2008, Fig. 5 shows that, despite some improvement in the first years of EMU, there was a noticeable decline thereafter.

These developments were pivotal to the poor performance of Greece in attracting foreign direct investment in spite of the substantial fall in interest rates and the facilitation of capital flows within the Eurozone. As depicted in Fig. 6, FDI expressed as percent of GDP hardly improves during the last decade relative to the 1980s. The

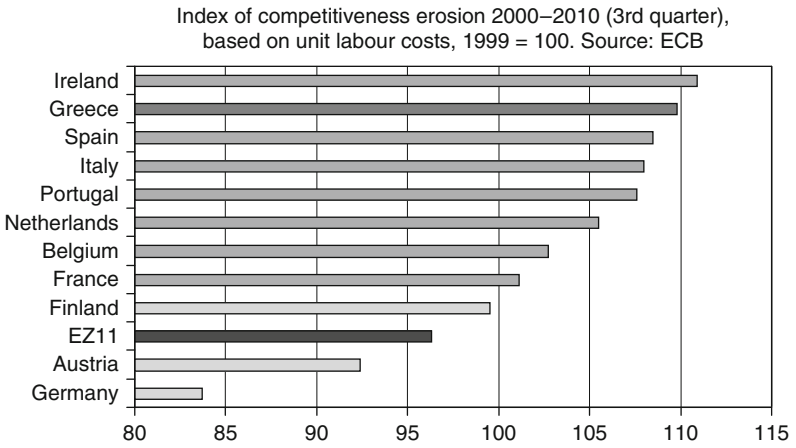



**Figure 4** Development of unit labour costs in the Eurozone 2000–10. Source: ECB, Competitiveness indicators, 2011. For Portugal the ULC index was missing for 2010 and replaced by the CPI index adjusted for differences from ULC by using the estimates for 2011. In the more recent editions, the effect of Greek ULC on competitiveness is even less pronounced, due to the wage-cuts implemented in the last quarter of 2010 and through 2011.

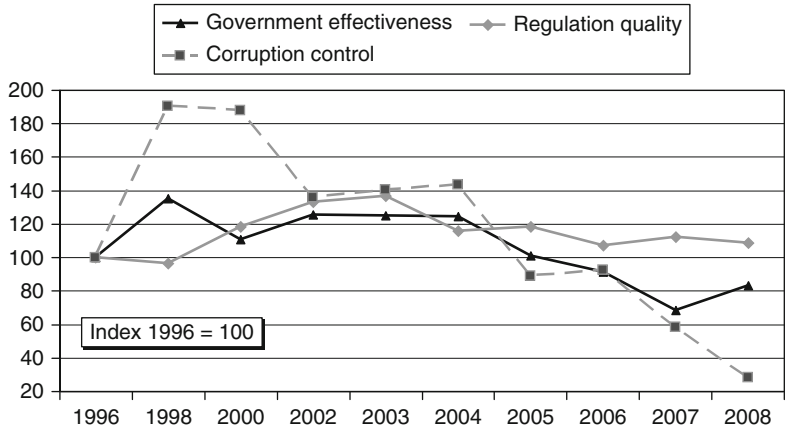



**Figure 5** Quality indicators affecting the economic climate. Notes: Indicators are measured in various units with higher values corresponding to better outcomes; to ease comparison all are here indexed at 100 in 1996. Source: World Bank, WGI various editions.

composition has also changed, as most of the FDI inflows were directed to non-manufacturing sectors and, pointedly, with an increasing allocation to real estate that further aggravates the strain in the Current Account.

It is a well established fact that when new investments are directed mainly to the tradeable sectors this leads to substantial productivity improvements and favours net exports. In contrast, investments going mostly into the real-estate sector boost aggregate demand, raise prices, cause the real exchange rate to appreciate and hinder competitiveness. These developments manifest a major failure of Greece – and for that

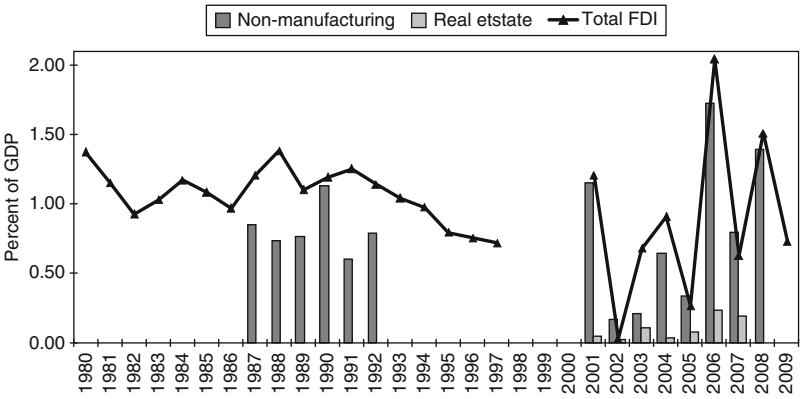



**Figure 6** FDI inflows to Greece expressed as percent of GDP. Note: Missing observations are due to non-availability and do not necessarily imply poor flows. Source: OECD, FDI statistics.

matter of other Eurozone countries – to exploit the post-EMU capital flows in order to upgrade and expand production; for details see a study by Christodoulakis and Sarantides (2011) who use the differentiation in composition and the asymmetry in the volumes of FDI to explain the diverging patterns of external balances in the Eurozone countries.

## 4. Unprepared for the 2008 crisis

The fiscal decline started with the disappearance of primary surpluses after 2003 and culminated with rocketing public expenditure and the collapse of revenues in 2009, as shown in Fig. 2. Revenues declined as a result of a major cut in corporate tax rate from 35% to 25% in 2005 and extensive inattention to the collection of revenues.

Such decisions were making increasingly evident that stabilizing the economy was not a policy priority of the Government, and further actions soon confirmed the assumption: concerned over the rising deficits in 2007, it sought a fresh mandate to redress public finances but – despite securing a clear victory – no such action was taken after the election whatsoever. Only a few months before the global crisis actually erupted, the Government claimed that the Greek economy was *"sufficiently fortressed"* and would stay immune to the reverberations of international shocks. Even after September 2008, the Government was for a long time ambivalent as to whether implement a harsh program to stem fiscal deterioration or to expand public spending to fight off the prospect of recession. A final compromise at the end of the year included a consumption stimulus combined with a bank rescue plan of Euro 5 bn and a pledge to raise extra revenues. The first two were quickly implemented, whilst the latter was forgotten.

Weakened by internal divisions, the Government continued to be indecisive on what exactly to do and, after a defeat in the European elections in June 2009, it opted for yet another general election in October 2009 asking for a fresh mandate to address the mounting economic problems. In practice, the election period turned to be an opportunity for further largesse rather than of preparation on how to contain it. The fiscal

consequences were stunning: total public expenditure was pumped up by more than 5 percentage points exceeding 31% of GDP at the end of 2009. (In levels, it exceeded Euro 62 bn, i.e., twice the size in 2003). The rise was entirely due to consumption as public investment remained the same at 4.1% of GDP; details on how public spending was ballooned are given in Christodoulakis (2010).

Total receipts in 2009 collapsed by another 4% of GDP as a result of widespread neglect in collection and the fact that privatization proceeds turned negative since the Government had to finance the emergency capitalization of Greek banks. The deficit of General Government spiraled and its figure was serially revised from an estimated 6.7% of GDP before the elections to 12.4% in October 2009, and finally widening to 15.4% of GDP by the end of the year. It was only then that European authorities stopped their onlooking attitude and issued a number of warnings against the spending.

*Post-election inaction*

In spite of the gathering storm in the Autumn 2009, the newly elected Government was far from being determined to achieve immediate fiscal consolidation, constrained as it was by its pre-electoral rhetoric that *"money exists"* and its ideological aversion to controlling trade union demands in public enterprises. Trapped in such unrealistic mentalities, the December Budget for 2010 surprised everybody by including an *expansion* of public expenditure and completely *excluding* privatizations, rather than the other way around. Seeing that no appropriate action had been taken to deal with the situation, rating agencies downgraded the economy, this sparked massive credit default swaps in international markets and the crisis loomed.

The problem Greece faced at that time was an acute shortage of financing for the deficit, not yet one of debt sustainability as it later turned out to be. In this regard, a significant opportunity to diffuse the crisis was missed by the Government and European authorities alike. In order to reduce the risk of spillovers to other markets after the credit crunch in 2008, the ECB had invited private banks of Euro member states to obtain low-cost liquidity by using sovereign bonds from their asset portfolio as collateral securitization; see De Grauwe (2010) for a positive assessment of this policy. As a result of this credit facilitation, yields on Treasury bills remained exceptionally low. But instead of borrowing cheaply in the short term as a means of gaining time to redress the fiscal situation, the Government kept on issuing long maturities despite the escalation of costs. This had dramatic consequences on the perception of the crisis by international markets. Commenting on the cost of confusion, Feldstein (2012) aptly notes that:

> "What started as a concern about a Greek *liquidity problem* – in other words, about the ability of Greece to have the cash to meet its next interest payments – became *a solvency problem*, a fear that Greece would never be able to repay its existing and accumulating debt", (my emphases).

Adding injury to misjugment, the situation was further undermined when the ECB threatened to refuse collateral status for downgraded Greek bonds, hence fuelling fears that domestic liquidity would shrink and precipitating a capital flight from Greek banks. Three months later the rating requirement was dropped for all Eurozone

countries, but the damage was no longer reversible. In early 2010, borrowing costs started to increase for both short and long term maturities, Greece had become a front page story worldwide and the count-down began. Despite the belated ECB generosity, the Government was financially exhausted and in April 2010 sought a bailout.

*The role of external deficits*

The global financial crisis in 2008 revealed that countries with sizeable Current Account deficits are vulnerable to international market pressures because they risk having a "sudden stoppage" of liquidity. Recent studies show that highly indebted EMU countries with large external deficits are found to experience the highest sovereign bond yield spreads. Along this line, Krugman (2011) recently suggested that the crisis in the southern Eurozone countries had rather little to do with fiscal imbalances and rather more to do with the sudden shortage of capital inflows required to finance their huge external deficits.

This explains why immediately after the crisis sovereign spreads peaked mainly in economies with large external imbalances, such as Ireland, Spain, Portugal and the Baltic countries, which were under little or no pressure from fiscal deficits; for a discussion of the effects of credit crunch in emerging markets with large Current Account deficits see Shelburne (2008). In contrast, countries with substantially higher debt burdens but without external imbalances, such as Belgium and Italy, experienced only a small increase in their borrowing costs at that time.

Greece happened to have a dismal record on both deficits and its exposure to the international credit stoppage was soon transplanted into a debt crisis. The Current Account went in free-fall after 2006 when three factors intensified: domestic credit expansion accelerated and disposable incomes were enhanced by the tax cuts, while capital inflows from the Greek shipping sector peaked as a result of the global glut and the huge rise in Chinese freight. The external deficit exceeded 14% of GDP in 2007 and 2008 and still no warning was voiced by any authority, domestic or European. In fact quite the opposite happened: Responding to pleas of car dealers, the Greek Government decided to reduce surcharges on imported vehicles in an attempt to revive the market, while other fellow Governments – at least those from car-making countries – failed to notice the pro-cyclical character of the measure. Replicating history back in 1989, the unfortunate act to facilitate car purchases in order to favour particular groups caused again a significant deterioration of both the external and the public deficit. Additionally, nobody missed the signalling about the true priorities of the Government and the pre-electoral spree followed as described above.

## 5. Two important policy facts

Two stylized facts emerge from the historical account of fiscal developments in Greece. One is the fact that in periods of recession counter-cyclical activism usually takes the form of increased consumption, not public investment and this has detrimental effects on public and external deficits without contributing to higher growth. Another recurring characteristic is the propensity of Governments to increase public spending and to tolerate lower revenues in election years.

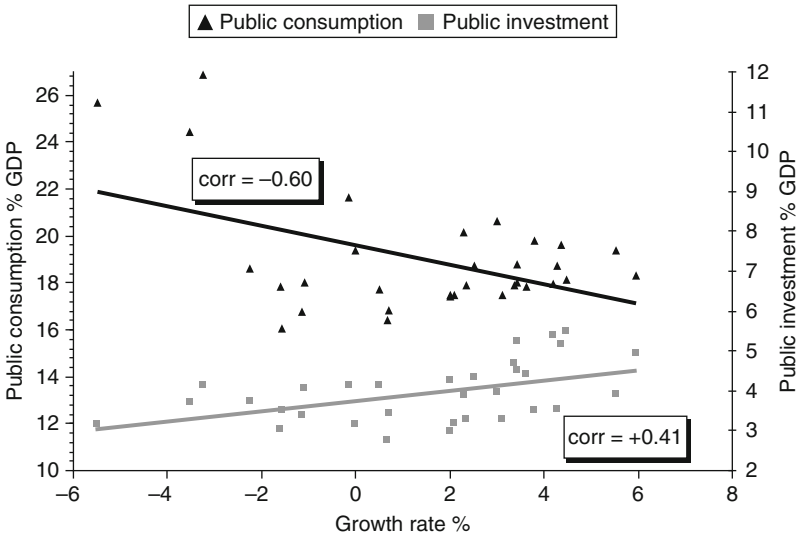



**Figure 7** Growth rate correlations with public consumption (Lhs) and public investment (Rhs) expressed as percent of GDP. Source: Government budget, various editions.

*Cyclicality of public spending*

As an indication of how the two main components of Government spending behave over the economic cycle, public consumption and public investment expressed as proportions of GDP are correlated with the growth rate; see Fig. 7. Public consumption is found to have a strong negative correlation with growth rates, suggesting a countercyclical pattern. This finding implies that periods of economic downturn are likely to be associated with higher public consumption due to increased benefits and programs to contain unemployment. In a situation of fixed public employment and nominal wage resistance, public consumption is expected to rise further relative to GDP.

On the other hand public investment shows a strong positive correlation with growth. This implies that, in a downturn, public investment is likely to fall, thus hindering the resumption of growth and causing more recession in the economy.

A clear manifestation of such behaviour over the cycle took place in recent years. With recession deepening year after year, the Government rather than curtailing the public sector found it more expedient to cut public investment in order to control the deficit. As a result, recession was made worse.

*Electoral cycles*

The Greek economy was often subject to the electoral cycle, as incumbent Governments tried to appeal to voters by a variety of opportunistic policies, thus inflicting non-trivial fiscal losses. Practices included extra appointments of party affiliates, grants to favorable groups and allocation of petty projects to local constituencies, all of which affect current or next period expenditure.

It can readily be seen from Fig. 2 that spending rises during the election years in the 1980s and, as deficits widened, the economy had to enter a period of stabilization that

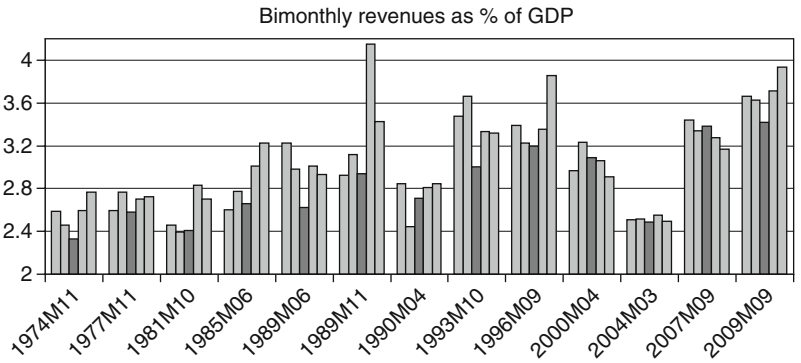



**Figure 8** Comparison of bimonthly tax revenues in pre-election periods. Revenues are calculated for the period of two months before each election as % of annual GDP. Each election year (N) is denoted by black and compared with revenues collected over the same period during the previous (N – 2, N – 1) and the following years (N+1, N+2) denoted by grey. Frequency is bimonthly to account for the fact that the pre-election period lasts for 30–40 days, thus it extends over the prior as well as the poll month. Data are not seasonally adjusted, thus they reflect within year variations. Source: Skouras and Christodoulakis (2011) where further details are available.

was usually terminated before the next election. During the debt escalation in 1980–93 there were four stabilization programs and ten Finance Ministers – usually one to pursue the program and then a successor to denounce it and prepare for the next period of spending rise. Though the electoral cycle subsided in the period before and after EMU membership, it returned full-steam in the elections of 2009.

Apart from direct actions on the expenditure side, the empirical evidence suggests that slacker tax auditing around elections causes further fiscal deterioration. An extensive investigation by Skouras and Christodoulakis (2011) found that flaws in tax collection arise either as a result of deliberate relaxation of audits as a signal to political supporters or as an indirect consequence of the slackness prevailing in public administration around elections.

Considering that a typical pre-election period has duration of circa 40 days, Fig. 8 compares the revenue in the two months of the election period in each electoral year with the same two months in adjacent years. Simple inspection shows that in most of the elections held between 1974 and 2009, average bimonthly revenues expressed as percent of GDP were lower than the average of the respective figures in the two adjacent years, (with only two slight exceptions in 2000 that coincided with the entry to EMU and 2007 because it is compared with another – and a lot worse – electoral period in 2009). In the same study it is estimated that pre-electoral misgovernance causes a loss in revenues equal to 0.18% of GDP in each election year. For the 13 elections taken place in the period 1974–2009, this amounts to more than 5 billion Euros at 2010 prices.

## 6. An ex post assessment of the Memorandum

EU authorities seemed to be unprepared to react promptly and concertedly to the Greek problem and undertook action only when they recognized the risks it posed for the banking systems of other European states. After difficult negotiations, a joint loan

of Euro 110 bn was finally agreed in May 2010 by the EU and the IMF to substitute for inaccessible market borrowing. The condition was that Greece follows a Memorandum of fiscal adjustments to stabilize the deficit and structural reforms to restore competitiveness. More details are given in the Appendix. In the eventuality of success, Greece would be ready to tap markets in 2012 and then follow a path of lowering deficits and higher growth. More than two years after implementation, the record remains poor and the economy is fiercely contracting. An explanation is attempted below.

*The failure in fiscal adjustment*

The decline of revenues as a share to GDP after 2007 and the collapse of the collection mechanism in 2009 in particular were instrumental for the explosion of public debt and deficit thereafter. Strangely enough, no serious effort was undertaken to remedy the situation after the elections. The ministerial post in the Inland Revenue remained empty for more than a year and two top executives resigned in protest that their proposals to beat tax evasion were turned down. The Government opted for an increase in the VAT rate from 19% to 23% in the spring 2010 and, as a result, CPI inflation jumped to 4.5%, further cutting purchasing power amid recession. The only result was that activity was reduced and revenues did not rise.

The Government continued to act in a positive feedback loop, with lower revenues prompting higher taxation and this in turn causing further evasion. Unable to raise efficiency and under pressure to collect revenues, it imposed a heavy increase in fuel tax, substantial consumption surcharges and finally a lump-sum tax in exchange for settling previous arrears. Once again tax revenues ended up far below the target in a typical manifestation of elementary Laffer-curve predictions.

Only by the end of 2011 was it recognized that further tax measures are no longer viable and attention should shift on collection efficiency. In its assessment of progress, the European Commission task force warned that "... tax and expenditure measures ... substantially compress the households' disposable income and significantly tighten their liquidity constraints", (European Commission, 2011, p. 2).

But that was no more than a void warning, because at the same time the Government was forced by the very same task force to retroactively raise the tax rate on the self-employed and impose a new levy on property in order to make up for the falling revenues.

Regarding public expenditure, a more optimistic picture emerged but at a huge cost in terms of growth and efficiency. Soon after the elections, the Government made clear signals that it had no real intention of containing the oversized public sector. Numerous appointments that were made before elections through a highly disputed process were nevertheless approved by the new incumbent, and a widely publicized operation to abolish and merge outdated public entities has made no real progress, to date. A novel scheme to push older staff onto a stand-by status with a fraction of their salary misfired as it was soon discovered that most on the list were exploiting the incentives of the system for an early retirement. After the fiasco the Government announced a lengthy process of evaluation in the public sector as a precondition for staff redundancies, but without setting a time limit it proved to be only an excuse to avoid actual decisionmaking.

In the absence of any structural adjustment in the public sector, the reduction of spending was achieved by imposing universal cuts in salaries and this led to widespread shirking practices. Another unusual tool for keeping expenditure low was to cut the budgetary co-financing of the European Community Support Framework, thus reducing public investment at a time when it was mostly needed to induce some growth in the economy. After the Decision by the European summit in July 2011, Greece was freed from the co-financing obligation, but when the new practice started to be implemented at the end of 2011 it was already too late to rectify the damage done to economic activity.

*The limits of structural adjustment*

In order to rebalance the economy onto a more competitive path, the Memorandum agreement envisaged a long list of structural reforms ranging from reforming the social security system to removing closed-shop vocational practices, and from cutting red tape to liberalizing the licensing process for lorry and taxi drivers. The pension reforms initially succeeded in harnessing the deficits in the social security funds, but soon they reappeared when a wave of retirement took place in anticipation of imposing further age extensions in the future.

Most of the reforms were either abandoned or backfired. For example, the opening-up of lorry licenses failed to reduce transportation costs and enhance competitiveness in practice despite the severity of clashes with trade union hardliners. The reason for this was that insiders took advantage of a two year postponement and decided to maximize rent-seeking by withdrawing previous price concessions. Besides, the economic gloom was thwarting potential investors by making the upfront cost of setting up a new business too high.

A similar attempt to open-up the taxi licensing system was abandoned after a protracted clash with insiders in the summer 2011 that seriously damaged tourism in its period of peak. In other professions, such as lawyers and pharmacists, there was only a token liberalization without any reduction in consumer prices. Recognizing this failure, the new conditionality program imposed a regressive mechanism with the aim of reducing the overall profit margin to below 15%, (see Memorandum II, 2012, para 2.8, *"Pricing of medicines"*). The results of this are still to be seen.

Seeing that the structural adjustment program was derailed, the Memorandum sought for alternatives. To enhance competitiveness in the labour market, liberalization measures extended part-time employment, imposed wage cuts across the board and removed collective bargaining agreements. Despite lowering labour costs by 12%, enterprises were overwhelmed by recession and unemployment became rampant, exceeding 17% of the labour force by the end of 2011. As with the positive feedback mechanism on the tax front, the rise in unemployment invited a new round of wage cuts in the private sector, shrinking further disposable income and fuelling new waves of social protest. Though the IMF mission in the autumn 2011 was explicit that "accelerated private sector adjustment ... would likely lead to a downward spiral of fiscal austerity, falling incomes and depressed sentiment", it nevertheless urged for further such measures in order to achieve a "... critical mass of reforms needed to transform the investment climate", (IMF, 2011). Bringing-up some growth to the real economy is still not a top priority for the program overseers.

*The failure of privatizations*

The failure of the privatization program is worth commenting on, as it reveals an unusual combination of strong rhetoric in theory with complete apathy in practice. Immediately after the elections in 2009, the Government showed that it had no intention of curbing the wider public sector. Its lack of resolve to tackle the excessive demands of public trade unions was made manifest in a dispute with a newly arrived investor in the Piraeus Port Company. The Government succumbed to paying enormous compensation for early retirement as a condition that the investment goes ahead. No privatization target was included in the 2010 Budget and none was actually implemented.

Thus it was viewed as a major shift of policy when the Government agreed in March 2011 to adopt a large-scale privatization plan of Euro 50 bn during the period 2011–2015, or roughly 4% of GDP per annum. The plan included extensive sales of public real-estate, privatizations of public enterprises in the energy sector and private partnerships in the operation of airports and ports throughout Greece.

After months of procrastination a market-friendly Privatization Fund was finally set up to replace the ineffectual authority that was in charge before, but its determination was this time hindered by adverse market conditions. With asset prices falling to abysmal levels, privatizations would be probably embarrassing in political terms and inadequate in terms of revenue, but in practice there was no real demand, as capital flight continued to be fuelled by fears of abandoning the Eurozone and funds from abroad were not coming for the same reason. Despite initial ambitions, the program achieved little in 2011, selling only an option on Greek Telecom, future rights to the National Lottery and publishing a preliminary tender for the re-development of the old Athens airport. In 2012 the program was downscaled to a meager Euro 2.8 bn, just a quarter of the amount initially announced.

## 7. The new Memorandum conditionalities and ways-out of the crisis

Faced with a deepening recession and a failure to produce fiscal surpluses sufficient to guarantee the sustainability of Greek debt, the European Union intervened twice to revise the terms of the Memorandum. In the first major intervention in July 2011, the amount of aid was increased substantially by Euro 130 bn and repayment was extended over a longer period of time. To implement the Private Sector Involvement (PSI) in debt restructuring, a cut of 21% of the nominal value of Greek bonds and re-profiling of maturities was decided upon with the tacit agreement of major European banks.

Crucially, the EU authorities this time fully recognized the perils of recession and allowed Greece to withdraw a total amount of Euro 17 bn from Structural Funds without applying the fiscal brake of national co-financing. The plan looked powerful, except for the typical implementation lags. The Agreement was only voted through by all member-state Parliaments only in late September 2011 and the release of structural funds was approved by European Parliament in late November. Participation in the PSI had reached only 70% of institutional holders amid speculation that post-agreement buyers of Greek debt from the heavily discounted secondary market were expecting a huge profit through their offer to cut it!

Thus, a new intervention looked inevitable and in October 2011 a revised restructuring (the so called PSI+) was authorized, envisaging cuts of 50% of nominal bond value that would eventually reduce Greek debt by Euro 120 bn and would allow it to be stabilized at around 120% of GDP by year 2020. In exchange, Greece would undertake further fiscal cuts of around 6% of GDP. Greek bonds held by public institutions (i.e. by the bail-out providers) will be fully honored, though social security funds and domestic public entities were forced to participate in the scheme.

The agreement was hailed as the definite solution to the debt conundrum, but euphoria turned sour a few days later when the Greek Government surprised everybody by seeking a referendum for its approval. Many feared that the outcome could in all probability be negative as an expression of current misgivings, and this would be quickly interpreted as opting for exiting the Eurozone. In the ensuing *furore*, the decision was annulled, the Prime Minister resigned and a coalition Government was formed in November 2011 to implement the restructuring of debt and negotiate the terms for the new round of EU-IMF loans. The Government acted quickly and concluded the PSI agreement in May 2012, but the extra fiscal package was not finalized as new elections were called for by the coalition partners, anxious to refresh their mandate before the political cost of further cuts becomes too inhibitive for them to remain in power. A caretaker Government followed and pre-electoral inaction adjourned privatizations so as not to excite opposition from the unions, and made collection authorities to slacken the processing of income tax statements so as not to infuriate the voters with an increased tax burden. Revenues dropped significantly, sadly confirming the cycle described in Section 4.

The situation was further aggravated after the first election round proved inconclusive, with the two mainstream parties saw their share of the vote collapse from around 80% of the electorate in 2009 to just one third of the total in 2012, while leftwing and rightwing parties, with a strong rhetoric against the bail-out agreements, saw their share of the vote soar; this prompted a new wave of speculation that Greece is likely to exit the Eurozone and capital flight drained significant amounts from the Greek banking system.

The second round was polarized by the Euro dilemma, thus helping to increase mobilization of voters and finally resulting in a tripartite coalition that vowed to take all necessary measures to safeguard Greece in the Eurozone.

Though the new Government exhibits some of the weaknesses typical of party coalitions, it strives to persuade domestic and European opinion that it means business. As a signal, it reaffirmed that privatization of several public companies will go ahead and announced that the state will abolish minimum holding rights in utilities as an incentive to potential buyers. The Agricultural Bank was swiftly privatized and other state-owned credit organizations are to follow suit. The fiscal package is expected to be voted on in Autumn 2012, though a new round of social tensions and political breakaways cannot be ruled out.

In the meantime, two factors have adversely affected the situation of Greece versus the Eurozone and the bail-out authorities: On the European front, the market pressure is currently directed towards the economies of Spain and Italy, and the concomitant threat to the very existence of the Euro has pushed the Greek problem to the sidelines.

Although European authorities responded to the challenge by designing a new defense mechanism in the banking system and the European Central Bank decided to intervene in the bond markets to stave off speculative attacks against member-states, European public opinion is characterized by a "*rescue fatigue*" and appears to be increasingly hostile towards granting any further support for Greece. However, Greece itself is overstressed: unemployment has rocketed above 24% of the labour force; recession has further deepened with the contraction of real GDP expected to exceed 7% in 2012, an awesome deterioration from milder estimates just a few months ago.

If this trend continues, debt stabilization will be jeopardized and a new cycle of futility and desperation will emerge. Since the dynamics of the debt-to-GDP ratio are sensitive to the prospects of growth, it is worth examining alternative debt paths that correspond to lower and higher growth profiles.

*Three alternative scenarios: Walking on a tight rope*

The alternatives revolve around a medium growth path, as has been calculated by European authorities in March (EC, 2012, Table p 30) to ensure sustainability of public debt. This is based on real growth resuming in 2014 and staying above 2% thereafter, while inflation as measured by the GDP deflator is consistently kept below 2%. A primary budget surplus is achieved in 2013 and stays above 4% of GDP until at least 2020. Privatizations are to follow the revised schedule shown in Fig. 9, and the cost of borrowing is set at 4% per annum.

The lower and higher growth scenarios are devised by assuming two nominal growth paths cut or augmented by two percentage units respectively, as depicted in Fig. 10. All other assumptions remain the same as in the EC scenario to facilitate comparisons. Debt profiles are shown in Fig. 11.

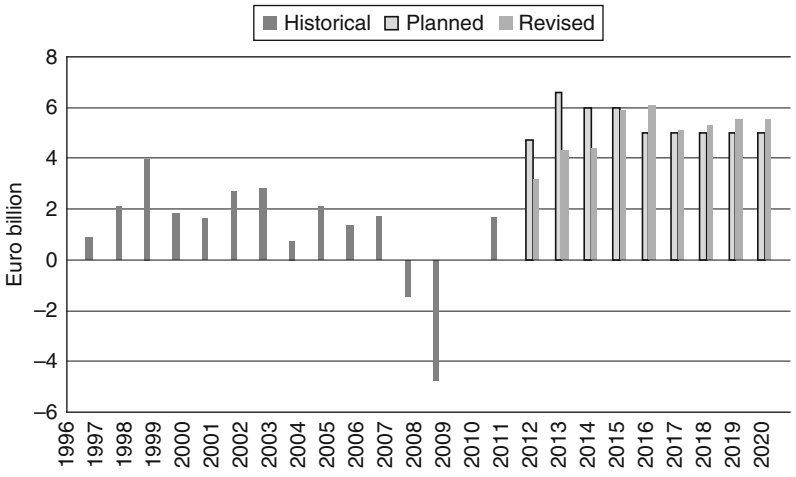



**Figure 9** Proceeds from privatizations. Note: Proceeds are net of capitalizations in state-owned enterprises. For 2008, 2009 and 2010 figures of proceeds are net of bank shares purchases, thus the negative sign. Source: Privatization Report, Ministry of Finance, 2008. Data for 1996 and 1997 are taken from Budget Reports. Planned figures were set in May 2011, but then they were revised in 2012 (Memorandum II, para 2.1).

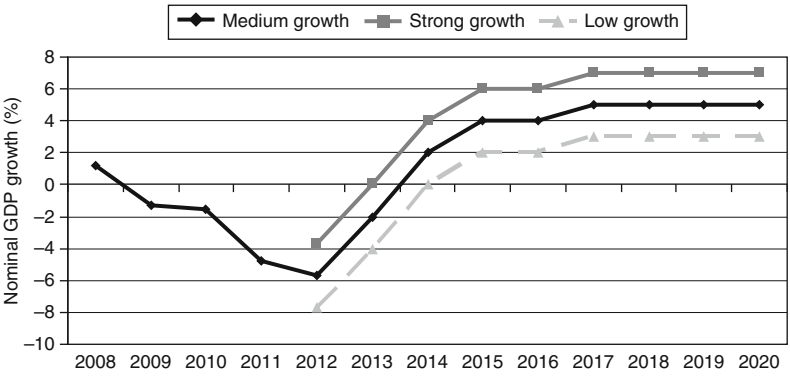



**Figure 10** The three alternatives for recovery. Growth rates of nominal GDP are taken as the sum of real growth and the projected rates of GDP deflator. The medium growth rate is taken from European Economy (2012, Table p. 30), The strong and low growth profiles are obtained by simply assuming plus and minus two percentage units per year over the medium path.

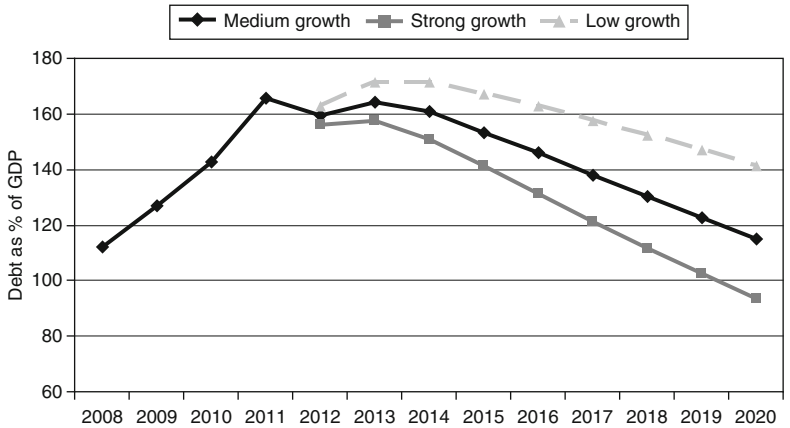



**Figure 11** Alternative paths for public debt as % of GDP. Note: Each scenario corresponds to an assumption of nominal growth in 2012–2020 shown in Fig. 10. All other assumptions remain the same. The medium growth scenario replicates EC (2012).

The low-growth case leads to a debt to output ratio in 2020 above the level it had in 2010 when the country initially asked for a bailout. Most probably, the economy will collapse before the end of the period as generating primary budget surpluses of above 4% of GDP or collecting privatization proceeds of more than 2% of GDP for six consecutive years is utterly unrealistic under anemic growth. The low growth scenario is not out of context though, and, for a start, it replicates what actually happened in 2012 with real growth rate plummeting at $-7\%$ rather than the $-4.7\%$ rate envisaged in the official scenario. If this gloomy trend continues, markets will pick up the conundrum and the situation will soon get out of control. There is no political force in Greece eager to undertake new painful cuts on top of the current ones, and the country will be forced to abandon the stabilization program. European Governments - unable to ignore

indignant public opinion – will insist on no more bailout aid without honoring previous obligations, and then Greece will be left impotent and unwilling to continue any further. End of game.

The higher growth scenario, on the other hand, leads to a debt to output constantly declining and reaching a level of around 95% at the end of the period, substantially below the medium scenario. The growth path is not unrealistic and real growth rates have just to be close to those that prevailed in the previous decade, though now based on deep market reforms and without the fiscal extravaganza. Other assumptions, such as those of substantial primary surpluses and uninterrupted privatizations, also become more realistic with higher growth. Sustainability may be further assisted if the bail-out funds currently allocated for the recapitalization of Greek banks are taken out of public debt accounts and become liabilities of the new banking authority that is scheduled to operate next year on a Eurozone-wide basis. This will mean a further reduction of the debt to output ratio by more than twenty-five percentage units and will firmly anchor Greece in the Eurozone.

*Is exit from the Euro an option?*

The crisis in Greece had profound ramifications for the Eurozone, both in political as well as in economic terms. In the Euro area, Greece is routinely considered not only as devouring European taxpayers, but also as the habitual wrongdoer especially when compared with the other two countries (Ireland and Portugal) which are undergoing similar adjustment programs with more efficacy. In such a politically unyielding and increasingly suspicious framework, a Greek exit from the Eurozone started to attract attention both at home and abroad.

Though complications and costs that would ensue in the banking sector will be enormous, the exit of Greece could prove opportunistically attractive to some European politicians who get angrier every time a new round of aid is discussed. However, they overlook the fact that a Greek exit would reverberate around other states and lead to an aggravation of the crisis; for how contagion will spread see Vehrkamp (2011). It may also serve as the convenient argument for consolidating and enforcing a two-tier model of Economic Governance, as has been advocated before the creation of EMU (e.g. Bayoumi and Eichengreen, 1992) and is recently suggested again by commentators and politicians betting on the "*Grexit scenario*" and assuming that other countries may follow suit. Based on an inner core of surplus economies in the north and a weaker periphery in the south, competitiveness in this model will be restored through the so called "*internal devaluation*" of labour costs, thus perpetuating the gap that is already widening between the Eurozone countries; for a description of divergences within the common currency see Christodoulakis (2009).

For Greece, exit would trigger a prolonged economic catastrophe. As the entire Greek debt will remain denominated in Euros, the rapid depreciation of the new national currency will make its servicing unbearable and the next move will be a disorderly default. Isolation from international markets would drive investors even further away, while the financial panic would drain domestic liquidity at a massive scale. The creditor countries of the EU would start demanding repayment of their aid loans, and this would soon deprive Greece of its claim on the EU cohesion funds. Tensions are

likely to produce further conflicts with EU agencies and the pressure to consider complete disengagement from the European Union will gain momentum both domestically and abroad.

*Stay in the Eurozone and grow more*

The cost would be so immense that the single available option for Greece is to complete the fiscal adjustment and become reintegrated into the Eurozone as a normal partner. This requires Greece to undertake concrete actions that produce visible results within a short timeframe, so that society becomes more confident to pursue further reforms. Some policy suggestions for this direction are as follows:

First, Greece needs to acquire credibility while also being properly understood abroad. The continuing fiscal shortfall is easily translated as reluctance, causing continual friction with the European Union and demands for a new battery of austerity measures. To escape this cycle, Greece must adopt a front-loaded policy as a matter of urgency to achieve key fiscal targets quickly and to change the impression of being a tactical waverer. This seems to be the line adopted by the new Government. If Greece succeeds in this front-loaded policy, it may be in a position to revise some of the pressing – although so far unattainable – schedules and ensure greater social approval and tolerance. To ensure that there will be no spending spree in future elections, the best option for Greece is to adopt a constitutional amendment on debt and deficit ceilings, just as Spain did in 2011, alleviating market pressures, at least for the time being.

Second, Greece desperately needs a fast-track policy for exiting the long recession. An amount of Euro 17 billion could be disbursed and routed immediately to support major infrastructural projects and private investment in export-oriented companies. The growth-bazooka should then be followed by structural reforms and privatizations that can attract significant private investment as market sentiment is restored. In addition, instilling growth will help to control the debt dynamics and reduce public deficits without ever-rising taxes that thwart private investment and make economic recovery and sustainability even more unattainable. Feldstein (2012) leaves no doubt about the mechanics of stabilisation when he warns that "(t)o achieve a sustainable path, Greece must start reducing the ratio of its national debt to GDP. *This will be virtually impossible as long as Greece's real GDP is declining*", (my emphasis).

The inevitability of the above thesis cannot be ignored anymore. Nor can it be circumvented by sermons on the necessity of front-loaded reforms on the assumption that will automatically restore growth and competitiveness. In fact, the need for further growth spreads fast to other countries either in or outside the Eurozone; for example, the UK Government recently decided to inject BPS 50 bn on infrastructural projects to speed economic recovery and one just hopes that the Eurozone will be equally responsive to the need.

## 8. Conclusions

Exactly three decades after becoming a fully-fledged member of the European Union and ten years after joining the Eurozone, Greece sought a bail-out agreement in 2010 to avoid bankruptcy. A long history of stabilization programs proved incapable of

achieving a lasting fiscal correction and adequately raising competitiveness, as fundamental weaknesses in the economic and political system continue to play a corrosive role. The oversized public sector and the frequent indulgence in pre-electoral spending sprees in exchange for political support led to protracted fiscal deficits and the accumulation of a large public debt. Equally, the chronic deterrence of productive investment by a multitude of regulatory inefficiencies resulted in a thin tradeable sector and large Current Account deficits. The economy remains vulnerable to political developments which are often dictated by short-term partisan considerations with far reaching fiscal implications. This explains why, in spite of substantial reforms taking place over the last two decades and achieving high growth rates, EMU participation and moderate debt stabilization, the situation went once more out of control.

Regarding the current crisis, the article described how prolonged external and fiscal deficits were allowed to reach uncontrollable levels and, in the aftermath of the credit crunch, led to a further escalation of debt and the subsequent bail-out. Two and a half years later, fiscal consolidation is still far from being sustainable in spite of augmenting the bail-out loans and implementing a substantial debt reduction on private holders.

The economy has contracted by nearly 20% since 2008, social tensions are multiplying and the future of Greece in the Eurozone is in jeopardy. Some consider such an outcome as a due punishment for past excesses, while others see it as an escape from further unemployment and recession. The article finds both angles of view as illusory, and argues that the only viable way out of the current crisis is to restore growth and then adopt a realistic plan for privatizations and reforms. The lesson of the past two years is that deep recession will otherwise continue to hinder any existing possibility for exiting the crisis. Greece, and other Eurozone countries too, are desperate for a *"corridor of confidence"*, to use Keynes' famous phrase, to put things in order before it is too late.

## Acknowledgement

I have benefited from various comments by V. Sarantides and A. Ntantzopoulos and I am also thankful to participants in a LSE seminar on an earlier version of the paper.

## Disclaimer

Views expressed in this article are solely those of the author, without implicating or representing any other person or organization.

NICOS CHRISTODOULAKIS

*See Also* **European central bank; European Central Bank and monetary policy in the euro area; European cohesion policy; European monetary integration; European monetary union; European Union budget; European Union trade policy; Eurozone crisis.**

### *References*

Bayoumi, T. and Barry Eichengreen, 1992, "Shocking aspects of European Monetary Unification", CEPR Discussion Paper no. 643, May.

Blanchard, O., 2006, "Current Account Deficits in Rich Countries", *IMF* Mundell-Fleming Lecture.

Blanchard, O., and Giavazzi, F., 2002, Current account deficits in the Euro Area: The end of the Feldstein-Horioka puzzle? *Brookings Papers on Economic Activity*, 33, 147–210.

Christodoulakis, N., 1994, "Fiscal Developments in Greece 1980–93: A critical review", *European Economy: Towards Greater Fiscal Discipline*, No. 3, pp. 97–134, Brussels.
Christodoulakis, N., 2009, "Ten Years of EMU: Convergence, Divergence and new priorities", *National Institute Economic Review*, Vol. 208, pp. 86–100, London.
Christodoulakis, N., 2010, "Crisis, Threats and Ways out for the Greek Economy", *Cyprus Economic Policy Review*, Vol. 4, 1, pp. 89–96, June.
Christodoulakis, N., 2012, "Market Reforms in Greece 1990–2008: Domestic limitations and external discipline", in Kalyvas, S. and Pagoulatos, G. (eds) *"Market Reforms in Greece"*, Columbia University Press (forthcoming).
Christodoulakis, N. and Sarantides, V., 2011, "External asymmetries in the euro area and the role of foreign direct investment", *Bank of Greece*, Discussion paper.
De Grauwe, Paul, 2010, "The Greek Crisis and the Future of the Eurozone", *Intereconomics*, No. 2, pp. 89–93.
European Commission, 2009, "Quarterly Report on the Euro area", 8(1), Brussels.
European Economy, 2011, "The Economic Adjustment Programme for Greece: Fifth review", Occasional Papers 87, October.
European Economy, 2012, The Second Adjustment Programme for Greece: Fifth review", Occasional Papers 94, March.
Feldstein Martin, 2012, "The Failure of the Euro: The Little Currency that Couldn't", *Foreign Affairs*, Vol. 91, 1, pp. 105–116.
IMF, 2011, "Greece: Third Review under the Stand-By Arrangement", International Monetary Fund, Country Report No. 11/68, March.
Krugman, Paul (2011), Origins of the Euro Crisis, *blog* "The Conscience of a Liberal," 23 September.
MemorandumII, 2012, "Memorandum of Understanding on Specific Economic Policy Conditionality", 9 February 2012, available at http://www.hellenicparliament.gr
Shelburne R. C., 2008, "Current Account Deficits in European Emerging markets", *UN Discussion Paper,* No. 2008.2
Skouras, S. and Christodoulakis, N., 2011, "Electoral misgovernance cycles: Evidence from wildfires and tax evasion in Greece and elsewhere", LSE, Hellenic Observatory, *GreeSE Paper* No. 47.
Vehrkamp, R., (2011), "Who's next? The Eurozone in an Insolvency Trap", *Bertelsmann Stiftung*, No. 2.

## Appendix: A brief description of the conditionality programs for Greece

The adjustment program for Greece was laid out in three phases. The first Memorandum was signed in May 2010 and aimed at reducing the fiscal deficit to 3% in 2013. Specific measures that were actually implemented included universal cuts in public salaries and all pensions, a rise in VAT from 19% to 23% and similarly in other consumption surcharges, the abolition of collective agreements in favor of firm-level contracts, the lowering of private sector wages by 12% and a reform in the Social Security system. It also included the liberalization of red-tape practices in the transport sector, pharmacists and lawyers, but the outcome was heavily compromised through a series of delays and back offs. Fiscal deficit for 2010 ended up close to 11% of GDP, substantially lower than the horrendous 15.4% in the year before but still away from the initially set target.

Thus, in early 2011 a new round of negotiations resulted in a second round of measures voted by Parliament in June 2011. They included further taxation on past incomes, a lump-sum tax on professionals, further rises in indirect taxes and a new property levy that was imposed two months later. The program demanded the abolition of outdated public entities, the reduction in the number of civil servants and a further

curtailment in their salaries. It also envisaged ambitious privatizations on utilities and public real-estate that could trim down public debt by Euro 50 bn within a four-year period. Fiscal deficit for 2011 is provisionally estimated to be 9.8% of GDP, revealing a major difficulty in further adjustment in the absence of growth.

The third round of adjustments was voted for in February 2012 as Memorandum II. (For the full text see "Memorandum of Understanding on Specific Economic Policy Conditionality", 9 February 2012, available at http://www.hellenicparliament.gr).

This time it was approved by the two major parties, but only after a line-up was imposed to avoid desertions and rising internal protest. Measures included a reduction of minimum wages in the private sector by 22%, an additional cut by 10% to new entrants as a means to beat youth unemployment, 15% cuts in various pensions, the abolition of several tax credits and explicit targets for cutting employment and entities in the wider public sector. Policies will start to be implemented in the final quarter of 2012.

# Great Depression

## Magnitude

Figure 1 shows the fall in industrial production during the Great Depression in the four largest national economies at that date. Industrial production declined by almost half in the United States and Germany. It fell more slowly and continuously in France, and paused rather than fell in Great Britain. National incomes did not fall as far as industrial production since services did not contract as much, but they decreased sharply; real per-capita GNP in the United States fell by one-third. National experiences in the depression varied greatly, but very few countries in the world escaped the economic hardship of the 1930s. One task for any account of the Great Depression is to explain its worldwide impact.

Figure 2 shows the fall in wholesale prices for the same four countries. Prices fell at the same time as production, by the same amount or more. Unemployment grew dramatically in almost all countries. Rates for the four largest economies are shown in Table 1. Only in the United Kingdom were unemployment rates approximately as high in the 1920s as in the 1930s, due to depressed conditions in Britain during the 1920s and a mild depression in the 1930s. Other countries for which we have data fit the more common pattern of higher unemployment in the 1930s.

Unemployment meant distress in the 1930s, most visible in Europe and North America. Diets in Europe became very monotonous despite the presence of home-grown vegetables in some areas. Families ate meat only rarely, starches were the basis of most diets, and sugar frequently was replaced by cheaper saccharine. Even this poor diet consumed almost all the family income. Families with children bought milk, most families bought coal for heat, but there was little money left over for clothes and other expenses. Shoes in particular were a problem. Families typically could not afford to replace shoes that had worn out, and so they were patched and patched again. Some families even restricted the activities of their children to save the wear and tear on their shoes.

While spending was channelled into food, and food into bread and coffee, personal travel was reduced to journeys to local neighbourhoods and villages. Trips to towns and town centres had been increasing during the 1920s, to go to the theatre, do Christmas shopping, or attend school. With unemployment, the money to undertake these journeys vanished. Even tram and train fares became a burden, and people relied more heavily on their bicycles. The isolation of rural villages, alleviated by the railways and the prosperity after the First World War, reappeared in the depression.

Unemployed men were exceedingly idle; an increase of apathy reduced all forms of recreational activity. Men passed their time doing essentially nothing; when asked, they could not even recall what they had done during the day. They sat around the house, went for walks – walking slowly – or played cards and chess. Most men went to bed early; there simply was no reason to stay awake. Women were far more active.

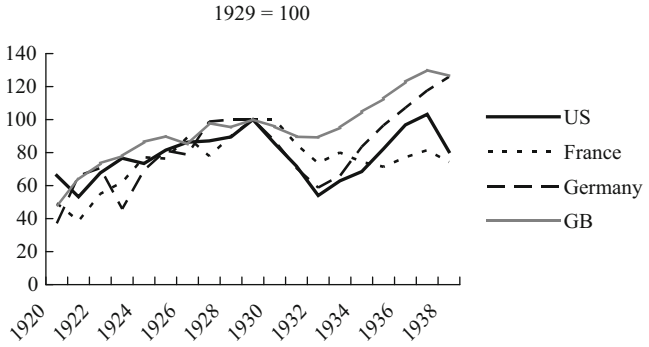



**Figure 1** Industrial production, 1920–1939. *Source*: Temin (1989).

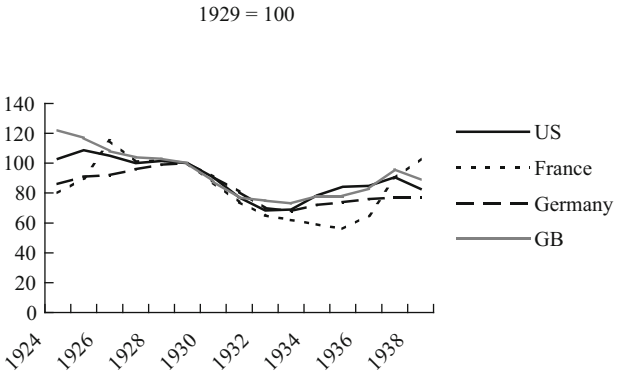



**Figure 2** Wholesale prices, 1924–1939. *Source*: Temin (1989).

They spent time cooking, mending clothes to make them last longer, and managing their budgets. Men contributed less to the running of the household than before, sometimes not even turning up on time for meals, and women had the full responsibility for maintaining the household. Even though women previously had struggled to complete their housework after working, they uniformly would have preferred being back at work.

Sociologists observed that most European unemployed families were resigned to their condition. Such families were hanging on, preserving as much of their life and family as they could on their meagre budgets. All their activity was dedicated to getting by; no thought was given to the future. Some families still planned as before, but others collapsed entirely into mental and physical neglect and conflict.

Beyond Europe and North America, the story of destitution was the same, although the workers' issues typically were more related to physical survival. Rural families in Asia and Africa suffered from the low prices that their crops received in the depressed world markets. They do not seem to have lapsed into idleness like unemployed urban workers, but rather continued to produce crops in the hope of increasing their incomes. Consumers in India, no longer able to afford imported cloth, gave a boost to domestic,

**Table 1** *Industrial unemployment rates, 1921–1938.*

| Country | 1921–29 | 1930–38 | Ratio |
|---|---|---|---|
| France | 3.8 | 10.2 | 2.7 |
| Germany | 9.2 | 21.8 | 2.4 |
| United Kingdom | 12 | 15.4 | 1.3 |
| United States | 7.9 | 26.1 | 3.3 |

*Source*: Temin (1989).

beleaguered handloom weavers. Workers in Latin America retreated from cities and organized agriculture back into the countryside, and little is known of their living conditions. Latin American governments divided into active states that tried to insulate their economies from the outside world and passive states that waited for better times. Governments were surprisingly stable under this economic stress, but they collapsed in some countries, ranging from Germany to Burma.

## Analysis

The first question to ask about this contraction is whether the shocks that produced it were demand or supply shocks. The simultaneous fall in production and prices indicates that the shocks were demand shocks, that the economies of the world were moving down along their upward-sloping aggregate supply curves in response to downward shifts of aggregate demand curves. The apathetic reaction to unemployment in the Great Depression confirms the hypothesis that the depression was due to a demand shock. Had it been due to a supply shock, families would have been unemployed by choice, happy with their extra leisure. The psychological depression also put great strains on the social structure, and even the political structure in some countries. It was in soil such as this that the noxious weed of National Socialism grew in Germany.

A second question about the Great Depression is how so many countries could have had negative demand shocks at the same time. The answer is that all these countries were adopting deflationary policies according to the dictates of the gold standard. The gold standard was characterized by the free flow of gold between individuals and countries, the maintenance of fixed values of national currencies in terms of gold and therefore each other, and the absence of an international coordinating or lending organization such as the International Monetary Fund. Under these conditions, the adjustment mechanism for a deficit country was deflation rather than devaluation – that is, a change in domestic prices instead of a change in the exchange rate. Lowering prices and possibly production as well would reduce imports and increase exports, improving the balance of trade and attracting gold or foreign exchange. (This is the price-specie-flow mechanism first outlined by Hume in 1752.)

A recession began at the end of the 1920s in the United States and Germany. Both countries began to contract economically, at least partly as a result of central bank pressure. The initial downturns appear to be independent in each country, but their economies were connected, and it is hard to be sure about this. In any case, it was gold-standard policies that transformed the downturn into the Great Depression and pulled the rest of the world down. The choice of deflation over devaluation was the

most important factor determining the depth of the Great Depression. The choice was seen clearly and supported by contemporaries in all industrial countries who insisted that the way out of depression was to cut wages and thereby the costs of production and the prices of goods and services. Devaluation was not a respectable option.

Less developed countries were less likely to be on the gold standard than those in Europe or North America. They suffered from the depression nonetheless because of their ties to gold standard countries. As industrial countries reduced their demand for imports, exports from less developed countries declined. As industrial countries stopped exporting capital, less developed countries found their balance of payments deteriorating further. A few countries, such as Spain and Japan, devalued their currencies early and avoided the worst of the depression, but many more countries were not in a position to do this or where it would have had a large effect.

A third question that economists ask about the Great Depression is why the fall in demand was not absorbed entirely in falling prices. In other words, why did prices not fall more and production less than shown in Figures 1 and 2? The relative stability of wages caused production and employment to fall; falling prices and wages did not absorb the full brunt of the fall in demand. Falling prices also put pressure on financial institutions, whose failures reduced production as well.

Governments and central banks could not easily deflate their economies in the aftermath of the First World War. Workers, who had borne the burdens of international stability mutely in the past, expected and even demanded a voice in policy after their sacrifices during that war. The inability of economic policymakers to force wages down rapidly created the conditions for the Great Depression. The political strains generated by attempts to lower wages caused investors to fear for the stability of the gold standard even as policymakers struggled to maintain it. One reason the gold standard worked well before 1914 was that labour had no voice. The spread of democracy both cast doubt on the monetary authorities' commitment to the gold standard and reduced price flexibility.

Banks failed right and left in the midst of deflation and currency crises. Widespread banking failures were restricted to countries on the gold standard, showing that the strain of the gold standard was the principal cause of bank distress. All banks suffered as economic activity and prices declined, but the diversion of central banks from the support of commercial banks to the defence of the currency made the difference between banks in difficulty and banking crises. The German government took over the country's great banks in June and July 1931; American banks were allowed to fail continuously as economic decline continued. It seems that a slow crisis was more destructive of economic activity than a rapid one, though there are not enough observations to test this hypothesis.

## Narrative

The narrative of the Great Depression properly begins with the First World War. The dislocations of the war and the peace agreements meant that many adjustments had to be made in the international economy. Strains were evident in the immediate aftermath of the war, resulting in hyperinflations in several countries, most notably Germany.

The response was to return to the gold standard in the mid-1920s in the hopes of regaining pre-war stability. Alas, the cure proved worse than the disease.

Federal Reserve policy became contractionary at the start of 1928 in order to combat speculation in the New York stock market and to arrest a gold outflow begun in part by previous financial ease. The gold outflow was a prominent determinant of the policy change, even though it was tiny relative to US reserves. The Federal Reserve's primary aim in 1928 and 1929 was to curb speculation on the stock exchange while not depressing the economy. Even though this policy did not impede stock-market speculation, it reduced the rate of growth of monetary aggregates and caused the price level to turn down. The monetary stringency was even tighter than it seems from examining the aggregate stock of money because the demand for money to effect stock-market transactions rose, leaving less for other activities.

The German economy was heavily dependent on imported capital in the 1920s. Popular history regards the capital imports as a necessary offset to Germany's outflow of war reparations payments; they were needed to solve the transfer problem. The reality was quite different. Germany managed to avoid paying reparations by a variety of economic and political manoeuvres that succeeded in postponing its obligations until they could be repudiated entirely. The capital inflow therefore represented a net increase in the resources available to the German economy. The Reichsbank paradoxically worried that this capital inflow was unhealthy and acted to curtail it, sharply reducing the amount of credit available on the German market at the end of the 1920s. The capital flow from the United States to Germany ceased at the end of the 1920s, but the downturn in Germany preceded this fall and derived largely from German economic policies.

At its inception, the Great Depression was transmitted internationally by a gold-standard ideology, a mentality that decreed that external economic relations were primary and that speculation like the booming stock markets in New York and Berlin was dangerous. As the American, British and German economies contracted, they depressed other economies through the mechanism of the gold standard. These countries reduced their imports as they contracted, reducing exports from other countries. They also reduced their capital exports or increased their capital imports in response to the tight credit conditions at the end of the 1920s.

A bad recession turned into the Great Depression in the summer and autumn of 1931. A series of currency crises led both to what we now regard as perverse policy responses and to failures of financial institutions. A warning came in May 1931 when the main bank of Austria, the Credit Anstalt, failed, taking the Austrian schilling with it. This was a preview of things to come, but not a cause of them. The German mark had been under pressure since the German recession began in the late 1920s and the Weimar government began to run increasingly large deficits. They were covered by foreign lending, of which the American Young Plan was the most famous. The Weimar government, however, scared its foreign creditors by a series of statements for domestic consumption about a customs union with Austria and a possible repudiation of First World War reparations. The Reichsbank lost reserves precipitously in late May, and free trading in the mark was suspended in July 1931.

**Table 2** *Standard deviation of changes in 21 gold-standard countries, 1930–1932.*

| Year | Prices | Industrial production |
|---|---|---|
| 1930 | 0.037 | 0.081 |
| 1931 | 0.055 | 0.078 |
| 1932 | 0.090 | 0.123 |
| 1932* | 0.035 | 0.039 |

*Seven countries still on gold in 1932.
*Sources*: Bernanke and James (1991); Temin (1993).

The British government found itself in similar trouble as its deficits followed Germany's. The Bank of England, unwilling to raise the bank rate above six per cent and further depress the domestic economy, abandoned the gold standard, floated the pound, and devalued in September 1931. The Federal Reserve, facing similar problems and adverse speculation, chose to raise its discount rate by 200 basis points in October 1931. This dramatic measure saved the dollar but killed the domestic economy. It was, however, loudly applauded by the American financial community as the correct gold-standard action.

The effects of fixed exchange rates can be seen in a comparison of Figures 1 and 2. Figure 1 shows that industrial production in four major countries declined at quite different rates. Figure 2 shows that the rate of decline in prices in the same four countries was strikingly similar. The fixed exchange rates of the gold standard led to uniform changes in prices even though other factors affected the change in production. The standard deviation of price changes was smaller than the standard deviation of production changes for 21 countries on the gold standard in 1930–2, as shown in Table 2. The standard deviation of price changes was smaller than the standard deviation of changes in the industrial production index in each year, even though the standard deviation of both series rose in 1932 as some countries abandoned gold. The final row of Table 2 shows the standard deviations in 1932 for seven countries that stayed on gold in 1932. Even though data for these countries are indistinguishable from the rest of the sample in 1930 and 1931, they are far more uniform in 1932.

No country on the gold standard, however large, could escape the discipline of this harsh regime in the depression. In almost all cases, deflation was accompanied by depression as declining aggregate demand moved countries down upward-sloping aggregate supply curves. Banking systems in many gold-standard countries collapsed under this deflationary pressure, further reducing economic activity. The Federal Reserve sharply raised the US discount rate in October 1931 in response to a threatened outflow of gold, even though the US economy was contracting rapidly and had massive gold reserves. The primary transmission channel of the Great Depression was the gold standard.

It follows that abandoning the gold standard was the only way to arrest the economic decline. Going off gold severed the connection between the balance of payments and the domestic price level. Countries could lower interest rates or expand production without precipitating a currency crisis. Changes in the exchange rate rather than changes in domestic prices could eliminate differences between the level of domestic

and foreign demand without a painful deflation. Any single devaluation could beggar neighbours under some conditions, but universal devaluation would have increased the value of world gold reserves and allowed worldwide economic expansion.

Great Britain abandoned the gold standard in September 1931 after a speculative attack on the pound prompted by bad budgetary news and by contagion from the German currency crisis of July 1931. Great Britain and the countries that followed Britain off gold were not large enough for their actions to arrest the world decline, and they were criticized at the time for abandoning gold; but the world would have been far better off if others had followed them off gold.

Even in the United States, with its vast economic resources and gold reserves, going off gold was a necessary prerequisite for economic expansion. Great Britain avoided the worst of the Great Depression by going off gold in 1931, as shown in Figure 1. Spain avoided the depression by never being on the gold standard; Japan by a massive devaluation in 1932. At the other extreme, the members of the gold bloc led by France endured contractions that lasted into 1935 and 1936. The single best predictor of the severity of the depression in different countries is how long they stayed on gold. The gold standard was a Midas touch that paralysed the world economy.

Real wages stayed high in countries on the gold standard. Macroeconomic policies to preserve the value of the currency reduced prices faster than wages, and real wages stayed high or even rose. Bank failures also were widespread in gold-standard countries, further depressing production. Both high real wages and bank failures show up as explanatory variables for low incomes around 1935, and the prevalence of financial crises in countries on gold suggests that a counterfactual with more rapid deflation and no devaluations would not have resulted in the maintenance of something close to full employment.

## Complications

The influence of the gold standard determined the spread and the depth of the Great Depression, but the story has many dimensions not captured in this stark description. The literature can be contentious, although apparently competing views may represent elements in a more comprehensive view.

One view of the Great Depression sees it as an American contraction that was transmitted to the rest of the world. In *A Monetary History of the United States, 1867–1960* (1963), Milton Friedman and Anna Schwartz argued that the Federal Reserve System in the United States acted with such ineptness that it plunged the world into depression. They attributed this incompetence to the death of Benjamin Strong (president of the Federal Reserve Bank of New York) in 1928, and they describe several alternative monetary policies that they argue would have eased or even eliminated the economic contraction.

Even their story cannot separate the United States from the rest of the world, however. The Federal Reserve raised interest rates in October 1931 to defend the dollar, as noted above, even though the economy was contracting. Friedman and Schwartz characterized this action as an inept mistake, but they acknowledged the power of the gold standard to unite the financial community behind this perverse policy. This

contractionary policy in the midst of rapid economic decline was the classic central bank reaction to a gold-standard crisis.

Charles Kindleberger put forward a more international explanation in *The World in Depression* (1986). He argued that the lack of central bank leadership in the operation of the restored gold standard was key to the spread of the Great Depression: the proposition summed up in the phrase 'no longer London, not yet Washington'. The diminished financial status of Great Britain meant that London was unable to act as sole conductor of the international orchestra – or, in more modern terminology, to operate as the 'hegemon' – while the United States was not yet willing to take over this role despite the enormous improvement in its international economic standing.

Another factor which has been put forward as the primary explanation for the problems of the interwar period is the absence of international cooperation between the United States, Britain, France and Germany. Barry Eichengreen, in *Golden Fetters* (1992), identified this behaviour as a central feature of the period, manifest particularly in the attempt of each of the main powers to secure for itself a disproportionate share of the world's limited stocks of monetary gold. Prior to the collapse of the gold standard in 1931 their non-cooperative behaviour involved the imposition of tight monetary policies not only by countries in deficit, but also by those – notably the United States and France – which were in surplus. This added to the deflationary pressures on the world economy and increased the vulnerability of the weak currencies, such as the pound and the mark, to speculative attack.

## Recovery

The world began to recover from the contraction in 1933, when the United States and Germany both abandoned the policies of the gold standard, but the Great Depression was far from over. Unemployment continued to be high in most countries, as indicated in Table 1. The world economy split up into competing currency and trading blocs, and domestic policies to combat the hardships of depression changed the role of government.

Unemployment continued to be high in most countries throughout the 1930s. Measures designed to help workers often perpetuated unemployment. The National Industrial Recovery Act of 1933 in the United States attempted to bring order to industries and income to workers by allowing industries to enforce codes of conduct that raised both prices and wages. Rising wages impeded the extension of employment, trading off the benefits to the unemployed for benefits to those working. Germany under the Nazis expanded government spending and, apparently, decreased unemployment dramatically. France and other members of the gold bloc continued to maintain contractionary policies in an effort to retain the convertibility of their currencies into gold. Only when France devalued in 1936 could its recovery begin.

Recovery, however slow and halting, did not approach the status quo ante. The world economy fragmented in the 1930s, and recovery took place within relatively isolated currency and trading blocs. The United States began the process of reducing world trade with the Smoot–Hawley tariff of 1930. The United Kingdom abandoned its tradition of free trade in 1932 in favour of protection for the British Commonwealth.

Germany under the Nazis adopted a complex set of bilateral trading arrangements that reoriented its trade towards south-eastern Europe. International trade was much reduced, and international capital flows virtually disappeared.

Countries were changed internally as well. Governments became active in the economy as they attempted to reduce unemployment or ease the condition of the unemployed. Unions grew in many countries, helped both by legislation and by unemployment. Regulation grew as governments substituted direct controls for those of the market, and the world war that followed the Great Depression caused governments to take control even more firmly of their economies. The mixed economies and large governments that were typical of the last half of the 20th century were the legacy of the Great Depression and its aftermath.

It is not possible to separate the long-run effects of the Depression from those of the Nazis and the Second World War, but it is instructive to ask whether the Great Depression could have been avoided. There were indeed stresses on the world economy at the end of the 1920s, and the control mechanisms used in earlier times were not in good shape. The downturns in the United States and Germany would have produced a serious recession in the early 1930s in any case. The currency crises of 1931 then turned this recession into the Great Depression. If Germany and the United States had abandoned gold after Britain had chosen devaluation over further contraction, the world economy would have begun to recover two years earlier and before unbearable strain had been put on economic and political institutions.

Historians today debate how much freedom policymakers had in 1931. The German cabinet discussed devaluation after Britain left gold, but the memory of hyperinflation less than a decade before inhibited – if it did not preclude – an expansionary policy such as devaluation. The United States was not under the same economic pressure as Germany, but the Federal Reserve nonetheless raised interest rates sharply in late 1931 in response to gold outflows following Britain's devaluation. The Federal Reserve was following the dictates of the gold standard in actions that were applauded by the local financial community. It was a world tragedy – one that escalated from economics to politics and war – that the hold of the gold standard was so strong in the early 1930s that policymakers in the major economies chose to continue deflationary economic policies long after the need for expansionary measures was clear.

PETER TEMIN

*See Also* **gold standard; Kindleberger, Charles P.**

### *Bibliography*

Balderston, T. 1993. *The Origins and Course of the German Economic Crisis*. Berlin: Haude and Spener.

Bernanke, B. 1983. Nonmonetary effects of the financial crisis on the propagation of the Great Depression. *American Economic Review* 73, 257–76.

Bernanke, B. 1995. The macroeconomics of the Great Depression: a comparative approach. *Journal of Money, Credit and Banking* 27, 1–28.

Bernanke, B. and James, H. 1991. The gold standard, deflation, and financial crisis in the Great Depression: an international comparison. In *Financial Markets and Financial Crises*, ed. R.G. Hubbard. Chicago: University of Chicago Press.

Borchardt, K. 1991. *Perspectives on Modern German Economic History and Policy*. Cambridge: Cambridge University Press.
Bordo, M., Goldin, C. and White, E.N. 1998. *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*. Chicago: University of Chicago Press.
Calomiris, C.W. 1993. Financial factors in the Great Depression. *Journal of Economic Perspectives* 7(2), 61–85.
Calomiris, C.W. and Mason, J.R. 2003. Consequences of bank distress during the Great Depression. *American Economic Review* 93, 937–47.
Choudhri, E.U. and Kochin, L.A. 1980. The exchange rate and the international transmission of business cycle disturbances: some evidence from the Great Depression. *Journal of Money, Credit, and Banking* 12, 565–74.
Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression*. New York: Oxford University Press.
Eichengreen, B. 2004. Still fettered after all these years. *Canadian Journal of Economics* 37, 1–27.
Eichengreen, B. and Sachs, J. 1985. Exchange rates and economic recovery in the 1930s. *Journal of Economic History* 12, 925–46.
Eichengreen, B. and Temin, P. 2000. The gold standard and the Great Depression. *Contemporary European History* 9, 183–207.
Feinstein, C.H., Temin, P. and Toniolo, G. 1997. *The European Economy between the Wars*. New York: Oxford University Press.
Field, A.J. 1984. Asset exchanges and the transactions demand for money, 1919–29. *American Economic Review* 74, 43–59.
Friedman, M. and Schwartz, A.J. 1963. *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton University Press.
Hamilton, J.D. 1988. The role of the gold standard in propagating the Great Depression. *Contemporary Policy Issues* 6, 67–89.
James, H. 1986. *The German Slump: Politics and Economics, 1924–1936*. Oxford: Oxford University Press.
Kindleberger, C.P. 1986. *The World in Depression, 1929–1939*. Rev. edn. Berkeley: University of California Press.
Margo, R.A. 1993. Employment and unemployment in the 1930s. *Journal of Economic Perspectives* 7(2), 41–59.
Mouré, K. 1991. *Managing the Franc Poincaré: Economic Understanding and Political Constraint in French Monetary Policy, 1928–1936*. Cambridge: Cambridge University Press.
Romer, C.D. 1990. The Great Crash and the onset of the Great Depression. *Quarterly Journal of Economics* 105, 597–624.
Romer, C.D. 1993. The nation in depression. *Journal of Economic Perspectives* 7(2), 19–39.
Rothermund, D. 1996. *The Global Impact of the Great Depression, 1929–1939*. London: Routledge.
Temin, P. 1989. *Lessons from the Great Depression*. Cambridge MA: MIT Press.
Temin, P. 1993. Transmission of the Great Depression. *Journal of Economic Perspectives* 7(2), 87–102.

# Great Depression, monetary and financial forces in

What caused the worldwide collapse in output from 1929 to 1933? Why was the recovery from the trough of 1933 so protracted for the United States? How costly was the decline in terms of welfare? Was the decline preventable? These are some of the questions that have motivated economists to study the Great Depression.

Cole and Ohanian (1999) document that US per capita GNP fell 38 per cent below its long-run trend path (of two per cent per annum growth) from 1929 to 1933. Real per capita non-durables consumption fell nearly 30 per cent, durables consumption fell over 55 per cent, and business investment fell nearly 80 per cent. On the input side, total employment fell 24 per cent and total factor productivity (TFP) fell 14 per cent. On the nominal and financial side, the GNP deflator fell 24 per cent; per capita M1 (currency plus deposits) fell 30 per cent; M1 velocity fell 32 per cent; the per capita monetary base rose 9 per cent; the currency–deposit ratio rose over 160 per cent (Friedman and Schwartz, 1963, Table B3); the loan–deposit ratio fell 30 per cent (Bernanke, 1983, Table 1); and *ex post* real commercial paper rates rose from six per cent in 1929 to a peak of 13.8 per cent in 1932.

What caused the Depression? For the United States, Friedman and Schwartz (1963, p. 300) argued that it was the decline in the stock of M1 – a consequence of Fed tightening and of a fall in the money multiplier induced by banking panics. According to Eichengreen (1992), international adherence to the gold standard transmitted the US monetary contraction to other industrialized countries. Specifically, high interest rates and low prices in the United States attracted foreign inflows of gold (in 1932 the United States and France held over 70 per cent of the world gold reserves), which the Fed largely sterilized (that is, sold domestic government debt and bought money). The outflow of gold from foreign countries implied that gold-backed money supplies of those countries had to decline in order to meet their cover ratios. Further evidence (see Bernanke and James, 1991, Table 4) of the importance of the gold standard in transmitting the contraction comes from the experience of countries like Britain, which suspended the gold standard in 1931 and recovered by 1932; from Spain, which never was on it and had a much less severe contraction than those on the gold standard; and from France, which was one of the last major countries to leave it and still faced declining industrial production past the 1933 trough. As Bernanke (1995, p. 3) puts it: 'The new gold-standard research allows us to assert with considerable confidence that *monetary factors played an important causal role*, both in the worldwide decline in prices and output and in their eventual recovery.'

However, much of this evidence is problematic in that it is in the nature of correlations between *endogenous* variables – a fact that makes it challenging to establish causality. Did the decline in M1 *cause* the decline in aggregate output or – as Temin (1976) argued early on – did M1 and aggregate output decline in response to some other common shock? If the 'monetary-cum-exchange-rate-policy' explanation is indeed correct, we ought to be able to demonstrate its correctness in a reasonably

calibrated, dynamic stochastic general equilibrium (DSGE) model. To paraphrase Lucas (1993, p. 271): 'If we know what a depression is, we ought to be able to *make* one.' The challenge of 'making' a depression has been taken up by various researchers and constitutes a noteworthy recent development in depression research.

The conventional explanation of why money affected output is sticky nominal wages – goods prices fell as a result of the monetary contraction but nominal wages adjusted slowly and the ensuing increase in the real wage depressed the demand for labour. One significant contribution to evaluating this conventional explanation is by Bordo, Erceg and Evans (2000). They calibrate a one-sector stochastic macro model with four-quarter nominal wage rigidity and find that 70 per cent of the output decline from 1929 to 1933 can be accounted for by feeding in the negative innovations to the actual M1 money supply process during that period.

Although the findings of Bordo, Erceg and Evans are striking, there are some unresolved issues. One is that the real-wage rise in the model was chosen to mimic the actual real-wage rise in the manufacturing sector while there is some indirect evidence that non-manufacturing real wages actually fell during the 1929–33 downturn. Cole and Ohanian (2000) re-examine the sticky-wage hypothesis in a multisector model and find much less support for it.

A second unresolved issue is that Bordo, Erceg and Evans do not take into account the evidence on aggregate labour productivity and TFP, both of which declined between 1929 and 1933. Ohanian (2002) argues that only about a third of the decline in labour productivity and/or TFP can be plausibly accounted for by mismeasurement of factor inputs. By itself, a decline in TFP could account for a substantial fall in aggregate output, consumption and investment. Unless a decline in TFP can be viewed as an endogenous response to the monetary shock (through, for example, aggregate increasing returns), the decline leaves less scope for a purely monetary explanation. Using a DSGE model where money is non-neutral due to imperfect information, Cole, Ohanian and Leung (2005) show that the decline in M1 accounts for only one-third of the decline in output from 1929 to 1933, while the effect of an exogenous decline in TFP accounts for two-thirds. They use a misperceptions model of monetary non-neutrality because such a model generates less of a counterfactual movement in labour productivity than a model with nominal wage rigidities.

Sticky wages and monetary misperceptions are not the only mechanisms through which money can affect real output. Irving Fisher (1933) pointed out that the unanticipated fall in prices during 1929–33 led to bankruptcies because it increased the real value of nominal debt of households, firms, and financial intermediaries. This 'debt-deflation' hypothesis was analysed by Mishkin (1978) for households and formalized by Bernanke and Gertler (1989) for firms. More generally, Bernanke (1983) argued that the reduction in borrower net worth increased the cost of obtaining external finance, while bank failures and tightened credit standards hampered the efficient allocation of capital. However, a quantitative DSGE model featuring this mechanism has yet to be implemented for the Great Depression. Such a model holds out the promise of explaining some portion of the puzzling decline in TFP during 1929–33 as an endogenous response to a misallocation of capital.

One of the most striking facts of the Depression was the reduction in the money multiplier from 1929 to 1933 associated with the flight from bank deposits to currency. Cooper and Corbae (2002) construct a model in which households have the option of saving in the form of currency or bank deposits, and in which bank deposits ultimately fund working capital for businesses. Because of increasing returns in the intermediation technology associated with fixed verification costs, their model admits multiple equilibria. In the good equilibrium the return on bank deposits is high, households hold small amounts of currency, and output is high. In the bad equilibrium, the return on bank deposits is low, households substitute into currency, and output is low. A shift from the good to the bad equilibrium replicates many of the salient nominal changes that occurred between 1929 and 1933. Although not quantitative, their work formalizes the idea that output, credit and money supply responded negatively to a loss in confidence – much as Irving Fisher (1933, p. 343) suggested it did.

Why was the recovery from the trough of 1933 so protracted for the United States? As noted by Cole and Ohanian (1999), aggregate US output was still below trend in 1939. The answer cannot be the gold standard or M1 because the United States left the gold standard in 1933 and the US money stock recovered rapidly thereafter. One explanation offered is that the National Industrial Recovery Act (NIRA) encouraged businesses to accept high real wages of industrial workers. Cole and Ohanian (2004) embed labour bargaining into a DSGE model and quantitatively explore the effect of the NIRA, giving more weight to workers in the bargaining process post 1933. Their model is reasonably successful in producing a slow recovery. Adverse labour market interventions also appear to have played a role in other industrialized countries such as Germany, France, the UK and Italy (Kehoe and Prescott, 2002).

How costly was the Depression in terms of welfare? Real per capita consumption of non-durables fell 30 per cent in the United States but it is not known how this decline was distributed across households. Chatterjee and Corbae (2006) analyse how households that can self-insure against uninsured earnings losses would fare through a depression. They found that the welfare cost of living in a world with a small likelihood of a Depression-like event is quite large – somewhere between one and seven per cent of consumption in perpetuity depending on the completeness of asset markets. Much of this cost is associated with the increased variability of individual consumption streams.

Was the Depression preventable? First, if the 'monetary-cum-exchange-rate-policy' explanation is correct, the right monetary policy could have prevented the decline. Christiano, Motto and Restagno (2003) estimate a DSGE model with many shocks but find that a liquidity preference shock inducing households to hold currency instead of deposits played the most important role in the contraction phase of the Depression. They then specify a policy rule that raises the monetary base as a function of liquidity shocks, and run a counterfactual experiment where they find that output would have declined only six per cent if such a reaction function had been in place. Second, if a portion of the decline in output was the result of a banking collapse stemming from a shock to confidence, then – as shown by Cooper and Corbae (2002) – an announcement by the monetary authority that it stands ready to supply liquidity to the banking system might have moderated the decline. Finally, with regard to the slow recovery in

the United States, the only credible explanation offered is adverse labour market intervention. If this explanation is correct, we know what *not* to do to prolong a severe decline in output.

SATYAJIT CHATTERJEE AND P. DEAN CORBAE

*See Also* **Great Depression; Great Depression (mechanisms); monetary business cycle models (sticky prices and wages); real business cycles.**

### *Bibliography*


Bernanke, B. 1983. Nonmonetary effects of the financial crisis in the propagation of the Great Depression. *American Economic Review* 73, 257–76.

Bernanke, B. 1995. The macroeconomics of the Great Depression: a comparative approach. *Journal of Money, Credit and Banking* 27, 1–28.

Bernanke, B. and Gertler, M. 1989. Agency costs, net worth, and economic performance. *American Economic Review* 79, 14–31.

Bernanke, B. and James, H. 1991. The gold standard, deflation, and financial crisis in the Great Depression: an international comparison. In *Financial Markets and Financial Crises*, ed. R. Glenn Hubbard. Chicago: University of Chicago Press.

Bordo, M., Erceg, C. and Evans, C. 2000. Money, sticky wages, and the Great Depression. *American Economic Review* 90, 1447–63.

Chatterjee, S. and Corbae, D. 2007. On the aggregate welfare cost of Great Depression unemployment. *Journal of Monetary Economics* 54, 1529–1544.

Christiano, L., Motto, R. and Rostagno, M. 2003. The Great Depression and the Friedman–Schwartz hypothesis. *Journal of Money, Credit, and Banking* 35, 1119–97.

Cole, H. and Ohanian, L. 1999. The Great Depression in the United States from a neoclassical perspective. *Federal Reserve Bank of Minneapolis Quarterly Review* 23(1), 25–31.

Cole, H. and Ohanian, L. 2000. Re-examining the contributions of money and banking shocks to the U.S. Great Depression. *NBER Macroeconomics Annual* 15(1), 183–227.

Cole, H. and Ohanian, L. 2004. New Deal policies and the persistence of the Great Depression: a general equilibrium analysis. *Journal of Political Economy* 112, 779–816.

Cole, H. Ohanian, L. and Leung, R. 2005. Deflation and the international Great Depression: a productivity puzzle. Working Paper No. 11237. Cambridge, MA: NBER.

Cooper, R. and Corbae, D. 2002. Financial collapse: a lesson from the Great Depression. *Journal of Economic Theory* 107, 159–90.

Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression 1919–1939*. New York: Oxford University Press.

Fisher, I. 1933. The debt–deflation theory of Great Depressions. *Econometrica* 1, 337–57.

Friedman, M. and Schwartz, A. 1963. *A Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.

Kehoe, T. and Prescott, E.C. 2002. *Great Depressions of the Twentieth Century*. Special Issue of the *Review of Economic Dynamics* 5(1).

Lucas, R. 1993. Making a miracle. *Econometrica* 61, 251–72.

Mishkin, F. 1978. The household balance sheet and the Great Depression. *Journal of Economic History* 38, 918–37.

Ohanian, L. 2002. Why did productivity fall so much during the Great Depression? *Federal Reserve Bank of Minneapolis Quarterly Review* 26(2), 12–17.

Temin, P. 1976. *Did Monetary Forces Cause the Great Depression?* New York: W.W. Norton and Company.

# International Monetary Fund

The International Monetary Fund (henceforth 'the IMF' or 'the Fund') was conceived at a conference at the Mount Washington Hotel in Bretton Woods, New Hampshire, in July 1944 and its Articles of Agreement entered into force in December 1945. The World Bank (henceforth 'the Bank') was set up at the same time. The IMF was established to promote international monetary cooperation and the elimination of exchange restrictions on current account transactions; to facilitate trade, economic growth and high levels of employment; to foster exchange rate stability; and to provide temporary financial assistance to countries so as to ease balance of payments adjustment. More specifically, it was given the role of supervising a system of pegged but adjustable exchange rates, which became known as the Bretton Woods system. In the first two sections of this entry we explain how the Bretton Woods system worked, and why it broke down in 1971. In the following sections we consider the roles which the Fund now plays, which differ from its original activities. They are: surveillance, ensuring stability for the international financial system and for individual economies within this system, and assisting the world's poorest economies. As part of each of these three activities, the Fund also provides policy advice and technical assistance. This is a much less clear collection of responsibilities, and, as a result, the future direction of the Fund is somewhat uncertain. The aim of this article is to review the achievements of the Fund, and also the challenges that lie ahead. A related overview of some of the issues discussed here can be found in Gilbert and Vines (2004).

## 1 The Bretton Woods system

### *1.1 Intentions*

As the Second World War drew to a close, the United Kingdom, the United States and their allies, inspired in part by the *General Theory* of John Maynard Keynes (Keynes, 1936), established a policy framework in which countries would be able to promote high levels of employment and output, by means of demand management policies, focused mainly on fiscal measures. This would – it was hoped – avert slumps in growth and would thereby prevent the re-emergence of the kind of global depression that had occurred in the 1930s. (See Williamson, 1983a; Moggridge, 1986.)

From early on, Keynes had seen that such policies would need global support. This is because they would have to be reconciled with the need for each country to be sufficiently competitive; that is, each country would need to be able to export enough to pay for the imports that would be purchased at full employment. In 1942, Keynes put forward plans for a new post-war international monetary system designed to make this possible, which he called a 'Clearing Union'. (See Keynes, 1971–88, vol. 25, pp. 41–67; van Dormael, 1978; Gardner, 1956.) His plan drew on the theoretical arguments in his *General Theory*, and also on the harsh practical example provided by the United Kingdom's return to the gold standard in 1925 (Eichengreen, 1992). He argued that,

for many countries, sufficient competitiveness would not be assured if the world returned to a gold standard after the war. Such a standard would require that any country with balance of payments difficulties, of the kind which Britain was likely to have, would need to rely on downward adjustment of its wages and prices in order to make its goods sufficiently attractive in world markets. Keynes judged that, in the political climate of the post-war world, such wage and price adjustments might not be possible. Nevertheless, because of the exchange rate instability of the early 1920s and the 1930s, he also showed no enthusiasm for floating exchange rates. The need for something different was discussed in much detail over the next two years with Harry Dexter White and others from the United States (Keynes, 1971–88, vol. 25, pp. 338 ff.), including during a visit that Keynes made to Washington in 1943.

The analytical content of these immensely difficult negotiations is explained in Meade, James Edward, and is discussed in more detail in Vines (2003), which draws on the wonderful historical account by Skidelsky (2000). Skidelsky makes clear that Keynes was propelled in these discussions by the knowledge that the generous provision by the United States of wartime funding to the United Kingdom ('Lend Lease') had put the United States in a position in which it would be able to dismember the British Empire after the war. Keynes, who had been accustomed to Britain managing the global economy, wanted to create a new global order in which prospects for Britain remained acceptable, even although global economic hegemony would pass to the United States. He feared that difficulties in the balance-of-payments adjustment process might impose, on deficit countries like Britain, an obligation to deflate demand below full employment, something which might not be matched by symmetrical over-expansion by surplus countries, and might thereby create pressures towards global deflation. This is why he wanted his Clearing Union to be able to create global liquidity. (Like a bank, it would 'clear' the overdrafts which countries could obtain from it.) He differed in this view from Harry Dexter White, who feared an outcome in which liquidity would be so freely available that there would be a great post-war worldwide inflation.

What emerged at Bretton Woods was a global system of pegged but adjustable exchange rates, to be overseen by an International Monetary Fund. The currency system was to have three major features. First, each country would establish a *par value* for its currency in terms of gold or dollars. Second, all exchange controls would be removed for current-account transactions and all currencies would be freely convertible into dollars, although controls on international capital flows would remain in place. Third, dollars would be freely convertible into gold. Thus, the system was to be a 'gold exchange standard'; it would differ from a gold standard in being a club rather than a unilateral pegging arrangement, and in allowing for occasional exchange rate changes.

The IMF would do two things in this system. First, exchange-rate pegs would only be adjusted if the approval of the IMF's Executive Board had been obtained. That approval would not be given unless there were deemed to be a 'fundamental disequilibrium'. This term was imprecisely defined, but it meant a situation in which an exchange rate was not at a level that would ensure that exports could equal imports at full employment. This kind of test was designed, with the 1930s in mind, to prevent countries pursuing a 'beggar-thy-neighbour' devaluation of their currencies so as to

steer towards full employment by 'stealing' jobs from other countries rather than by expanding expenditure at home. A country with longer-term difficulties would be declared to be in 'fundamental disequilibrium' and would be expected to devalue its currency by an appropriate amount after consulting with the Fund and getting the required approval. Similarly, a country with an excessively large and sustained balance of payments surplus would be expected to revalue its currency.

Second, the Fund would be set up like a credit union, into which members would place deposits; a country in temporary balance of payments difficulty rather than 'fundamental disequilibrium' would be able to draw on a short-term basis from the Fund to help it address the problem. It was thought that these loans would be repaid quite rapidly (that is, within three to five years), since more fundamental difficulties would be addressed by exchange rate adjustments. Each country in this credit union was to be given a 'quota', based on a nonlinear equation that took account of a country's national income, its international trade, and its official reserves; services, other external current account transactions, and a measure of volatility were further added to the quota formula in the 1960s. The quotas would define each country's capital contribution, its borrowing entitlement, and, in aggregate, the Fund's lending capacity. The US quota was initially about 20 per cent of the total (less than would have been implied by a strict calculation based on the variables noted above), and originally the United Kingdom had, by design, the second largest quota. This was *not* like Keynes's Clearing Union, and Keynes was dismayed at how little the Fund would be able to lend (see Vines, 2003). There have been a number of substantial increases in total quotas under regular quinquennial reviews, but they have not grown in such a way as to keep pace with the expansion of the world economy and international financial flows. In addition, as the relative size and importance of countries have changed, there has been a need to adjust both quota shares and the factors used in the calculation of these quotas. Both of these types of adjustment have been politically difficult; a (small and interim) adjustment for four emerging-market countries (China, Korea, Mexico, and Turkey) happened in September 2006.

The quota system partly determined the relative voting entitlements of countries on the Executive Board of the Fund. It seemed obvious, for a credit union to which money had been contributed, to make voting power depend partly on the amount contributed, and on the amount which could be borrowed at a time of difficulty, rather than using a one-member, one-vote system of governance like that adopted at the United Nations. However, there were also a number of 'basic votes' allotted equally to all members, whose effect was to mitigate a little the voting power of large countries.

The Fund's Articles and their subsequent amendments established that a member is allowed to borrow up to a certain proportion of its quota as of right, without policy conditions. This amount was referred to as the 'reserve tranche'; it was equal to 25 per cent of quota and corresponded to the amount that a member had paid into the Fund in hard foreign currencies. Beyond the reserve tranche, a country had an option to borrow up to four 'credit tranches', each of which represented 25 per cent of quota. Access to the first credit tranche was relatively easy; borrowing under the subsequent or 'upper' credit tranches was normally made available through what were (and still are) described rather quaintly as 'stand-by arrangements'.

*1.2 Consequences*

The international monetary system followed only imperfectly the intentions underpinning the Bretton Woods system, and only until 1971. (See de Vries, 1976.) Current-account convertibility, for most European currencies, was not achieved until 1958 (the year after a large US current account deficit). There was a reluctance to alter exchange rates even in the presence of 'fundamental disequilibrium'. And the Fund was unable to stop France from implementing a multiple currency system in 1948. One major currency, the Canadian dollar, floated from 1950 to 1962 and the Fund acquiesced in this. The Fund ratified British devaluations in 1949 and 1967 at short notice (though it was closely involved in discussions in the second case). It had little influence on US policies – and has had little influence ever since. It played virtually no role in the later US decision to end gold convertibility in August 1971, a decision which brought the Bretton Woods system crashing down. And it had limited influence on the policies of the principal surplus countries in the 1960s. On the other hand, the Fund did have a role in the exchange rate realignments of other currencies that took place in 1949, 1967 and 1971 as a result of the sterling and dollar devaluations, seeking to ensure 'orderly adjustment'. The most important point is that the IMF had an influence mainly through the conditions it could impose on those countries (such as the United Kingdom in 1976) which needed its funds.

When the Fund began providing stand-by arrangements in 1952 they were typically of short duration and did not feature any conditions. This may seem surprising now, given the close association in the popular imagination between conditional lending and the IMF. Policy conditions were first added to Fund-supported programmes in 1954, partially in light of the increase in the size of borrowing under stand-by arrangements, as compared with first-credit tranche financing. Quantitative targets or 'performance criteria' followed in 1957, in order to provide a clear baseline for policymaking under IMF-supported programmes, and an objective yardstick by which the effects of these policies – and the possible need for further adjustments – might be assessed. They were calibrated using the Fund's financial programming framework, developed by Polak (1957), and came to be a nearly universal feature of Fund-supported programmes by the mid-1960s. (See IMF, 1987; 2004a; Mussa and Savastano, 1999.) This combination of policy or 'structural' commitments and quantitative performance criteria came to characterize the 'conditionality' attached to IMF lending from the 1960s to the present. This was justified – then as now – not so much as a way of collateralizing IMF lending, and guaranteeing a turnover of the IMF's funds, but rather as a means of ensuring the viability of Fund-supported programmes and the quick adjustment of countries in crisis back to a balanced growth path.

The period from 1945 to 1971 was one of extraordinary dynamism (a 'golden age'): it was a time in which Europe and Japan were first rebuilt after the war and then proceeded to catch up with the United States. The Bretton Woods system appears to have played a part in ensuring that this happened. In this system, the Fund was helped by the World Bank, whose role was to lend money for longer periods than the Fund, first for reconstruction after the war, and then, later on, to help finance development. (Keynes once helpfully remarked that in order to comprehend the Bretton Woods institutions one has to understand that the Fund is a bank, and the Bank is a fund.)

The purpose of this World Bank lending was to enable these countries to borrow abroad (in a world in which there was little international mobility of private capital), to run balance of trade deficits, to invest, and to grow – with the expectation that the borrowing would then be repaid out of the increased export proceeds that investment and growth made possible. In addition, a conference in Geneva in 1947 established the General Agreement on Tariffs and Trade (or GATT) to supplement the Bretton Woods system by encouraging the growth of international trade. The GATT's role in promoting the liberalization of trade restrictions supplemented the Fund's role in promoting the liberalization of exchange restrictions on current account transactions. In due course, a series of GATT 'rounds' brought about tariff reductions, which helped to create markets for exports as countries expanded. With high employment, with balance-of-payments deficits dealt with as described above, and with many countries growing by exporting, there were clear incentives for most countries to support trade liberalization. That, in turn, made exports and imports more sensitive to exchange-rate levels and so made balance of payments adjustment easier to achieve by exchange-rate adjustments. Yet, these linkages between different aspects of the overall post-war policy framework are difficult to pin down empirically. This explains why economic historians still differ in their view as to how important the Bretton Woods system actually was in sustaining the golden age of growth observed in the 1950s and 1960s. (See Matthews, Feinstein and Odling-Smee, 1982; Matthews and Bowen, 1988; Temin, 2002; papers in Eichengreen, 1995; and Eichengreen, 2007.)

## 2 Breakdown and reconfiguration

Up to the 1960s the growth of gold reserves had been slow, and the need for additional international liquidity was increasingly met by the use of the US dollar as a 'reserve currency'. This led to calls for the IMF to create a more multilateral way to augment official reserves. The IMF's Articles of Agreement were eventually amended in 1969 to allow the Fund to create 'special drawing rights' (SDRs) that would act as the Fund's unit of account and which could be used as a source of credit for member countries. (See Corden, 1983a; Boughton, 2001.)

In the 1960s, imbalances also began to emerge: by the latter part of the decade, the United States had a large balance of payments deficit. A belief emerged that the dollar price of gold might rise as economic growth in Europe and Japan weakened the US dollar's role as anchor of the Bretton Woods system. In 1968, central banks ceased their efforts to control the dollar price of gold in private markets, which meant that the prevailing fixed price of gold applied only to central bank dealings. The market price of gold rose: in August 1971, following a massive speculative attack on the dollar, the United States ended the gold convertibility of dollars held by central banks and, as a result, the entire gold exchange standard broke down. A reluctant movement from a pegged exchange-rate system to a system with floating exchange rates followed. This outcome can best be explained by three sets of factors. (See Corden, 1993.)

First, many countries were unwilling to adjust the exchange rates for their currencies in the face of fundamental disequilibria. It was particularly problematic that the core country, the United States, behaved in this way. Because US productivity growth lagged

behind that of the countries which were catching up with it, the trade position of the United States was at risk by the late 1960s. In addition, the United States fought the Vietnam War and launched its 'Great Society' programmes at the same time, without adequately raising taxes. The result was a large balance of payments deficit for the United States, the correction of which required both real exchange rate depreciation and restraint of domestic expenditure. Neither of these actions was forthcoming.

Second, the growth of international capital flows – which was in part a result of the international stability associated with the golden age – helped to undermine the system. As first demonstrated by the 1967 sterling crisis, it was no longer possible for the IMF and national governments to set exchange rates without reference to the forward-looking perceptions of private markets about what sustainable exchange rates might be. With increasingly mobile capital, once a suspicion was generated that that there would be (or might need to be) a devaluation of a country's currency to preserve external balance, speculation could make it difficult or impossible for central banks to defend an existing rate. By 1971, the balance of payments deficit of the United States had caused a large build-up of mobile dollar holdings in offshore or 'Euro-dollar' accounts. These funds were used to finance the speculative attack on the dollar in 1971.

Third, the Keynesian macroeconomic policy framework established after the Second World War contained no clear responsibility for preventing inflation. Although there were periods of (generally unsuccessful) price controls or 'incomes policy', the seeds of incipient inflation were sown by this omission. Eventually, tensions generated by the oil price shock of 1973, and by the period of undisciplined inflation which followed it, led to more than the collapse of the Bretton Woods system. The entire structure of Keynesian, interventionist, high-employment policies, which had been at the centre of the post-war policy architecture, came tumbling down, both in the United States and in Europe. For the ten years after 1971, macroeconomic policy was in a state of worldwide disarray.

The great inflation of the 1970s led to significant movements in the real exchange rates between countries, which killed nearly all of the (many) attempts made at the time to reconstruct an international monetary system with pegged exchange rates. (See Williamson, 1977.) There was only one lasting, partial, attempt to reconfigure such a system, in Europe, which led to the European Monetary Union.

For a period of time it appeared that the Keynesian approach to macroeconomic policy might be replaced by monetarist policies of a non-interventionist kind. But this alternative proved unsuccessful. Instead, with great difficulty, activist macroeconomic policies were reconstructed by the 1990s within inflation-targeting regimes, in which an inflation target was pursued through interest rate changes. This new system quickly came to be allied with a system of floating exchange rates in which there was a high degree of international capital mobility. In this new set-up, a floating exchange rate would help to stabilize demand, and movements in the exchange rate would become an important part of the process of inflation control. If a country suffered from a shock which raised prices, then its monetary policymakers would set higher interest rates, and the nominal exchange rate of the country would appreciate. This would reduce net exports and import costs, and so inflation.

As a result of this reconfiguration of policy assignments, a second revision of the Fund's Articles of Agreement was made in 1976 and came into effect in 1978.

At Bretton Woods, the Fund had been set up to manage a pegged exchange rate system. But it came to be realized that a country cannot have, at the same time, an independent monetary policy, capital markets which are open to the rest of the world, and a pegged exchange rate. (These three things, taken together, have become known as an 'impossible trinity'. The reason that these things cannot occur together is to be found in the Mundell–Fleming macroeconomic model, which was developed by Fleming and Mundell, at the IMF, in the early 1960s.) As a result, the Fund's revised Articles ratified a new form of international monetary system in which a country did not have to establish a par value for its exchange rate, but could instead have exchange rate arrangements of its own choice.

Since 1978, the Fund has gradually been drawn into new roles, in support of this revised, and more flexible, system. As described in the introduction, its work now has three aspects. First, the Fund's Articles, as revised in 1976, require it to exercise surveillance and influence over macroeconomic policies, and to monitor and guard against the development of unsustainable conditions that could lead to financial crisis. The Fund still lends to countries in balance of payments difficulty, and its second activity has been to do this for emerging-market economies and for 'transition economies' moving from central planning to market-based systems. More than this, the Fund helps such countries to deal with, and to prevent, the financial crises that have afflicted a number of them. Third, the Fund has lent money to the poorest developing countries, which generally do not have capital-market access. In these cases, Fund lending has often been indistinguishable from other long-term concessional development assistance, and the Fund's main distinctive contribution has been to work with central banks and finance ministries in crafting credible macroeconomic frameworks that can elicit further support from aid donors. We consider each of these three activities in turn.

## 3 The IMF and policy surveillance

Countries that are creditworthy, and which have access to highly mobile international capital under floating exchange rate regimes, no longer need to borrow from the Fund in the way they did when the Fund was first established. Such countries can adjust to balance of payments disequilibria through exchange rate movements, supported by foreign borrowing from sources other than the Fund. (See Corden, 1983b, and Dam, 1982). At the time of writing, no advanced country had agreed a borrowing arrangement with the Fund since the substantial stand-by arrangements with the United Kingdom and with Italy in 1976. Fund lending is only required at a time when a country ceases to be perceived as clearly creditworthy, something which, as of mid-2007, had not happened in industrial countries since 1976. This was true even at the time of the crisis of the Exchange Rate Mechanism of the European Monetary System in 1992. The Fund did not at that time provide financing to assist Sweden, Italy, the United Kingdom, or France in a defence of their currencies. When crisis struck, these countries (eventually) allowed their currencies to float downwards, rather than using lending from the IMF to defend further their exchange rates.

Nevertheless, a world with a high degree of international capital mobility is not without difficulties. In such a system, the spending decisions of nations can move away

from permanently sustainable positions for very long periods of time, an outcome with an external current account deficit (or surplus) offset by an external capital account surplus (or deficit). The 'global imbalances' that can result have, as of mid-2007, been substantial at three points of time since the 1960s. In the late 1960s, as we have seen, the US ran a large current account deficit; current account surpluses of a number of European economies and of Japan, which, as noted above, were engaged in a process of export-led growth and 'catch-up', were the 'other side of the coin'. Nearly 20 years later, in the early to mid-1980s, President Reagan increased defence expenditures and cut taxes. Tight monetary policy was used to restrain demand in the United States, which caused the dollar to appreciate, and the result was a large current account deficit. Japanese current account surpluses were on the other side of this coin. Twenty years later, in 2007, the United States was again running a large fiscal deficit and an (unprecedentedly) large current account deficit; and again Japan was running the corresponding current account surpluses, along with China, other emerging-market economies in East Asia and elsewhere, and a number of oil-producing countries.

These global imbalances reflect decisions by countries to de-link income and spending over time. Of course, such 'intertemporal trade' can be welfare-improving. But such imbalances might instead reflect an urge by a deficit country to spend beyond its means. This was clearly the case for the United States in the late 1960s and the mid-1980s, and might also be the case from 2000 (and especially from 2005). Conversely, these imbalances might also partly reflect a desire by some countries to maintain their currencies at artificially devalued levels against the US dollar, in order to grow quickly through a process of export-led catch-up. This is something which, at one time, would have been called 'beggar-thy-neighbour' behaviour of the kind which the IMF was established to prevent. As noted above, one can argue that this may have been what was done by western Europe and Japan in the late 1960s. Some commentators have argued that a number of emerging-market economies in East Asia, and elsewhere, were behaving the same way in the early 21st century (Dooley, Landau and Garber, 2003; Roubini and Setser, 2005). These commentators, in recognition of the parallel, suggested that we were living under a 'Bretton Woods II' regime.

But global imbalances eventually unwind. They must do so if countries are eventually to repay what they owe. In 1971, global imbalances led to crisis, and to the collapse of the Bretton Woods financial system. By contrast, the imbalances of the mid-1980s were resolved in an orderly way. (See Eichengreen, 2004; Eichengreen and Park, 2006; Corden, 2007; Joshi, Lane, and Vines, 2006; Williamson, 2006.) Such orderly adjustment requires the deficit country to cut expenditure, and its currency to depreciate significantly (unless it grows its way out of difficulty). It also requires, in addition, that expenditure in surplus countries expands so that global expenditure is maintained, or, if this does not happen, that global interest rates fall so that global expenditure is stimulated by other means. If all of this happens, as it did in the late 1980s, then the benefits of intertemporal separation between spending and income may not be diminished by the costs of an adjustment crisis.

There are four main ways in which the existence of the Fund helps global imbalances to unwind in an orderly manner.

First, ever since the second amendment of the Fund's Articles described above, the Fund has been required to exercise 'firm surveillance' over the exchange rate and

macroeconomic policies of its members. As a result, the Fund regularly sends to each country an 'Article IV mission' whose purpose is to review the country's macroeconomic policies. This is done annually for most countries, and at interludes of up to 24 months in countries with active Fund-supported programmes. (For such countries the Article IV cycle is elongated since policies are reviewed frequently in the context of semi-annual or quarterly programme reviews.) All aspects of macroeconomic policy are considered on these occasions. Following the emerging-markets crises of the 1990s and early 2000s, the Article IV consultation process has been supplemented by detailed review of countries' financial sectors under the World Bank and IMF's joint Financial Sector Assessment Program (FSAP).

Second, the Fund provides a vast amount of published information and analysis, both about the world economy and financial system in general and about particular countries. The Fund's biannual *World Economic Outlook* provides a forecast for the world economy, and analyses multilateral and regional issues; this report is supplemented by *Regional Economic Outlook*s. These products are based in part on Article IV consultations and would not be possible without that process. The Fund also publishes a biannual *Global Financial Stability Report* which monitors markets, and several statistical publications that compile economic and financial data supplied by member countries, including *International Financial Statistics*.

Third, the Fund plays an important role in keeping the governments of all members in touch with developments in other countries and globally. The Article IV missions to the largest economies (and the related research, published in *Selected Economic Issues* papers that are companions to the Fund's Article IV staff reports) are particularly important in helping to keep governments informed of policies and developments that are likely to affect the world economy as a whole. Additionally, the Annual Meetings of the Boards of Governors of the IMF and the World Bank enable an informed exchange of ideas between countries, as do the Spring Meetings. The Fund thus provides a valuable global information network.

Finally, the Fund has also created a valuable global human network. Fund staff are of high quality, something which is necessary since they have to deal with senior officials in many countries. The offices of Executive Directors of the Fund in Washington act as valuable means of communication between the member nations of the Fund. And in many national capitals a large number of public servants and elected officials have served on the Fund staff earlier in their careers, or have been located in Washington as Executive Directors at the Fund or as members of staff in Executive Directors' offices. This experience has made many decision-makers more internationally minded than they might otherwise have been.

Nevertheless, some have argued that the Fund's 'firm surveillance' is not firm enough. Arriazu, Crow and Thygesen (1999) discuss the impact of Fund surveillance, country by county, in the Article IV consultation process. They note that, although these consultations have been 'taken seriously', it does not appear that these reviews by the Fund have had more than an occasional impact on national policy decisions in some countries. A more recent assessment of Article IV consultations by Meyer et al. (2004) reaches similar conclusions. When an Article IV mission goes to a country that does not borrow from the Fund (and which therefore does not require the Fund's

*imprimatur* in order to obtain loans from other official creditors or from banks), the mission is usually relegated to a mainly advisory role, for which 'surveillance' may be too grand a label. But this *de facto* situation is not inevitable, since the *de jure* position of the Fund is that it should assess and appraise as well as advise. Goldstein (2006) asserts that there are gaps in the current practice of bilateral surveillance and argues, in particular, that the Fund's dealings with China in the early 21st century have not been satisfactory in addressing and effecting remedies for exchange rate misalignments. He further observes that the Fund's Managing Director has only rarely used the power granted to him by the 1977 and 1979 Board decisions on ad hoc and 'supplemental' consultations with members to address cases where a country's exchange rate policies appear inconsistent with the exchange rate principles of the Fund's Articles. (See Boughton, 2001.)

It is important to note that these critics do *not* seek policy changes from countries, in the interests of the greater good, that such countries would find unattractive if left to make policy choices on their own. That is, it is not suggested that the Fund could enforce a 'cooperative' outcome in macroeconomic policymaking when countries would prefer a different selfish, or 'Nash,' outcome. (This difference between Nash and cooperative outcomes was much discussed in the 1980s literature on policy coordination, summarized by McKibbin, 1997). Instead, it is argued that the Fund could enable cooperative outcomes, so that any adjustments in countries' policies that need to happen in the face of global imbalances might happen in the right sequence rather than in a disorganized manner. The capacity to enforce even this modest form of coordination might occasionally be important in the adjustment processes. (See Kumar, 2006; Wolf, 2005; 2006; Joshi, Lane and Vines, 2006.)

There was action of this kind under the Plaza Accord of September 1985, although it was not coordinated by the Fund. At this time, the finance ministers of the world's five largest national economies agreed that the value of the dollar needed to go down. They also arrived at some (rather general) agreements on the monetary and fiscal policies that would be needed in order for this fall in the dollar to be achievable, and announced coordinated intervention in foreign-exchange markets to help bring it about.

To act effectively in this way requires the Fund to come to terms with the difficult tension between its strengths as a universalist institution and the need, on occasion, to bring together a more limited group of players. But it is an objective of the Fund's current Medium-Term Strategy that it should provide such a forum (IMF, 2005b). The Fund's Multilateral Consultation on global imbalances began by consulting with the United States, the European Union, Japan, China and Saudi Arabia, and it reported on its findings in April 2007. This work ran in parallel with similar discussions at summit meetings of Heads of Government of the Group of Eight Countries (or G8), and at meetings of the finance ministers and central bank governors of these countries. The G8 consists of the United States, Russia, Japan, Germany, Britain, France, Italy, and Canada. This is a powerful collection of countries, but it is not clear that these G8 meetings have had the right participants to deal with the global imbalances of the early 2000s. China and India have not been members of this group (though they have been observers), nor have many of the major oil-producing economies; by contrast, Canada and Italy, while committed to the G8 process, have been perhaps too small to

contribute substantially to coordinated efforts to unwind global imbalances. The Fund may therefore have more to offer than such G8 gatherings, since the Fund can act as a locus of coordination amongst subsets of its membership, convening small groups of countries to deal with particular problems.

Nevertheless there are three reasons why further progress may be slow on this front.

First, in the words of the IMF's Independent Evaluation Office (IEO) (IMF, 2006a, p. 2), 'As a result of its ... [country-by-country] orientation, multilateral surveillance has not sufficiently explored options to deal with policy spillovers in a global context'. Pursuing this theme, Mervyn King, Governor of the Bank of England, made it clear (King, 2006b) that more effective multilateral surveillance would require: (i) that countries made clearer commitments about their objectives for macroeconomic policies (that is, fiscal, monetary and financial); (ii) that the Fund's Article IV and the *World Economic Outlook* processes focused more transparently on cases when these policy commitments, and the countries' policy actions, are not globally consistent; and (iii) that this process also transparently demonstrated the negative spillover effects that come from such lack of consistency and proposed actions to reduce such negative spillovers. But, given the limits to the precision of what we know about the international economy at any given time, doing this would be difficult. And it should be noted that the Fund's management issued a rejoinder to the 2006 IEO report which explained this difficulty.

Second, there may well be governance limitations on such firm surveillance. As of 2007, Article IV consultations were not finalized by the Fund Staff sent on the Article IV mission, but by the Fund's Executive Board, whose views were conveyed to the authorities of the country concerned after discussion at the Board. It is possible that this has compromised the space for missions to assess and appraise frankly. If the process of IMF surveillance were made more independent of the IMF's Executive Board, then this might allow clearer messages to be delivered to the Fund's member countries. As against this, the messages might then lose political weight because they would no longer be seen as the views of the global community represented in the Executive Board.

Third, and fundamentally, the Fund is not an agent of a sovereign state in the way that central banks (except the European Central Bank) are, however 'independent' these central banks may be. As a result, the Fund has no actual instruments of its own with which its recommendations on global cooperation can be implemented. It must always rely on being able to persuade its members to act.

## 4 The IMF and crises in emerging markets since 1980

In the mid-to-late 1970s, after the rise in the price of oil in 1973, funds flooded from oil producers on to the international capital market and flowed to middle-income countries. The early to mid-1990s saw a further massive surge of private capital flows into emerging market economies, and this was repeated in the mid-2000s. The economic benefits of such international mobility are obvious: if capital flows from relatively rich to relatively poor countries, and if the rate of return is high in poor countries, the potential gains are high for both borrower and lender. But such funds

are not always used well, the volatility of these flows can be very high, and they can create dangerous mismatches in the maturities and currencies of assets and liabilities. Indeed, these flows contributed to three major waves of financial crises, in Latin America, East Asia and Russia, something which called into question the stability of the entire international financial system. Across these regions of the world, the IMF has been required to help prevent such crises through surveillance. It has also been required to assist in the orderly workout of crises, through lending and through ongoing engagement in the development of macroeconomic policies in the countries which it assists. We explain how the Fund's activities have evolved in these emerging-market economies, and how its role has broadened. We do this by examining the three generations of emerging-markets crises that occurred from the early 1980s onward.

*4.1 The Latin American debt crisis: a 'first-generation' crisis*

Oil money, facilitated by loans from international banks, financed a spending boom in Latin America and elsewhere during the 1970s. This led to a rapid increase in foreign debts (Little et al., 1993) in countries which were not in a position subsequently to adjust and service these debts. In due course, significant balance of payments problems emerged when, in 1980–82, real interest rates rose, driven by tight monetary policy in the United States and by a world recession which worsened the terms of trade for many emerging-market economies. These countries rediscovered the truth of what Keynes had maintained 40 years earlier: adjustment to external difficulties requires both good budgetary control and an appropriately competitive real exchange rate (Corden, 1990; Little, 1993). This turned out to be something which many policy-makers in Latin America, and elsewhere, were unable to engineer, and monetized fiscal deficits led to reserve losses, uncontrolled devaluations of currencies and inflation, and difficulties in meeting foreign-currency-denominated debt obligations. Currency and debt crises were triggered more or less mechanically as macroeconomic fundamentals drove reserves down to critical levels, resulting in what has become known as a 'first-generation' crisis.

Although Latin America is most closely associated with the debt crisis of the early 1980s, other countries, including Morocco, were also involved. The crisis placed the IMF at the centre of the world stage in a way which made it more prominent than it had ever been under the Bretton Woods system. The Fund played four roles. First, it offered financial support with stand-by arrangements and other lending facilities. Second, the Fund came to define the broad envelope of resources that a country could be expected to devote to meeting its residual obligations under a debt rescheduling. In turn, the Fund, together with the United States and other bilateral creditors in the Paris Club, pressed creditor banks to reschedule debts and to engage in 'concerted lending' programmes, threatening to provide no support for indebted countries if banks did not cooperate, and, hence, making defaults more likely. Third, the Fund's advice and conditionality, together with that of the World Bank, had significant effects on indebted governments' policies: they were encouraged to undertake growth-oriented structural reforms to escape from their debt problems. Fourth, the Fund's reports and conditionality provided the 'seal of good housekeeping' on the basis of which banks and bilateral creditors could justify rescheduling existing debt and providing new funds.

This use of the Fund, and the broader strategy surrounding it, is usually associated with James Baker, then Secretary of the US Treasury. It was a success only to the extent that it made the financial crisis manageable. The strategy avoided explicit debt reduction and insisted that indebted countries meet their obligations, although over an extended period of time. (This lengthening of the repayment profile did, of course, lead to some reduction in the net present value of debt.) Such an approach was advocated by the governments of major industrialized countries, especially the United States, that were concerned about systemic risks to their own banking systems arising from widespread write-downs of debt. The Fund was criticized in some quarters for agreeing to this strategy and for acting as an 'enforcer' of debt service on behalf of private banks.

A policy shift took place in 1989. Under the Brady Plan, also initiated by the US administration, the Fund and the World Bank provided encouragement and some financial support for debt reduction programmes for those countries (notably Mexico) where major policy reforms were being undertaken. The shift from the Baker Plan to the Brady Plan represented a tilt in favour of debtor countries relative to creditor banks. But this came only after a long period in which these banks were able to rebuild their balance sheets, thereby putting them in a position to weather debt restructuring. The US Treasury induced creditors to grant write-downs to debtor countries by collateralizing the debt that emerged from these restructurings. The Fund backed up this carrot by concluding financing packages with debtor countries before the terms of debt reschedulings had been determined: a practice that came to be known as 'lending into arrears'. This acted as a stick to weaken creditor leverage in the negotiation process, and it also greatly strengthened the role of the Fund in debt work-outs since, during the negotiations, Fund staff came to play a major role in influencing debtor countries' macroeconomic policies.

*4.2 The Mexican 'Tequila' crisis: a 'second-generation' crisis*

The Latin American debt crisis of the early 1980s had been caused by *public*-sector overspending. But in 1994 something new happened. A major financial crisis, caused by the outflow of *private* capital, of the kind which had brought down the Bretton Woods system in 1971 and the European Monetary System in 1992, happened in Mexico. The Mexican crisis was different from the Latin American turmoil of the 1980s in that it was set off not just by fundamental weaknesses, such as unsustainable fiscal and current account deficits, but also by currency mismatches on the public-sector balance sheet. (See Calvo and Mendoza, 1996.) These caused a 'second-generation crisis' in the form of a self-fulfilling currency run. This crisis presented new challenges for the IMF since it marked the first of a series of crises in emerging markets that originated in the capital account, rather than the current account, of the external balance of payments. The IMF was called on to assist Mexico despite the fact that its Articles of Agreement provide it with only limited jurisdiction over capital account issues.

Mexico had implemented a comprehensive reform programme in the early 1990s, which included financial liberalization and the completion of the North American Free Trade Agreement (NAFTA) in 1993. This led to a surge in investment financed mainly by foreign capital flows. The result was a large (real) overvaluation of the peso and a

very large current account deficit. Initially, the government maintained prudent fiscal policy. But during 1994 many began to question the sustainability of the exchange rate, the fiscal position and current account deficit. By December 1994 there was a massive reversal of capital flows, and the peso plummeted. The consequences for Mexico were severe: inflation rose from 7 per cent in 1994 to 35 per cent in 1995; and GDP fell by 6.2 per cent in 1995 compared with a growth rate of 4.4 per cent in the preceding year.

The pain inflicted on Mexico by private investors led to a view that pegged exchange-rate regimes are unviable everywhere, not just in advanced industrial countries. (Mexico had a 'crawling peg' at the time.) And in Mexico there was a new emerging-market feature. Much of the Mexican government's debt was denominated in US dollars (for example, the '*tesobonos*') because of the difficulty and high costs of borrowing in local currency; much of the government's revenue stream, by contrast, was peso-denominated (although oil revenue was denominated in dollars). This mismatch meant that the collapse of the peso led the government to the verge of default in early 1995.

The Fund played a critical role in stabilizing the crisis. In particular, drawing on financing from bilateral creditors, it coordinated assistance, mainly from the United States, that totalled more than five times Mexico's quota entitlements at the IMF. After a significant real devaluation of the peso and fiscal correction, exports rebounded, the economy grew, although only slowly, and Mexico earned enough foreign exchange to repay the exceptional financing that had been provided to it during the crisis.

Some subsequent analyses (see, for example, Calvo and Goldstein, 1996) were critical of the IMF's role in both surveillance and in crisis management for Mexico. But the arguments cut both ways.

On surveillance, it was claimed that IMF reports prior to the crisis placed insufficient emphasis on the vulnerabilities of public-sector and financial-sector balance sheets to the possibility of a run on the currency. Some authors argued that the Fund should have been more frank in conveying its views on macroeconomic and exchange-rate policy to its members, and that it should publish these appraisals. But there may well have been inadequate provision of information by Mexico to the Fund, as well as to the public. In particular, it appears that incomplete data may have been provided on official international reserves and liabilities (although the Mexican authorities disagreed with this claim when it was made). As a result, following the Mexican crisis, the Fund began a drive to get countries to sign on to transparency standards, such as the Fund's Special Data Dissemination Standards (which were established in 1996; see Fischer, 2004, p. 127). Additionally, the Fund began the practice of publishing Board documents, except when the authorities of a country objected. But this heightened focus on transparency left the Fund unclear on whether it should assist countries confidentially to prevent crises or spur corrective action by bringing bad news to the market. Given the sometimes self-fulfilling mechanics of second-generation currency crises, solving this dilemma is critical in defining the future role of the Fund in crisis prevention.

On crisis management, no clear conclusions emerged, either. *Ex post* it appeared that the private sector should have been prepared to lend short term to the Mexican government in the way that the IMF and the United States did. Overcoming such a market failure is surely a role of the IMF and national governments, and giving the

IMF the capacity to provide such big loans seemed important to many observers. From this experience, Sachs (1995) concluded that the Fund should be given an explicit international lender-of-last-resort capacity, well beyond that formally possible under its 'credit-union' status, so as to enable it to be ready to respond forcefully and quickly to emerging crises, as it had done in the Mexican crisis. (See also Fischer, 1999.) With such firm IMF action, currency crises could be contained as liquidity crises rather than becoming solvency crises. Indeed, it appears that the combination of large-scale IMF financing, combined with significant adjustment by the authorities, prevented the development of a solvency crisis in Mexico. However, some authors began to warn that, if the IMF always acted as a lender of last resort in the face of crisis, then this might create moral hazard on the part of lenders to emerging markets, who might expect to be able to lend virtually risk-free with any possibility of default prevented by IMF action. (The Fund-led bailout of *tesobonos* holders strengthened these fears.) These critics suggested that efforts be made to make sovereign debt rescheduling easier and more orderly (Eichengreen and Portes, 1995), thereby containing the threat of creditor moral hazard.

*4.3 The Asian financial crisis of 1997–98: the 'third generation' of crises*

Two and a half years later these issues re-emerged in Asia, in a crisis which interrupted a long period of sustained economic growth financed by exports and foreign capital inflows. Unlike the earlier Latin American debt crisis, or even in Mexico, fiscal profligacy played no *explicit* part in the East Asian crisis. But there were two other main policy failings. (See Bluestein, 2001; Corbett and Vines, 1999a; 1999b; Corbett, Irwin and Vines, 1999.)

First, much more than in Mexico, an under-developed financial system and over-protected financial sector in some Asian economies meant that the private sector had to rely on borrowing, rather than equity issuance, to raise investment funds. As a result, firms became highly leveraged, but banks continued to lend because they were underpinned by *implicit* government guarantees. When growth slowed, as it first did in Thailand in 1996, and then in other East Asian economies, these banks were exposed to the inability of borrowers to repay loans.

Second, a further difficulty arose, as so many times before, from the existence of fixed exchange-rate systems in some East Asian economies, but with a new twist. Banks financed much of their domestic corporate lending by borrowing in foreign exchange from abroad, often at shorter maturities than those employed when they lent onwards in domestic currency. Very little of this borrowing was hedged as a result of the implicit guarantee on the exchange rate. As noted in the previous paragraph, the financial sector was already in difficulty after the initial slow down in growth in 1996. Currencies fell in mid- to late 1997 because of foreign investors' concerns about these difficulties; as a consequence, widespread bankruptcies and potential bank failures loomed because of the unhedged foreign-currency obligations. Fear grew that fiscal systems would be unable to bear the cost of large-scale bank rescues (Irwin and Vines, 2003).

The East Asian debacle marked the advent of 'third-generation' crises in which currency crises and banking crises are intimately intertwined – situations in which

vulnerabilities in the private balance sheet can quickly translate into a public debt crisis.

As in Mexico, the Fund played a large part in resolving the crises. The IMF moved quickly to lend very large sums to Thailand, Korea and Indonesia. Nevertheless, there has been widespread criticism of the Fund's behaviour before and after the crisis. (See, for example, Stiglitz, 2002.)

Two difficulties must be acknowledged in the Fund's *crisis prevention* work in East Asia. First, the Fund may have underestimated the risks associated with capital account liberalization. Second, the Fund may not have been firm enough in warning of the difficulties inherent in maintaining a fixed exchange-rate peg. Nevertheless, Thailand, for instance, was warned privately by the Fund several times in the year leading up to the 1997 currency crisis. The Fund, like some private-sector analysts, saw problems looming in Thailand, but its advice was not heeded.

Concerning the Fund's work on *crisis management*, there are three points to consider.

First, as the Fund has acknowledged in both its own reviews of the East Asian crisis and in the evaluations performed by its Independent Evaluation Office (IEO) (IMF, 2003), its programmes may have placed too much emphasis on tightening budgets in countries that were already running prudent fiscal policies. Stanley Fischer, then the Fund's First Deputy Managing Director (FDMD), argues, however, that this approach was driven by a need to boost government savings to support the current account and provision for the impending cost of bank restructurings. (See Fischer, 2004.) Furthermore, the credibility of an adjustment programme at a time of crisis may hinge on policy erring towards being too tight, in order to send a clear signal to markets. Once the scale of the economic downturn became apparent in East Asia and current account balances improved, Fischer argues that the Fund programmes shifted to addressing structural problems. (See also Corden, 1999; Boorman et al., 2000.)

Second, monetary policy was also tightened in an attempt to defend currencies. There is an inevitable trade-off between raising interest rates in order to moderate exchange rate depreciations and lowering interest rates so as to ease the stress on both the banking system and on corporations that depended on domestic credit. Stiglitz (2002) argues that the tightening was too forceful. However, it does appear that this tightening was essential in order to stem capital flight. Nevertheless, this tightening was not followed by a concerted move to an inflation-targeting regime of a kind that might have allayed concerns of further depreciation. Hence, pressure on the region's currencies continued. And rather than stimulating recovery, these depreciations proved contractionary, at least initially, owing to their effects on external debt burdens. (See Krugman, 1999.)

Third, the Fund did not have a mandate to declare 'standstills' on external debt payments during the crisis. In corporate bankruptcies, standstills force creditors to share in the burden of crisis and agree to reasonable debt reschedulings. In the context of a currency crisis, a standstill mechanism would similarly 'bail in' foreign private-sector creditors and then make reschedulings possible to reduce debt to sustainable levels. The fact that a standstill was not imposed in Thailand, Korea or Indonesia enabled creditors to race to get their assets out of these countries. Negotiations with foreign creditors to

Korea and Indonesia did ensure some rollover of existing short-term lending, with effects similar to those that might have resulted from standstills. In both cases, however, negotiations were pursued too late and without sufficient coordination to maximize their impact (though they did stave off collapse in Korea). The only comprehensive brake on external payments was that imposed in Malaysia through the implementation of capital controls rather than a standstill by the government of Prime Minister Mahathir bin Mohamad in late 1998, a move that contravened the Fund's advice. But this was done only after substantial capital outflows from Malaysia had already taken place.

Because the Fund lacked a mandate to impose standstills, it lent countries money in an attempt to allay the concerns of foreign creditors and to stem capital flight. Given the scale of the external capital-account movements in these countries, the size of IMF financing packages soared, especially after it became clear that smaller lending programmes would be unlikely to produce adequate results. In the case of Korea, the authorities of the IMF's large shareholder governments, notably the United States and Japan, also made a key decision to pursue a debt rollover plan and to exert moral suasion on creditor banks. These banks presumably realized that the alternative would have been partial default. The IMF played a useful role in facilitating communication among the different actors, in providing information, and in certifying that the policies to be pursued by the Korean authorities were appropriate. The IMF's Independent Evaluation Office writes, 'No single national government, nor any private sector institution, could have played this role as effectively' (IMF, 2003, p. 115).

Although the Fund's work in Korea showed that the IMF could effectively manage a debt workout, its conduct elsewhere in the East Asia crisis had the effect of shifting the balance of power in debt workouts back toward creditors. IMF programmes did *not* reduce the debt overhang in Indonesia and Thailand. Instead, governments rescued banks and corporations by shifting their debt to the public balance sheet. Taxpayers in these countries still bear the burden of this debt. Rather than 'bailing in' private creditors, the Fund's handling of the crisis in these countries may have provided creditors with an even bigger bailout than they might have expected under the terms established in the 1990s' Brady Plan.

Partially out of dissatisfaction with this result, Anne Krueger, who followed Fischer as the Fund's FDMD in 2001, proposed a bankruptcy or standstill procedure for countries, the 'Sovereign Debt Restructuring Mechanism' (SDRM) (Krueger, 2002). The US Treasury and financial markets both opposed this proposal out of a concern it would create unrestrained debtor moral hazard. Under what came to be known as the 'Taylor Doctrine' (after John Taylor, then US Treasury Under Secretary for International Affairs), the US government argued that countries should be left on their own to negotiate with their creditors. But this is only feasible when the number of external creditors is small, which for most countries has not been the case since the 1980s when external borrowing was provided mainly under loans from banks. To help remedy this problem, the US supported the introduction of 'collective action clauses' (CACs) in bond contracts with commercial creditors. These clauses prevent rogue creditors from holding out in restructuring negotiations in order to extract a premium from the bond issuer; they work by enforcing a restructuring if a pre-specified

minimum proportion of creditors have agreed to its terms. CACs do not, however, provide a framework to guide the allocation of losses between borrowers and lenders, which is necessary in any restructuring. In the absence of a clear means of sharing these losses, it may prove impossible to renegotiate debt owed to commercial creditors. When faced with debt-servicing problems, debtor countries may then decide to borrow from official sources (including the IMF, whose debt is senior to other external liabilities and not reschedulable) in order to repay private sector creditors, as happened in Korea, Thailand and Indonesia. Since private-sector creditors are likely to believe that this will happen, the Taylor doctrine's approach, even when coupled with CACs, might promote creditor moral hazard, something which has been feared ever since the Mexican crisis. Thus, although the Taylor doctrine's approach has the virtue of minimizing debtor moral hazard, it appears to go in the opposite direction by promoting creditor moral hazard.

*4.4 Default: the Russian and Argentine crises*

*Russia*. The fall of the Berlin wall in 1989 and the dissolution of the Soviet Union in 1991 enabled the IMF at last to become a (nearly) universal institution. In three years, membership increased from 152 countries to 172, the most rapid increase since the influx of African members in the 1960s. The IMF supported programmes in most former Eastern Bloc countries and newly independent ex-Soviet Republics to help ease the transition to a market economy. The contribution the IMF made to the speed and relative smoothness of this transition is, perhaps, one of its most singular and least-heralded achievements.

Russia, however, got off to an inauspicious start under the first stand-by arrangement with the Fund in 1992. The IMF encountered intense difficulties in influencing the Russian leadership (Odling-Smee, 2004). GDP fell for several years under the IMF-supported combination of macroeconomic stabilization and industrial restructuring. Although the IMF can claim credit for helping to instil some monetary discipline by the mid-1990s, the process took time, foreign direct investment remained low, tax collection was poor, and the fiscal deficit remained large. Growth in real GDP did re-emerge by 1997. But, following the onset of the East Asian crisis, the ruble came under speculative attack in November 1997. Pressure on the ruble was compounded by foreign investors' attempts to hedge their ruble holdings, as well as by a drop in the price of oil, which accounted for about one-third of Russia's foreign-exchange inflows.

Russia sought additional IMF financing in early 1998, but agreement on the terms of a new programme could not be reached owing, in part, to a failure by the Russian authorities to secure an increase in fiscal revenue. As a result, foreign investors began to unload Russian assets and about US$4 billion fled the country in the summer of 1998. By the time additional IMF financing was agreed in July 1998, fears of a devaluation led to such a pronounced sell-off of Russian securities that the authorities were forced to devalue the ruble and halt payments on both domestic and foreign debt.

Although the Fund is routinely criticized for providing cover for private capital flight from Russia in the first half of 1998, private investors who maintained faith that the Fund would rescue Russia sustained even greater losses when the ruble was devalued. This was perhaps the largest case to that point where the Fund stepped away

from a floundering member, declared a solvency crisis, and let private creditors sustain substantial losses. It marked a different approach to the challenge of balancing creditor and debtor interests from that which the Fund had adopted in East Asia. And in some ways it set a precedent for the Fund's handling of the Argentine crisis in 2001.

*Argentina*. After a sustained period of hyperinflation in the 1980s, Argentina decided in 1991 to peg its currency, the peso, to the US dollar under a quasi currency-board regime at a one-to-one parity. Although the Fund cautioned that Argentina had neither the fiscal discipline nor the robust export sector needed to sustain such a system, it went along with the authorities' plans and supported their macroeconomic programme under a series of lending arrangements. By the late 1990s, Argentina was widely hailed as a model of successful economic reform as the rate of inflation fell to single digits and growth increased. In addition, the economy had successfully weathered the global turbulence caused by the East Asian crisis of 1997–8, and the Russian crisis of 1998.

But the seeds of the problems identified by the Fund back in the early 1990s were beginning to bear fruit by the end of the decade. Fiscal policy remained insufficiently tight owing to the lack of effective central government control on provincial borrowing, and this stimulated domestic demand for imports. Argentina's export sector remained too small to finance these imports, and its real exchange rate made its goods uncompetitive on regional and international markets. As a result, Argentina chose to borrow substantial amounts in US dollars to finance its imports. Brazil's decision to float the real in 1999 in response to pressure from the Russian crisis made it even harder for Argentina to compete under its quasi currency-board regime. The Argentine authorities allowed the peso to float in January 2002, and it quickly collapsed from parity with the US dollar to an exchange rate of nearly 3.9 to the dollar in June 2002. Output fell sharply, inflation reignited, the government defaulted on its debt, and the banking system was largely paralysed.

The Argentine debacle rightly cast several doubts on the Fund's conduct of both crisis prevention and crisis management in emerging markets. At the outset of the 1990s, the Fund proved incapable of resisting Argentina's arguably doomed effort to impose its quasi currency board. Subsequently, the Fund endorsed Argentina's exchange rate peg in a series of programmes through the 1990s that coincided with an accumulation of macroeconomic vulnerabilities. When the regime became unsustainable in 2001 (or earlier), the Fund maintained lending until the end of that year in an attempt to save the peg. After the crisis, the Fund resumed lending to an insolvent Argentina in 2003 at the behest of the Executive Board, even although misgivings were expressed by the Fund staff. IMF lending ceased again later in 2003 and Argentina pursued an aggressive 'take it or leave it' strategy with private creditors. The Argentinean authorities achieved a roughly 75 per cent write-down on the country's defaulted foreign bonds, while leaving nearly US$20 billion in unexchanged bonds in default (IMF, 2005a).

The Fund's experience with Argentina demonstrates at least four things. First, it can be very difficult for Fund staff to resist Executive Board pressure to support a country with IMF lending, either when inappropriate policies are being pursued (for example, the creation of the quasi currency board) or when a country is insolvent (as Argentina was by 2003). Second, the Fund has sometimes found it just as hard as its members to take a stand against an inappropriate fixed-exchange-rate regime. Third, the absence of

any international standstill process or debt restructuring mechanism makes it difficult and time consuming to reconstruct a financial system and to reach a balanced solution with creditors once a crisis has occurred. The Taylor doctrine has not worked out wholly as planned. Fourth, once damaged, the quality of the policy dialogue between the Fund and its members is difficult to restore. Since the crisis, Argentina's policies have appeared unsustainable: Argentina has contrived to keep its exchange rate at a level at which its exports seem to be excessively competitive, while relying heavily on high international primary commodity prices to sustain its balance of payments. These policies do not seem consistent with the world envisaged in the second amendment of the Fund's Articles, a world in which the Fund exercises firm surveillance over member countries' policies in its role as steward of the international financial system.

*4.5 Conclusions*

The capital account crises of the 1990s and 2000s represent a new chapter in the Fund's history: they mark a distinct shift from the Fund's previous bread-and-butter work of dealing with current account crises. These capital account crises created new challenges and strains on the Fund – some of which it responded to well, some less so.

On *crisis prevention* the Fund has learned much. After the Mexican crisis it promoted regulatory reform, increased transparency, and better monitoring in emerging market economies. The Fund's Articles prevent it from pronouncing on countries' particular choice of exchange-rate regimes. But in its policy advice the Fund has made clear that the trilogy of floating exchange rates, carefully sequenced liberalization of capital accounts and financial systems, and inflation targeting can work well (Blejer et al., 2001; Corden, 2002; Batini, Kuttner and Laxton, 2005); by contrast, the Fund has given clear advice about the difficulties faced by fixed exchange-rate regimes. The Fund has also attempted to reinvent itself as a lender of 'first resort' through the creation of contingent or 'pre-approved' lending facilities aimed at crisis prevention. These lending windows would provide members with an added incentive to pursue sound policies and a signalling framework under which they could commit to these policies. But the Fund's first effort in this direction – 1999's Contingent Credit Lines (CCL) – expired in 2003 after four years without use, owing to somewhat stringent qualification criteria, less than full automaticity in disbursements, and concerns amongst members that a request for a CCL might send a negative signal to capital markets. New effort was invested in the design of such an instrument, initially called the Reserve Augmentation Line (RAL), during 2006–07.

On *crisis management*, much work has been done to understand better how to construct, balance and sequence macroeconomic policy restraint at a time of crisis. The Fund has developed a detailed debt sustainability framework and complemented its traditional analysis of financial flows with a 'balance sheet approach' to analysing stock imbalances, so as to enable it to understand the financial vulnerabilities of countries. This tool was designed to help Fund staff draw a clearer distinction between liquidity crises and solvency cases. (On this see Irwin and Vines, 2005; Cohen and Portes, 2004; Portes, 2004.) But from the early 1980s onward, the three generations of crises outlined above also threw into sharp relief the problem of moral hazard arising from IMF lending. The need to balance better debtor moral hazard and creditor moral hazard became

one of the key challenges facing the Fund in the design of its lending facilities and its accompanying policy responses to crises. This article has highlighted the manner in which the Fund has occasionally oscillated between favouring creditor interests and favouring debtor interests, in an attempt to balance these interests in an acceptable way.

The Fund's experience with crisis management in the 1990s revealed difficulties with Fund conditionality. By then the conditions attached to Fund loans had grown far beyond what had earlier been thought necessary to ensure adequate macroeconomic adjustment, and came to include substantial structural conditionalities. Some of these concerned macroeconomic issues of proper concern to the Fund. But there was also an explicit concern with a range of microeconomic reform issues, and, even more broadly, with poverty-reduction questions. Many observers, including Arriazu, Crow and Thygesen (1999), IFIAC (2000) and Williamson (2000), have questioned the wisdom of this policy creep, although it should be said that, in some cases (for example, poverty reduction), the spread of IMF conditionality reflected the concerns of member countries rather than an attempt by the Fund to expand its mandate. Following member country dissatisfaction with the comprehensive conditionalities included in their programmes (Indonesia's programmes in the late 1990s are particularly relevant cases), there has been much work at the IMF since 2000 on streamlining conditionality, and on pulling back from a range of concerns about structural issues that are not deemed 'macro critical'. This led to a careful restatement during 2002 of the principles governing the IMF's design and implementation of conditionality, with a view to ensuring that the conditions attached to IMF lending focus only on policies essential to the macroeconomic viability of Fund-supported programmes. (See IMF, 2002a; Boughton and Mourmouras, 2004.)

At the time of the preparation of this article (2007) there was a lull in the frequency of crises, and a significant decline in the volume of Fund lending. The Asian, Russian and Argentinean borrowings which originated in the crises described above had all been repaid. There is a striking parallel here with the end of the 1980s, when the Fund's stock of outstanding loans to emerging markets was also quite modest. At that time, the Latin American arrangements that had originated in the crisis years 1980–83 had been repaid. But, just as then, risks remain; the international community must remain engaged in the task of ensuring that the Fund is prepared to respond to and manage crises when they occur.

Dissatisfaction with the Fund's crisis management in the 1990s and early 2000s cast a long shadow over the Fund's relations with many emerging-market economies, which may have some consequences. A number of East Asian countries, over the ten years following the East Asian crisis, accumulated in excess of a trillion US dollars of reserves. This massive reserve accumulation reflected a persistent excess of saving over investment across these economies, which may, at least in part, represent a conscious choice to amass reserves as a form of self-insurance against future crises. These countries went about a pooling of some of these reserves into a common fund, a process which began in 2000 when ASEAN, Japan, China and the Republic of Korea agreed to set up a bilateral currency swap scheme known as the Chiang Mai Initiative. There were some suggestions that this might one day form the basis of an Asian regional

alternative to the IMF that would be designed to help these countries to co-insure and spread risks. But taking this step would require difficult decisions by these countries in order to make surveillance between the pool's members effective and enforceable. And such a common pool of reserves might also create its own form of moral hazard if it were to encourage countries to take excessive risks with foreign borrowing.

## 5 The IMF and low-income countries

Until the mid-1970s, the Fund's work in its role as coordinator and monitor of the international monetary system was concerned mainly with monetary, exchange-rate and trade issues. To the extent that the IMF also functioned as a credit union for countries in balance of payments difficulties, its lending focused on the provision of short-term, self-liquidating loans to buttress central banks through temporary balance of payments difficulties. The Fund's cornerstone principle of equal treatment of member countries dictated that finance to low-income countries was provided largely under stand-by arrangements on the same terms as those approved for emerging markets and industrialized countries. The oil crises of the 1970s, however, made it increasingly clear that intractable structural issues in many low-income countries needed to be tackled if balance of payments difficulties were to be addressed. As a result, the 1970s saw a lengthening of the average maturity of stand-by arrangements in both emerging markets and low-income countries, accompanied by the advent of lending on concessional terms, with lower interest rates, to low-income countries. This created some tension between the Fund's essentially monetary character and its deepening role in the provision of longer-term resources in support of broad macro-economic adjustment in developing countries.

In order to provide member countries with more breathing room to enact structural economic reforms, the Fund created a series of new lending instruments from the mid-1970s onward. The first amongst these, the Extended Financing Facility (EFF), provided greater financing and longer maturities than traditional stand-by arrangements, but its terms were not concessional. The Fund's Articles of Agreement did not provide for the use of IMF resources for concessional lending to a subset of the Fund's membership, and the EFF's market-linked interest rates were identical to those of other Fund arrangements. An EFF did, however, typically carry more stringent conditionality than a stand-by arrangement in response to concerns that the EFF's greater financing implied a need for greater adjustment.

The obstacle to financing concessional lending posed by the Fund's Articles was overcome in the 1970s by the solicitation of donor funds and the sale of a portion of the IMF's gold. Concessional IMF lending began under the 1975 Oil Facility Subsidy Account, in which contributions from 25 countries were used to reduce the interest cost of borrowing from a Fund facility set up to assist countries deemed to have been most severely affected by the sudden rise in oil prices. In the following year, the IMF created a Trust Fund for all low-income countries out of profits from the sale of a portion of the Fund's stock of gold. The Trust Fund offered long-term low-interest loans to low-income countries from 1976 until its resources were fully committed in 1981. Borrowing under the Trust Fund was similar to financing under the first credit

tranche: in order to obtain financing, low-income countries had only to demonstrate a balance-of-payments need and explain the efforts they were taking to reduce it.

These new financing windows provided concessional loans to developing countries, but it was feared that the weak conditionality attached to these loans did not induce sufficient adjustment (Boughton, 2001). In the early to mid-1980s prices for many primary commodities collapsed, and several developing countries faced new external balance of payments challenges. The Fund moved to reinvigorate its concessional lending by using the repayments of Trust Fund loans to finance a new round of concessional credit under what, in 1986, came to be known as the Structural Adjustment Facility (SAF). The SAF marked a determined attempt by the Fund to integrate concessionality with conditionality. In part, this twinning of concessionality with conditionality allowed the Fund to lobby for new donor loans and grants, which expanded the SAF some threefold into the Enhanced SAF (ESAF) in 1987.

Boughton (2001) contends that the ESAF became one of the IMF's great success stories, as it allowed the Fund to send billions of dollars to the world's poorest counties on concessional terms with longer maturities than was possible under previous IMF facilities. (See also Tarp, 1993.) The ESAF also had a catalytic effect on lending from other official creditors, and IMF collaboration with the World Bank and the regional development banks, as well as with, inter alia, the UN, UNICEF, UNDP and bilateral donors, all appeared to improve under the ESAF process (Boughton, 2001). In addition, IMF technical assistance to many developing countries on monetary, fiscal, and trade policy, as well as debt management, also expanded substantially in order to help countries achieve their programme commitments. This increase in technical assistance has been very valuable.

Despite these gains, and even although the ESAF was technically distinct from the Fund's general resources, some critics have charged that the ESAF marked an unfortunate departure from the Fund's monetary focus. Others have questioned the strict conditionality on adjustment agreed under ESAF-supported programmes, especially because some of the structural conditions have appeared to intrude on the traditional territory of the World Bank. In reply it might be said that this has happened partly because the Bank has not proved capable of devising appropriate macroeconomic conditions for its own loans. (See Gilbert and Vines, 2000.)

Despite the Fund's efforts – both to revive its concessional lending in 1986 and 1987, and to increase its accompanying technical assistance – it was clear by 1988 that many low-income countries would find it impossible to grow without debt relief. Under the auspices of the Paris Club of bilateral creditors, a series of progressively more concessional refinancing terms for bilateral debts were agreed from 1988 onward, for both emerging market, and relatively poor, indebted countries. Nevertheless, even with this bilateral debt relief, many low-income countries had trouble meeting the payment obligations on their stand-by arrangements and EFFs. But the absence of a serious lobby of private creditors (most low-income countries' external debt was owed to the Paris Club and other public creditors) may have delayed efforts to find a comprehensive solution to the debt problems of developing countries until the late-1990s.

By the 1990s, the Fund's engagement in low-income countries had become the target of a rising chorus of concern. Some civil society organizations and academics, as

well as some low-income governments themselves, contended that IMF conditionality and programme design in low-income countries tended to prioritize adjustment over poverty reduction, growth, and income distribution concerns. This criticism is summarized by Easterly (2005). It arose despite the fact that the Fund has been helping to produce, in many low-income countries, a marked stabilization in macroeconomic indicators, and in some cases the beginning of sustained periods of growth. In response to critics' concerns, and in a further step in the evolution of Fund lending, IMF Managing Director Michel Camdessus advocated in the mid-1990s a fresh model of engagement with low-income countries in which there would be a renewed role for the Fund in reducing global poverty and in promoting high-quality growth in developing countries.

This new strategy featured three main elements. First, along with bilateral donors and other international financial institutions, the Fund recognized that catalysing growth in low-income countries would require more profound debt relief, including treatment of previously unrescheduled multilateral concessional debt. The 1996 Heavily Indebted Poor Countries' (HIPC) Initiative represented the concerted efforts of the international community to address the external debt overhang in poor countries; the Initiative was later enhanced in 1999 to provide deeper and faster debt reduction. The HIPC Initiative was novel, particularly in that debt relief was explicitly tied to plans to spend debt-service savings on poverty-alleviating social expenditure. From 1999, these plans were articulated in a country-based Poverty Reduction Strategy Paper (PRSP). This approach, initiated by the Fund in conjunction with the World Bank, formed the second prong of the Fund's renewed engagement with low-income countries. The PRSP approach aimed to provide a clear country-owned link between national policy frameworks, donor support, and development outcomes. The PRSP approach also dovetailed neatly with the United Nations' Millennium Development Goals (MDGs). These goals were articulated at the UN Millennium Summit in 2000 and were centred on halving global poverty by 2015. The PRSPs were also intended to form the basis of the targets and policy conditions in programmes supported by the IMF's Poverty Reduction and Growth Facility (PRGF). This was the successor in 1999 to the ESAF and formed the third element of the Fund's new approach to low-income countries.

The results of these initiatives by the early 21st century were mixed. Reviews of the PRGF by IMF staff in 2002 (IMF, 2002b) and by the IMF's IEO in 2004 (IMF, 2004b) found that PRGF-supported programs had become more accommodating to higher public expenditure, in particular pro-poor spending. Nevertheless, a review of PRGF programme design by the IMF Executive Board in September 2005 (IMF, 2005c) found that per capita income and growth rates remained low despite some improvements in a range of macroeconomic indicators. More recently, the IEO found in its evaluation of Fund engagement in sub-Saharan Africa (IMF, 2007b) that the PRGF and PRSP approaches had not had a significant positive effect on catalysing new aid flows. This is despite the fact that commitments to increase such flows were made in 2002 under the 'Monterrey Consensus' and at the Gleneagles G8 summit in 2005. The IMF's Spring 2007 *Regional Economic Outlook* noted, however, that Sub-Saharan Africa's growth performance since 2004 had been the best in more than three decades (IMF, 2007d). In sum, the impact of the PRGF and PRSP on aid and spending in low-income countries remained inconclusive, but their growth effects appeared increasingly positive by 2007.

The advent of the HIPC Initiative, the PRSP and the PRGF together intertwined the work of the IMF and World Bank in developing countries to an unprecedented extent. The Multilateral Debt Relief Initiative (MDRI) agreed at the Gleneagles G8 Summit in 2005, and which provided a framework for the write-off of nearly all remaining HIPC-country debts to the IMF, World Bank and African Development Bank, represented a major step forward in this collaboration. While the MDRI drew a welcome line under the multilateral debt relief process, it left several questions about the next phase of IMF and World Bank support for low-income countries unanswered. Having written off so much concessional debt, the MDRI implied that future multilateral support for low-income countries should be provided only as grants, not loans. The source of financing for such grants remained unclear. And in some cases, financing, whether by grants or loans, may not be the most crucial contribution that the international financial institutions could make to development. The Fund's 2005 Policy Support Instrument (PSI), essentially a 'no money' programme, acknowledged that Fund macroeconomic advice, rather than short-term balance of payments financing, might be a valuable channel of support for developing countries. These matters have been complicated by the growth of 'South–South' flows in development assistance from new donors such as China and Brazil. These flows have raised doubts about the future necessity of concessional financing from the Bretton Woods institutions. But they have also called into question the conditionality that comes attached to IMF and World Bank money. Such financing from non-traditional donors could also complicate future debt restructurings, should they prove necessary, since most new donors have not been members of the Paris Club.

Throughout this section we have noted the latent tension between the Fund's monetary character and its long-term support for low-income countries. This tension is heightened by the intertwining of the work of the Fund and the World Bank, which we have just reviewed. The report of the external review committee on Bank–Fund collaboration (IMF, 2007c) provided some suggestions on strengthening Bank–Fund collaboration, while reducing overlap between the two institutions.

## 6 The Future of the IMF: next steps

In mid-2004 the Fund's Managing Director, Rodrigo de Rato, launched a review of the role of IMF in light of the challenges posed by a changing and increasingly complex global economic system. Stemming from this review, De Rato presented the aforementioned Medium-Term Strategy for the Fund (IMF, 2005b) to the World Bank–IMF Annual Meetings in September 2005, and shortly thereafter followed up with a plan for the Strategy's implementation (IMF, 2006b). The plan focused on specific proposals to ensure that the Fund:

- provides more effective surveillance and better monitoring of policies in advanced economies, with a renewed emphasis on exchange rates;
- provides better monitoring of emerging markets economies, re-explores financing mechanisms to help prevent crises, and reconsiders issues regarding capital account liberalization;
- enhances the role of IMF in low-income countries, and sharpens its focus;
- reforms IMF governance, particularly country representation; and

- restructures the IMF's own budget, including by broadening the Fund's income base, and its management practices.

The plan also expressed an intention to expand the role of the IMF as a provider of technical assistance and training, while improving Fund communications and transparency to ensure that the Fund would play a more central role in global policy debates.

The Fund's Medium-Term Strategy is a clear response to the three dominant tasks it has assumed following the collapse of the Bretton Woods system of fixed exchange rates in 1971, tasks which we have reviewed in Sections 3–5 of this article. But if the Fund is to be able to act effectively in relation to these tasks it will need to have: (i) a better system of governance; (ii) a more secure and robust source of income so that it can cover its operating expenses; and (iii) a larger stock of resources to lend for crisis prevention and resolution. We conclude this article by briefly discussing these three issues. (See also Lane, 2006.)

### *6.1 Governance*

The first subsection of the Fund's Articles of Agreement made clear that its founding purpose was 'to promote international monetary co-operation through a permanent institution which provides the machinery for consultation and collaboration on international monetary problems'. At the time of the Fund's creation, most countries stood a reasonable chance of alternating between being a creditor to and borrower from the Fund over time. Since then, the ranks of creditors and borrowers have diverged as industrial countries have stopped using IMF financing, a role which has instead been filled by emerging market economies and low-income countries. A number of reformers such as Woods (2006) argue that the Fund's capacity to facilitate solutions to international monetary problems depends on the Fund's decision-making structure being made more reflective of the interests and voices of the emerging markets and developing countries which borrow from it, and which see their public policy frameworks at least partly determined by Fund conditionality. The demand for such reform is bolstered by the fact that the relative distribution of quotas, which determine the voting power in the Fund, has become separated from the relative economic (and political) weights of many emerging markets in the global economy. In addition, the relative power of basic votes, which were intended to provide some measure of fairness to poorer countries, has been substantially eroded relative to the contribution of quotas to voting weights at the Executive Board. The ad hoc provision of increased quota shares to China, Korea, Mexico, and Turkey in 2006 under the Fund's Medium-Term Strategy was a first step toward realigning voting power in the Fund with emerging markets' growing share of the world economy; further steps will be more difficult since increased voting shares for some countries will inevitably mean painful decisions to reduce the shares of others. It may, however, be possible for countries to change the way in which the 24 chairs on the IMF's Executive Board are allocated in order to compensate partly for changes in relative voting shares.

Changing the Fund's voting structure would not in and of itself alter the way in which the Fund operates, suddenly making it better able to deliver on the objectives set out in its 2005 Medium-Term Strategy. De Gregorio et al. (1999); King (2006a, p. 12);

Dodge (2006a; 2006b) and Kenen (2006) have all argued, however, that parallel changes in the Fund's governance arrangements might help the Fund in its push towards these objectives.

One proposal would put the responsibility for the delivery of improved policies more firmly in the hands of the management of the IMF. Up to 2007, the Executive Board of the Fund had involved itself in day-to-day reviews of Article IV reports, approved all lending decisions, and reviewed the design of the Fund's lending programmes. Stepping back from this activity would enable Directors to pay proportionately more attention to strategic issues. That would move the governance structure of the Fund closer to the relationship between management and advisory boards that one sees in the private sector, where non-executive directors bring dispassionate external views to broad questions of corporate operations and strategy, and clearly delegate day-to-day operations to management.

Evolution in this direction could strengthen the accountability of the Managing Director and his Deputies. In one version of this type of arrangement, all of the Managing Director, the Deputy Managing Directors, and Department Directors would report on a regular basis to the Board, but Executive Directors would be more removed from many of the day-to-day decisions of the institution. Doing this could have an effect – even if only implicit or indirect – on the Fund's ability to function better in its pursuit of more dispassionate surveillance. It might also lead to more effective crisis prevention and resolution through a careful balancing of debtor moral hazard and creditor moral hazard in Fund lending; and also to a clearer focus in the Fund's work with low-income countries.

A move to a non-resident Executive Board would draw a clearer line between the work of Directors and management. Such a move would leave the Managing Director in control of the execution of the Fund's work since the Executive Directors would give only part-time oversight and direction. Making this change would take the governance of the Fund closer to Keynes' original vision. (See King, 2006a.) Directors would be the senior public servants that steer policy in their national capitals, and not, as in 2007, their proxies resident in Washington. In contrast with 1946, the ease of modern travel makes a non-resident Board, with meetings some six to eight times a year, entirely feasible. Any move in this direction would, however, need to ensure that the nexus of communication between capitals, which the Board currently provides, is preserved in some other way.

### *6.2 Income*

In May 2006, the Managing Director established a committee (the 'Crockett Committee'), chaired by a former General Manager of the Bank for International Settlements, Andrew Crockett, to study options for sustainable long-term financing of the IMF. The Committee's report, released on 31 January 2007 (IMF, 2007a), argued that the IMF's current funding model was unsustainable and that a more diversified income stream needed to be developed in order to guarantee the institution's financial future.

The IMF's revenue stream had been primarily based on income derived from its lending for crisis resolution (IMF, 2007a, Annex 2, p. 2). This financing mechanism was not entirely appropriate, because, as Crockett said during the press briefing to

launch the Committee's report, 'it's a concentrated income source ... It's volatile, because when the Fund is lending a lot ...it generates large resources. When the Fund is not lending, it doesn't generate resources.' In a low-lending environment, as existed in the early 21st century, the Fund's income model appeared untenable over the longer term; in the shorter term, it could also be inconsistent with sound incentives to minimize moral hazard in Fund lending.

The Committee considered some alternative sources of income for the Fund. In assessing these possibilities, the committee observed that the Fund's activities could be broken down into three types of functions that cut across the full membership of industrialized countries, emerging markets, and low-income economies: financial intermediation, the provision of global public goods (for example, data, standards and codes, and combating terrorist financing), and the provision of bilateral services, in the form of capacity building and technical assistance.

The Committee concluded that revenue from Fund lending should be sufficient to cover its ongoing costs arising from financial intermediation. The Committee also noted that this income should not be used to cross-subsidize the provision of global public goods because (i) this income was too volatile for this purpose and (ii) cross-subsidization could cause IMF lending to become too expensive compared with private financing.

In order to ensure that the Fund could continue to provide its key global public goods, the Committee noted that the Fund could, like the United Nations, assess a periodic levy on member countries. The Committee did not, however, favour this source of income, as it 'would risk politicising the activities of the Fund' by making its work subject to regular financing calls. Nevertheless, the Committee did note that charging fees for some services might generate a small amount of additional revenue.

The Committee's core proposal concerned the creation of an endowment for the IMF that would provide a reliable income stream without relying on annual requests to member countries. The Committee suggested a further sale of IMF gold as a possible source of endowment funds. Such sales had been mooted at various points in the past for a variety of purposes; this was done to finance the establishment of the trust funds that underwrote the 1996 HIPC Initiative. But other plans for such sales have usually failed to gain enough support in the face of opposition from the United States and from gold-producing countries. To allay these fears, the Committee report suggested a 'balanced' approach, in which the Fund would also invest some of its quota resources in highly rated securities so that the burden of creating an investment endowment would not fall exclusively on the sale of gold.

As this article was being drafted, discussion was continuing on the exact form an endowment for the Fund could take. In meantime, the Fund had begun to invest some of its retained earnings from lending in investment grade securities in an effort to supplement its income.

### *6.3 Resources*

The relative size of the Fund shrank markedly from the 1970s ownard in comparison with, inter alia, global reserves, international trade, financial flows, stocks of financial assets and world output. This decline in pecuniary stature has distorted some of the

debates about the Fund's work, most notably on creditor and debtor moral hazard. Much of the debate over the implications of jumbo or 'exceptional access' arrangements in the 1990s (arrangements in which lending was equivalent to 300 per cent of quota or more) would be moot if regular quota increases had maintained the Fund's relative size in the global economy. Indeed, had the Fund grown through regularly scheduled quota increases, very few of the arrangements of the 1990s and 2000s would have been deemed at all exceptional. This suggests a simple yardstick for an appropriately-sized IMF: at any given time, the sum of the Fund's quotas should enable a risk-adjusted subset of its membership to borrow from the Fund on non-exceptional terms to finance their adjustment needs.

Accepting the validity of such a yardstick depends critically, however, on one's ultimate view of the role the IMF should play in the international system: trusted macroeconomic advisor, catalyst for private capital inflows and foreign assistance, or potential lender of last resort at time of crisis? To some extent the Fund played all of these roles at the turn of the 21st century, though its reduced relative size meant that the lender-of-last-resort function was credible only for its smaller members. The Fund staff, its shareholders, and those who care about the future of the multilateral system will need to decide which of these roles the IMF should continue to play.

BRETT HOUSE, DAVID VINES AND W. MAX CORDEN

*See Also* **Bretton Woods system; currency crises; development economics; emerging markets; Keynes, John Maynard; monetary approach to the balance of payments; World Bank.**

### *Bibliography*

Agénor, R., Miller, M. Vines, D. and Weber, A., eds. 1999. *The Asian Financial Crises: Causes, Contagion, and Consequences*. Cambridge: Cambridge University Press.

Arriazu, R., Crow, J. and Thygesen, N. 1999. *External Evaluation of IMF Surveillance*. Washington, DC: IMF. Online. Available at http://www.imf.org/external/pubs/ft/extev/surv/eval.pdf, accessed 6 June 2007.

Batini, N., Kuttner, K. and Laxton, D. 2005. Does inflation targeting work in emerging markets? In *World Economic Outlook*, September. Washington, DC: IMF. Online. Available at http://www.imf.org/external/pubs/ft/weo/2005/02/pdf/chapter4.pdf, accessed 6 June 2007.

Blejer, M., Leone, A.M., Raubanal, P. and Schwartz, G. 2001. Inflation targeting in the context of IMF-supported adjustment programs. Working Paper No. 01/31. Washington, DC: IMF.

Bluestein, P. 2001. *The Chastening: Inside the Crisis that Rocked the International Financial System and Humbled the IMF*. New York: Public Affairs.

Boorman, J., Lane, T., Schulze-Ghattas, M., Bulir, A. et al. 2000. Managing financial crises: the experience in east Asia. Working Paper No. 00/107. Washington, DC: IMF.

Boughton, J. 2001. *Silent Revolution: The International Monetary Fund 1979–1989*. Washington, DC: IMF.

Boughton, J. 2004. The IMF and the force of history: ten events and ten ideas that have shaped the institution. Working Paper, No. 04/75. Washington: IMF.

Boughton, J. and Mourmouras, A. 2004. Whose programme is it? Policy ownership and conditional lending. In *The IMF and its Critics: Reform of Global Financial Architecture*, ed. D. Vines and C. Gilbert. Cambridge: Cambridge University Press.

Bordo, M.D. 1993. The Bretton Woods international monetary system: an historical overview. In *A Retrospective on the Bretton Woods System, Lessons for International Monetary Reform*, ed. M. Bordo and B. Eichengreen. Chicago and London: University of Chicago Press.

Calvo, G. and Goldstein, M. 1996. What role for the official sector? In *Private Capital Flows to Emerging Markets After the Mexican Crisis*, ed. G. Calvo, M. Goldstein and E. Hochreiter. Washington, DC: Institute for International Economics.

Calvo, G. and Mendoza, E. 1996. Reflections on Mexico's balance of payment crisis: a chronicle of a death foretold. *Journal of International Economics* 41, 235–64.

Cohen, D. and Portes, R. 2004. Towards a lender of first resort. Discussion paper No. 4615. London: CEPR.

Corbett, J., Irwin, G. and Vines, D. 1999. From Asian miracle to Asian crisis: why vulnerability, why collapse? In *Capital Flows and the International Financial System*, ed. D. Gruen and L. Gower. Sydney: Reserve Bank of Australia. Repr. in Irwin and Vines (2001).

Corbett, J. and Vines, D. 1999a. Asian currency and financial crises: lessons from vulnerability, crisis, and collapse. *World Economy* 22, 155–77.

Corbett, J. and Vines, D. 1999b. The Asian crisis: lessons from the collapse of financial systems, exchange rates, and macroeconomic policy. In Agénor et al. (1999).

Corden, W.M. 1983a. Is there an important role for an international reserve asset such as the SDR? In *International Money and Credit: The Policy Roles*, ed. G.M. von Furstenberg. Washington, DC: IMF.

Corden, W.M. 1983b. The logic of the international monetary non-system. In *Reflections on a Troubled World Economy*, ed. F. Machlup, G. Fels and H. Muller-Groeling. London: Macmillan.

Corden, W.M. 1990. Macroeconomic adjustment in developing countries. In *Public Policy and Economic Development*, ed. M. Scott and D. Lal. Oxford: Clarendon Press.

Corden, W.M. 1993. Why did the Bretton Woods System break down? In *A Retrospective on the Bretton Woods System*, ed. M. Bordo and B. Eichengreen. Chicago: University of Chicago Press.

Corden, W.M. 1999. *The Asian Crisis: Is There a Way Out*? Singapore: Institute of South Asian Studies.

Corden, W.M. 2002. *Too Sensational: On the Choice of Exchange Rate Regimes*. Cambridge, MA: MIT Press.

Corden, W.M. 2007. Those current account imbalances: a sceptical view. *World Economy* 30, 363–82.

Dam, K. 1982. *The Rules of the Game*. Chicago: University of Chicago Press.

De Gregorio, J., Eichengreen, B., Ito, T. and Wyplosz, C. 1999. *An Independent and Accountable IMF*. Geneva Reports on the World Economy, No. 1. London: CEPR.

de Vries, M.G. 1976. *The International Monetary Fund 1966–71: The System under Stress*. Washington, DC: IMF.

de Vries, M.G. 1985. *The International Monetary Fund 1972–78: Cooperation on Trial*. Washington, DC: IMF.

de Vries, M.G. 1987. *Balance of Payments Adjustment, 1945 to 1986: The IMF Experience*. Washington, DC: IMF.

Dodge, D. 2006a. Global imbalances: why worry? What to do? Speech made to the New York Association for Business Economics, 29 March. *BIS Review*, 4 April. Online. Available at http://www.bis.org/review/r060404a.pdf, accessed 19 June 2007.

Dodge, D. 2006b. The evolving international monetary order and the need for an evolving IMF. Speech given at the Woodrow Wilson School of Public and International Affairs, 30 March. *BIS Review*, 5 April. Online. Available at http://www.bis.org/review/r060405a.pdf, accessed 19 June 2007.

Dooley, M., Landau, F. and Garber, P. 2003. An essay on the revived Bretton Woods system. Working Paper No. 9971. Cambridge, MA: NBER.

Easterly, W. 2005. What did structural adjustment adjust? The association of policies and growth with repeated IMF and World Bank adjustment loans. *Journal of Development Economics* 76, 1–22.

Edwards, S. 1989. The international monetary fund and the developing countries: a critical evaluation. In *IMF Policy Advice, Market Volatility, Commodity Price Rule, and Other Essays*, ed.

K. Brunner and A.H. Meltzer. Carnegie-Rochester Conference Series on Public Policy. Amsterdam: North-Holland.
Eichengreen, B. 1992. *Golden Fetters: The Gold Standard and the Great Depression 1919–1939*. Oxford: Oxford University Press.
Eichengreen, B., ed. 1995. *Europe's Postwar Growth*. New York: Cambridge University Press.
Eichengreen, B. 2004. The dollar and the new Bretton Woods system. Henry Thornton Lecture, delivered at the Cass School of Business, London, 15 December.
Eichengreen, B. 2007. *The European Economy Since 1945: Coordinated Capitalism and Beyond*. Princeton: Princeton University Press.
Eichengreen, B. and Park, Y.C. 2006. Global imbalances and emerging markets. Paper presented to the Asia-Europe Economic Forum Conference in Beijing on 13 July, 3, 7–68.
Eichengreen, B. and Portes, R. 1995. *Crisis, What Crisis? Orderly Workouts for Sovereign Debtors*. London: CEPR.
Finch, C.D. 1989. The IMF: the record and the prospects. Essays in International Finance No. 175. Princeton: International Finance Section, Department of Economics, Princeton University.
Fischer, S. 1999. On the need for an international lender of last resort. *Journal of Economic Perspectives* 13(4), 85–104.
Fischer, S. 2004. *IMF Essays from a Time of Crisis: The International Financial System, Stabilization, and Development*. Cambridge, MA: MIT Press.
Gardner, R. 1956. *Sterling Dollar Diplomacy*. Oxford: Oxford University Press.
Gilbert, C. and Vines, D., eds. 2000. *The World Bank: Structure and Policies*. Cambridge: Cambridge University Press.
Gilbert, C. and Vines, D. 2004. The IMF and international financial architecture: liquidity and solvency. In Vines and Gilbert (2004).
Goldstein, M. 2006. Currency manipulation and enforcing the rules of the international monetary system. In *Reforming the IMF for the 21st Century*, ed. E. Truman. Special Report No. 19. Washington, DC: Institute for International Economics. Online. Available at http://www.iie.com/publications/chapters_preview/3870/05iie3870.pdf, accessed 6 June 2007.
Gwin, C., Sachs, J., Kenen, P., Feinberg, R. and Nelson, J. 1990. *The International Monetary Fund in a Multipolar World: Pulling Together*. US–Third World Policy Perspectives No. 13, Overseas Development Council. New Brunswick: Transaction Books.
Horsefield, J.K. 1969. *The International Monetary Fund 1945–65: Twenty Years of International Monetary Cooperation*, 3 vols. Washington, DC: IMF.
IFIAC (International Financial Institution Advisory Commission). 2000. *Report of the International Financial Institution Advisory Commission*. Washington, DC: IFIAC, US Congress.
IMF (International Monetary Fund). 1987. *Theoretical Aspects of the Design of Fund-Supported Adjustment Programs*. Occasional Paper No. 55. Washington, DC: IMF.
IMF. 2002a. Guidelines on conditionality. Prepared by the Legal and Policy Development and Review Departments, 25 September. Online. Available at http://www.imf.org, accessed 30 May 2007.
IMF. 2002b. Review of the poverty reduction and growth facility (PRGF) and the poverty reduction strategy paper (PRSP) Approach, 15 March. Online. Available at http://www.imf.org, accessed 30 May 2007.
IMF. 2003. *The IMF and Recent Capital Account Crises: Indonesia, Korea, Brazil*. Report of the Independent Evaluation Office of the IMF. Washington, DC: IMF. Online. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2004a. Policy formulation, analytical frameworks and policy design. Prepared by the Policy Development and Review Department. Online. Available at http://www.imf.org, accessed 30 May 2007.
IMF. 2004b. *Evaluation of the IMF's Role in Poverty Reduction Strategy Papers and the Poverty Reduction and Growth Facility*. Report of the Independent Evaluation Office of the IMF. Washington, DC: IMF. Online. Available at http://www.imf.org, accessed 6 June 2007.

IMF. 2005a. *Argentina: 2005 Article IV Consultation – Staff Report*. IMF Country Report No. 05/236. Washington, DC: IMF.
IMF. 2005b. The Managing Director's report on the fund's medium-term strategy, 15 September. Online. Available at http://www.imf.org, accessed 30 May 2007.
IMF. 2005c. Review of PRGF Program Design – Overview. 8 August. Policy Development and Review Department. Online. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2006a. *Multilateral Surveillance*. Report of the Independent Evaluation Office of the IMF. Washington, DC: IMF. Online. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2006b. The Managing Director's report on implementation of the fund's medium-term strategy. 5 April. Online. Available at http://www.imf.org, accessed 30 May 2007.
IMF. 2007a. Committee to study sustainable long-term funding of the IMF: final report. Online. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2007b. *The IMF and Aid to Sub-Saharan Africa*. Report of the Independent Evaluation Office of the IMF. Washington, DC: IMF. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2007c. Final report of the external review committee on bank–fund collaboration. February. Washington, DC: IMF. Online. Available at http://www.imf.org, accessed 6 June 2007.
IMF. 2007d. *Regional Economic Outlook: Sub-Saharan Africa*. Washington, DC: IMF. Online. Available at http://www.imf.org/external/pubs/ft/reo/2007/AFR/ENG/sreo0407.htm, accessed 6 June 2007.
Irwin, G. and Vines, D., eds. 2001. *Financial Market Integration and International Capital Flows*. Cheltenham: Edward Elgar.
Irwin, G. and Vines, D. 2003. Government guarantees, investment, and vulnerability to financial crises. *Review of International Economics* 11, 860–74.
Irwin, G. and Vines, D. 2005. Policies for the resolution of international financial crises: how to avoid moral hazard. *International Journal of Money and Finance* 10, 233–50.
Joshi, V., Lane, T. and Vines, D. 2006. The US, East Asia and Europe: how to achieve an orderly resolution of global imbalances. Mimeo, Oxford University.
Keynes, J.M. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.
Keynes, J.M. 1971–1988. *The Collected Writings of J.M. Keynes*, ed. E. Johnson and D. Moggridge. London: Macmillan.
Kenen, P. 2006. *Comments on the address of the Managing Director of the IMF*. Washington, DC: Institute for International Economics. Online. Available at http://www.iie.com, accessed 30 May 2007.
Killick, T. 1986. *The Quest for Economic Stabilisation: The IMF and the Third World*. Aldershot: Gower.
King, M. 2005. The international monetary system. Remarks at Advancing Enterprise 2005 Conference, London, 4 February. Online. Available at http://www.bankofengland.co.uk, accessed 30 May 2007.
King, M. 2006a. Reform of the international monetary fund. Speech given to the Indian Council for Research on International Economic Relations. New Delhi, 20 February. Online. Available at http://www.bankofengland.co.uk, accessed 30 May 2007.
King, M. 2006b. Through the looking glass: reform of the international institutions. Inaugural International Distinguished Lecture, Melbourne Centre for Financial Studies, Australia, 21 December. Online. Available at http://www.bankofengland.co.uk, accessed 30 May 2007.
Krueger, A. 2002. A new approach to sovereign debt restructuring. Address given at the Indian Council for Research on International Economic Relations, Delhi, 20 December. Online. Available at http://www.imf.org, accessed 6 June 2007.
Krugman, P. 1999. Balance sheets, the transfer problem, and financial crises. *International Tax and Public Finance* 6, 459–72.
Kumar, M. 2006. Economic policy cooperation and coordination: analytical issues and historical experience. Mimeo, Fiscal Affairs Department, IMF.

Lane, T. 2006. Tensions in the role of the IMF and directions for reform. *World Economics* 6(2), 47–66.

Little, I.M.D. 1993. *Macroeconomic Analysis and the Developing Countries, 1970–1990*. Occasional Paper No. 41. San Francisco: International Centre for Economic Growth.

Little, I.M.D., Cooper, R., Corden, W.M. and Rajapatirana, S. 1993. *Boom, Crisis and Adjustment: The Macroeconomic Experience of Developing Countries*. Oxford: Oxford University Press.

Matthews, R.C.O. and Bowen, A. 1988. Keynesian and other explanations of postwar macroeconomic trends. In *Keynes and Economic Policy*, ed. W.A. Eltis and P.J.N. Sinclair. London: NEDO.

Matthews, R.C.O., Feinstein, C.H. and Odling-Smee, J.C. 1982. *British Economic Growth 1956–73*. Oxford: Clarendon Press.

McKibbin, W. 1997. Empirical evidence on international economic policy coordination. In *Handbook of Comparative Economic Policies Volume 5: Macroeconomic Policy in Open Economies*, ed. M. Fratianni, D. Salvatore and J. Von Hagen. London: Greenwood Press.

Meyer, L., Doyle, B., Gagnon, J. and Henderson, D. 2004. International coordination of macroeconomic policies: still alive in the new millennium? In *The IMF and its Critics: Reform of Global Financial Architecture*, ed. D. Vines and C. Gilbert. Cambridge: Cambridge University Press.

Moggridge, D. 1986. Keynes and the international monetary system 1909–1946. In *International Monetary Problems and Supply Side Economics: Essays in Honour of Lorie Tarshis*, ed. J.S. Cohen and G.C. Harcourt. London: Macmillan.

Mussa, M. and Savastano, M. 1999. The IMF approach to economic stabilization. In *NBER Macroeconomics Annual 1999*, ed. B. Bernanke and J. Rotemberg. Cambridge, MA: MIT Press.

Odling-Smee, J. 2004. The IMF and Russia in the 1990s. Working Paper No. 04/155. Washington, DC: IMF.

Polak, J. 1957. Monetary analysis of income formation and payments problems. *IMF Staff Papers* 6 (November), 1–50.

Portes, R. 2004. Resolution of Sovereign debt crises: the new old framework. Discussion paper No. 4717. London: CEPR.

Roubini, N. and Setser, B. 2005. Will the Bretton Woods 2 regime unravel soon? The risk of a hard landing in 2005–2006. Unpublished manuscript. Online. Available at http://pages.stern.nyu.edu/~nroubini/papers/BW2-Unraveling-Roubini-Setser.pdf, accessed 19 June 2007.

Sachs, J. 1995. Do we need an international lender of last resort? F.D. Graham Lecture at Princeton University, 20 April.

Skidelsky, R. 2000. *John Maynard Keynes: Fighting for Britain, 1937–1946*. London: Macmillan.

Stiglitz, J. 2002. *Globalization and Its Discontents*. New York: W.W. Norton & Company.

Tarp, F. 1993. *Stabilization and Structural Adjustment: Macroeconomic Frameworks for Analysing the Crisis in Sub-Saharan Africa*. London: Routledge.

Temin, P. 2002. The golden age of European growth reconsidered. *European Review of Economic History* 6, 3–22.

van Dormael, A. 1978. *Bretton Woods: Birth of a Monetary System*. London: Macmillan.

Vines, D. 2003. John Maynard Keynes 1937–1946: the creation of international macroeconomics. *Economic Journal* 113, F338–61.

Vines, D. and Gilbert, C., eds. 2004. *The IMF and its Critics: Reform of Global Financial Architecture*. Cambridge: Cambridge University Press.

Williamson, J. 1977. *The Failure of World Monetary Reform 1971–74*. New York: New York University Press.

Williamson, J. 1983a. Keynes and the international economic order. In *Keynes and the Modern World*, ed. G.D.N. Worswick and J. Trevithic. Cambridge: Cambridge University Press.

Williamson, J., ed. 1983b. *IMF Conditionality*. Washington, DC: Institute for International Economics.

Williamson, J. 2000. The role of the IMF: a guide to the reports. Policy Brief 00–5. Washington, DC: Institute for International Economics.

Williamson, J. 2006. The potential of international policy coordination. Paper presented at a seminar on Implications for the IMF's Role in Surveillance and Policy Coordination, at a Roundtable on International Economic Cooperation for a Balanced World Economy, Chongqing, China, 12–13 March. Online. Available at http://www.iie.com, accessed 30 May 2007.

Wolf, M. 2005. Multilateral leadership can right the ship. *Financial Times*, 28 June.

Wolf, M. 2006. The world needs a tough and independent monetary fund. *Financial Times*, 22 February.

Woods, N. 2006. *The Globalisers: The IMF, the World Bank and their Borrowers*. Ithaca and London: Cornell University Press.

## international monetary institutions

Domestic money is conceived of by society as a device to facilitate transactions in the marketplace, as a temporary store of value, and as a unit of account for contracts. Given the possibilities of fraud and counterfeiting, domestic monetary authorities have been established to regulate the quality of the domestic monetary unit in most countries. Such regulations attempt to guarantee the interchangeability of the different media, such as currency and the deposits of different banks, as well as stability in the value of the monetary unit, under conditions of prosperity.

*International* monetary arrangements are required under conditions of international trade when residents of different countries must make payments to each other, and yet wish to hold most of their assets in terms of domestic currency. Such arrangements are designed to guarantee *convertibility* of assets denominated in different currencies, so that payments may be made independent of country of residence, thus facilitating a free and open trading system. International monetary *institutions* such as the International Monetary Fund are designed to support international monetary arrangements by enforcing rules of behaviour, assisting countries in difficulties, and encouraging good practices.

### Alternative exchange rate mechanisms

Under a *gold standard*, domestic residents and foreign residents may freely convert domestic currency into gold at a fixed rate of exchange. This type of convertibility was eliminated in the 1930s in favour of a gold *exchange* standard, which allowed only foreign monetary authorities to exchange domestic currency for gold. Gold convertibility of both types was ended as part of the Smithsonian Agreement of 1971 (see below).

Under a system of *pegged* exchange rates between different currencies, as established by the Bretton Woods system (see below), convertibility implies that domestic residents are free to obtain foreign currency at a *fixed* rate of exchange for the purchase of foreign goods and services, inclusive of normal trade credit. Likewise, foreign residents are free to sell domestic currency obtained by sale of goods and services or to use it for purchase of domestic goods and services, at the same fixed rate of exchange. This definition does not require free convertibility for capital account transactions (those arising from exchanges of financial assets only).

Under a system of floating or flexible exchange rates, convertibility still implies that both domestic and foreign residents may freely convert domestic and foreign currency at the same rate of exchange for current account transactions, but the exchange rate at which this may be done is determined on a daily basis by market transactions, rather than being guaranteed by the domestic monetary authorities of the respective countries.

In 2005, only 20 out of the 184 member countries of the International Monetary Fund (IMF) declined to accept the obligations to current account convertibility. But in a large number of countries various types of restrictions limited convertibility in some way or created differences in the exchange rates applying to exports and imports.

Non-unified exchange rates lead to inefficient allocation of resources, as previously documented by Bhagwati (1978). For example, 70 countries required repatriation and surrender of proceeds of exports or invisible transactions, 57 countries had payments arrears of one kind or another, and 11 countries maintained either dual or multiple exchange rates for different types of transactions. With respect to capital account transactions, the situation is much more restrictive: 126 countries had controls on international transactions in capital market securities, and 143 countries maintained controls on direct investment flows.

### Reserve assets

In order to guarantee convertibility of the domestic currency into other convertible currencies, monetary authorities hold stocks of *reserve assets*, which are liquid assets held in readily accepted international media of exchange, such as dollars, euros, and a few other currencies. In addition, IMF member countries have access to unconditional borrowing rights to obtain additional reserve assets in the form of their reserve positions in the Fund and Special Drawing Rights. These, together with reserve asset holdings, make up *international liquidity*.

Since most international payments are handled by inter-bank transactions, banks have sought to minimize transactions costs by channelling their foreign exchange transactions through one or more *vehicle* currencies, the pound sterling in earlier days, but more recently the US dollar and to some extent the euro. Because the dollar is so widely used in private exchange transactions, monetary authorities also find it convenient to operate in dollars to ensure the convertibility of their currencies.

### Adjustment mechanisms

The existence of different national currencies and the need to maintain convertibility of the different currencies lead to the concept of balance of payments adjustment mechanism. At a given exchange rate, as long as the amount of foreign exchange earned through exports of goods and services and capital inflows just pays for imports and capital outflows, no external imbalance exists. If international capital markets were perfect and if investors were risk neutral so that assets denominated in different currencies were perfect substitutes for one another in private portfolios, there would in practice be a single world interest rate for short-term borrowing. Then imbalances between foreign exchange earnings and payments could simply be *financed* by borrowing in the international capital market. There would be no real distinction between the convertibility characteristics of the official liabilities of different borrowers.

But, in fact, countries face very real limits on the amount of foreign currency they can borrow abroad in exchange for domestic currency because of *exchange rate* risk, which limits the willingness of risk-averse foreign lenders to acquire domestic currency assets. According to the doctrine of *original sin*, countries with a history of convertibility problems are unable to issue foreign debt in their own currency (Eichengreen and Hausmann, 2005). The ability to repay foreign currency debt is dependent on balance of payments adjustment. *Political* risk involves the possibility that exchange controls may be imposed in the future, preventing the repayment of foreign currency debt on

the promised terms. Thus it is desirable for countries to have access to a variety of adjustment mechanisms to eliminate external imbalances, as well as a variety of sources of official financing in the form of international liquidity. The primary mechanisms of balance of payments adjustment are through movements in exchange rates and adjustments of income and price levels via monetary and fiscal policies. The need for adjustment can be postponed by imposition of tariffs and subsidies, quantitative restrictions on current account or capital account transactions, or controls over the allocation of foreign exchange. But tariffs, quantitative restrictions, and exchange controls generally involve inefficiencies in the allocation of resources, including in the latter case loss of convertibility of the domestic currency. Changes in monetary and fiscal policies or exchange rates have their own costs in terms of domestic policy objectives forgone.

## Financing

Thus, a mixture of adjustment policies and financing mechanisms is provided in a system of international monetary arrangements. *Official* financing is provided either by drawing on holdings of official reserve assets or by borrowing from international institutions. *Private* financing can be arranged by a monetary authority borrowing from foreign banks or the international bond market. Either provides the ability to postpone adjustment. The optimum mix of adjustment and financing for an individual country depends on the costs of the various alternatives. By setting the costs of these alternatives, international monetary arrangements influence the behaviour of the world economy.

## A model of adjustment versus financing

In the theory of adjustment versus financing, a country is faced with random balance of payments deficits and surpluses, which it may either finance by drawing on reserve assets or adjust by one of the adjustment mechanisms mentioned above. In one branch of the theory, due to Heller (1966) and others, the cost of adjustment is assumed to be a linear function of the size of the adjustment, so that any adjustments are postponed to the last minute, at which time full adjustment takes place. Alternatively, one may assume a nonlinear cost of adjustment, leading to a theory of partial adjustment. Kelly (1970) and Clark (1970) assume that the country's welfare function depends on the mean and variance of income, so that gradual adjustments are preferred. The analysis determines both the optimum level of reserve holdings, $R^*$, and the optimum rate of adjustment $\alpha$ to that level, according to the equation

$$\Delta R = \alpha[R^* - R_{-1}] + u, \tag{1}$$

where $u$ is normally distributed with mean zero and variance $\sigma^2$ and $R_{-1}$ is the stock of reserves at the end of the previous period. This equation assumes that changes in the stock of reserves arise from both the random shocks in the balance of payments and the desired rate of adjustment to the optimal level of reserves. From eq. (1) we find that the variance of reserve holdings decreases as the speed of adjustment $\alpha$ increases from zero to one.

Tchebychev's inequality then enables one to show that, for a given probability of not exhausting reserves and given opportunity cost $r$ of holding reserves, the optimum

reserve holding $R^*$ decreases with increasing $\alpha$. As $\alpha$ increases, the need for more frequent adjustments raises the variance of income. Therefore the speed of adjustment should be chosen such that the welfare loss from increased variance in income due to a small increase in $\alpha$ is just counterbalanced by the welfare saving due to holding slightly smaller reserves.

According to this theory, international monetary institutions will strongly affect the behaviour of national policies concerning balance of payments adjustment and acquisition of reserves. Specifically, international money institutions will determine the opportunity cost of holding reserves, the penalty attached to running out of reserves, and the availability of different types of adjustment policies. By influencing countries' balance of payments adjustment policies, international institutions will also influence their domestic policies, since there is a trade-off between internal and external objectives of policy.

## The role of markets and institutions

An optimal design for the international monetary system depends on balancing among a group of conflicting objectives: growth of real income and employment, stable prices, efficient allocation of resources, maintenance of convertibility of currencies, improving the distribution of income, and growth of world trade. The relevant trade-offs can be understood in the context of an economic model. According to the model of adjustment and financing outlined above, reductions in the opportunity cost of holding reserves will lead to increased reserve holdings, a reduction in the speed of adjustment to imbalances, increased use of financing, and a decline in the variability of income. The slowdown in the speed of adjustment implies a change in the allocation of resources among countries. The increased use of financing may imply an increase in the rate of inflation. An optimal international system should balance these various considerations. For discussion of efforts to design such a system, see Solomon (1982) and the documents of the IMF's Committee of Twenty (IMF, 1974).

In a purely laissez-faire system, market borrowing instead of official reserves would be the source of financing to postpone adjustment. Fluctuations in market interest rates would determine the terms of trade between adjustment and financing. As is usual in market solutions, the wealthy are in a better position to negotiate terms on loans. By contrast, a more institutionalized system provides access to financing at lower rates to those with a weaker market position, with more conditions on the use of the funds. Evaluating the difference between two such systems is a complex task. For an attempt, see Jones (1983).

## The evolution of international monetary institutions

Between the close of the Napoleonic Wars and 1880, the international monetary system gradually moved onto the gold standard, which was fully achieved during the period 1880–1914. Under the leadership of Great Britain, sterling operated as a vehicle currency during this period, allowing an efficient international payments mechanism to develop. The increasing substitution of bank deposits for currency allowed an ever-larger volume of payments to be supported by a gradually rising supply of gold. Despite the best efforts of the Bank of England and other central banks, periodic crises interfered with the continued convertibility of individual currencies. And the system

was characterized by substantial fluctuations in employment and prices, albeit about a rising trend of employment with no trend in prices.

Following the First World War, gold convertibility was resumed on a limited basis, until the Great Depression of 1929–33 brought it to an end. A period of fluctuating exchange rates, competitive devaluations, and increasing use of trade restrictions to promote domestic employment ensued. It is generally believed that the economic difficulties of the interwar period were major factors bringing on the Second World War.

*The Bretton Woods system*

The United States and Great Britain took the lead in constructing the post-war international monetary institutions, with Harry Dexter White and John Maynard Keynes drawing up rival designs for the new system agreed at the Bretton Woods Conference in 1944. The Articles of Agreement of the International Monetary Fund provided for a system based on pegged, but adjustable, exchange rates and an institution which would lend reserve assets to countries that were having temporary difficulties in maintaining convertibility. Resort to floating exchange rates, competitive devaluations, and trade restrictions to promote domestic employment were explicitly to be avoided, in the light of the problems of the 1930s. Convertibility for current account transactions was promoted, while capital account convertibility was required only for those transactions necessary for financing current payments.

The lending power of the IMF was based on *quotas* of gold and domestic currency contributed by each member country. Only the gold was to be paid in initially, but, if the Fund needed convertible currency to lend out, it would obtain it from any member whose currency was considered strong enough to be *usable*. Members could borrow automatically up to the amount of the gold portion or *tranche* of the quota, but only on demonstration of balance of payments need, and thereafter they could borrow more subject to meeting conditions on economic and financial policies. For further discussion of IMF policies, see Williamson (1983), Kenen (2001), and Truman (2006).

The initial post-war problem involved the establishment of a payments system that would promote economic recovery and the growth of trade among the former combatants. The International Monetary Fund limited itself to establishing a set of agreed par values for pegged exchange rates which could promote the growth of trade, leaving the provision of loans and grants for economic recovery to the United States, the strongest economy. Under this system, which was a form of gold exchange standard, countries declared their par values in terms of the US dollar, which in turn was convertible into gold at $35 an ounce. Thus the dollar became the key currency of the system, and most foreign exchange reserves came to be held in the form of dollars. Within Europe, convertibility remained limited until 1958, and the European Payments Union was established to facilitate intra-European payments. The re-establishment of convertibility led to fears that the IMF might have inadequate resources to deal with the problems of large member countries. In 1962 the General Arrangements to Borrow were created, to enable the Fund to mobilize additional resources from its largest members, the Group of Ten.

With the recovery of the European economies in the 1950s and the achievement of convertibility in 1958, the US dollar became gradually overvalued relative to gold and

other currencies. As Robert Triffin (1960) pointed out, the key currency system required the United States to continue to run balance of payments deficits in order to supply other countries with increased foreign exchange reserves. As it did so, the gold reserve of the United States became increasingly inadequate to guarantee gold convertibility of growing US official dollar liabilities at $35 an ounce.

A variety of solutions to this problem were proposed, including the creation of an artificial reserve asset to substitute for dollars, an increase in the dollar price of gold, and the adoption of floating exchange rates. In 1968 the First Amendment to the Articles of Agreement of the International Monetary Fund permitted the creation of Special Drawing Rights (SDRs), which have twice been allocated to member countries in proportion to their existing quotas in the Fund. SDRs, when utilized, permit the user to acquire convertible currencies from other members, upon the payment of interest. They represent a centralized mechanism for increasing the stock of reserves. By the early 1970s the gold convertibility of the dollar was under increasing pressure, for a variety of reasons. In August 1971 the dollar was unilaterally set loose from gold. The Smithsonian Agreement of December 1971 attempted to save the Bretton Woods system by multilateral realignment of exchange rates, including a devaluation of the dollar against gold and a widening of the narrow bands of fluctuation permitted around the newly fixed values. Some members of the European Communities (EC) agreed to maintain narrower margins of fluctuations versus each other's currency, in an arrangement that became known as the 'EC Snake'. Despite these efforts, the revised Bretton Woods system lasted only a little more than a year.

*Floating exchange rates*

In March 1973, exchange rates of most of the major industrial countries began floating. At the same time, most developing countries continued to peg their currencies to the dollar or another developed country currency, and the EC maintained the 'Snake'. About this time, a major effort to reconstruct international monetary institutions on the basis of pegged exchange rates began under the auspices of the IMF's Committee of Twenty. This effort collapsed in 1974, in part under the impact of the quadrupling of world oil prices by the Organization of Petroleum Exporting Countries.

In Jamaica in January 1976, the Interim Committee of the Board of Governors of the International Monetary Fund agreed on a Second Amendment to the Fund's Articles of Agreement, ratifying the system of floating exchange rates. First, stability of exchange rates was to be sought through stability of underlying monetary and fiscal policies rather than through pegging. Second, floating rates should be subject to a process of 'firm surveillance' by the IMF. Third, it was hoped that the SDR would 'become the principal reserve asset', with the role of gold and the dollar being reduced. Fourth, the fixed official price of gold was abolished and one-third of the IMF's gold was disposed of. Acceptance of the status quo was all that could be accomplished. The result, according to Corden (1983), was an international laissez-faire system.

In 2005 some 88 countries made use of floating exchange rates, while 51 had pegged exchange rates of one type or another and 48 operated within currency unions with other countries.

*Increased capital mobility, the Asian crisis and reform proposals*

Beginning in the 1970s, international capital mobility increased significantly, as middle-income developing countries found new access to foreign borrowing and industrialized countries increasingly opened production facilities in each others' markets. In the early 1990s, the IMF began discussions of a possible amendment that would promote capital account convertibility as an additional goal of the international monetary system, on the argument that improved allocation of capital would lead to increased economic growth. But a series of crises in emerging market economies interfered with this project, most notably the Asian financial crisis of 1997, followed by the Russian crisis of 1998 and the Argentine crisis of 2001. Each of these events was preceded by substantial capital inflows seeking higher returns, which overwhelmed under-regulated and under-prepared domestic economies and financial systems. The convertibility of affected currencies was often temporarily impaired (Black, Christofides and Mourmouras, 2006). In some cases the IMF was seen as creating a permissive environment prior to the crisis, followed by harsh demands for domestic reforms subsequently, in attempts to restore confidence and bring an end to capital outflow.

A substantial body of criticism on one side argued that, by its willingness to provide large amounts of financing to countries in crisis, the Fund had created 'moral hazard', encouragement to over-borrowing and over-lending in expectation of a bailout (International Financial Institution Advisory Commission, 2000). On the other side, others claim that the Fund by its harsh requirements for reform was stifling economic recovery and growth (Stiglitz, 2002). Both of these viewpoints may have had some validity, but in a sense they cancel each other out (see Kenen, 2001). The Fund itself proposed creation of an international Sovereign Debt Restructuring Mechanism to assist defaulting countries in negotiations with creditors (Krueger, 2003). This was rejected in favour of a more modest approach encouraging the use of collective action clauses in bond indentures requiring minority bondholders to accept terms of repayment agreed to by a majority.

Another criticism of the IMF is that its voting shares and representation appear outdated, as compared with the changing economic importance of different groups of countries (Truman, 2006). In particular, large emerging market economies such as China, India, and Brazil are under-represented, while the European Union countries with 32 per cent of the voting power are over-represented. Obviously, changes in representation are extremely difficult to achieve, but will still be necessary to remedy a situation in which the rich creditor countries that do not utilize the Fund's resources have disproportionate voting power relative to the debtor nations that have greater need for use of its facilities.

*The 'new' Bretton Woods and Asian monetary cooperation*

Following recovery from the Asian crisis of 1997, countries such as Korea, China, Malaysia, Taiwan and India sharply increased their accumulations of international reserves, as developing Asian countries in total raised their reserves (minus gold) from SDR 414 billion to SDR 1,039 billion between the ends of 1998 and 2004. China, Hong Kong and Malaysia in particular sought to maintain exchange rates pegged to the US dollar, while the other countries managed their floating exchange rates so as to avoid undue appreciation against the US dollar, accumulating enormous reserves in the

process. An influential paper by Dooley, Folkerts-Landau and Garber (2004) argued that this relationship was a new version of the old Bretton Woods system, whereby other countries pegged their exchange rates to the US dollar, enabling the United States to run large current account deficits, while the creditor nations increased their exports to the United States. Alternatively, the vastly increased reserve holdings of Asian countries could be regarded as a precautionary response to insure the availability of financing to avoid the prospect of another sharp adjustment, following the unpleasant experiences of the 1997 Asian crisis.

The combination of increased regional reserve holdings and recent bad experience with internationally supervised adjustment has led Asian countries to embark on steps towards regional monetary cooperation, culminating in the so-called Chiang Mai Initiative for regional currency swaps among the Association of South East Asian Nations (ASEAN) plus China, Japan, and Korea (see Park and Wang, 2005). ASEAN members realized that the industrial countries of the Group of Ten had previously used currency swaps among central banks to lend each other money in times of crisis and thus avoid the need for borrowing from the IMF with its conditionality. With growing availability of reserves in Asia, the ASEAN + 3 concluded that they might similarly help each other out in future. Under the leadership of the Asian Development Bank, further steps are contemplated, possibly including an Asian Monetary Fund and an Asian Currency Unit.

*The European Monetary Union*

The enlargement and strengthening of the EC 'Snake' in 1978, which was in the process renamed the European Monetary System (EMS), gradually led to the creation of the European Monetary Union with a unit of account, the European Currency Unit (ECU). The objectives of the enlarged EMS were to reduce intra-European exchange rate fluctuations, to promote convergence of macroeconomic policies within Europe, and to reduce European dependence on US monetary policies. Over a period of 15 years, the EMS succeeded in these objectives, at the cost of a series of exchange rate realignment crises culminating in a major collapse of the system in 1992–3, when the narrow margins (plus or minus $2\frac{1}{4}$ per cent) were expanded (to plus or minus 15 per cent). The crisis was brought on by a combination of increasingly rigid exchange rates within the system, increased capital mobility as a component of the Single Market programme of the European Union, and stresses brought on by the unification of East and West Germany.

In response to these factors, and to further strengthen the integration of European markets and achieve a more symmetrical sharing of decision making in monetary policy, the Maastricht Treaty ratified in 1993 brought into being in 1999 the European Monetary Union, with a single currency, the *euro*, with monetary policy controlled by a European Central Bank (ECB) in Frankfurt, Germany, replacing the currencies of the 12 member countries of the eurozone. While the euro has been quickly accepted as an international currency, in both the member countries and their neighbours, the relatively conservative operations of the ECB together with the constraints on member countries' fiscal policy embodied in the Stability and Growth Pact have proven controversial in the light of slow economic growth in the eurozone.

The euro is gradually becoming more important in international transactions and in the foreign exchange market as a rival to the US dollar. In 2006 the IMF redefined the SDR currency basket reflecting the importance of currencies in international trade and finance to be composed of 44 per cent US dollars, 34 per cent euro, 11 per cent Japanese yen and 11 per cent pound sterling, as compared with the previous weights of 45 per cent US dollars, 29 per cent euro, 15 per cent yen and 11 per cent pound sterling.

STANLEY W. BLACK

*See Also* **capital controls; gold standard; international capital flows; international reserves.**

### *Bibliography*

Bhagwati, J. 1978. *Anatomy and Consequences of Exchange Control Regimes*. Cambridge, MA: Ballinger.

Black, S. 1985. International money and international monetary arrangements. In *Handbook of International Economics*, vol. 2, ed. R. Jones and P. Kenen. Amsterdam: North-Holland.

Black, S., Christofides, C. and Mourmouras, A. 2006. Convertibility risk: the precautionary demand for foreign currency in a crisis. *Annals of Finance* 2, 141–65.

Clark, P. 1970. Optimum international reserves and the speed of adjustment. *Journal of Political Economy* 75, 356–76.

Corden, W. 1983. The logic of the international monetary non-system. In *Reflections on a Troubled World Economy*, ed. F. Machlup et al. London: Macmillan.

Dooley, M., Folkerts-Landau, D. and Garber, P. 2004. The revived Bretton Woods system. *International Journal of Finance and Economics* 9, 307–13.

Eichengreen, B. and Hausmann, R., eds. 2005. *Other People's Money: Debt Denomination and Financial Instability in Emerging Market Economies*. Chicago and London: University of Chicago Press.

Heller, H. 1966. Optimal international reserves. *Economic Journal* 74, 333–52.

IMF (International Monetary Fund). 1974. *International Monetary Reform: Documents of the Committee of Twenty*. Washington, DC: IMF.

International Financial Institution Advisory Commission. 2000. *Report*. Washington. United States Congress.

Jones, M. 1983. International liquidity: a welfare analysis. *Quarterly Journal of Economics* 98, 1–23.

Kelly, M. 1970. The demand for international reserves. *American Economic Review* 60, 655–67.

Kenen, P. 2001. *The International Financial Architecture: What's New? What's Missing?* Washington, DC: Institute for International Economics.

Krueger, A. 2003. Sovereign debt restructuring: messy or messier? *American Economic Review* 93(2), 70–4.

Park, Y. and Wang, Y. 2005. The Chiang Mai Initiative and beyond. *World Economy* 28, 91–101.

Solomon, R. 1982. *The International Monetary System, 1945–1981*. New York: Harper & Row.

Solomon, R. 1999. *Money on the Move: The Revolution in International Finance Since 1980*. Princeton: Princeton University Press.

Stiglitz, J. 2002. *Globalization and its Discontents*. New York and London: Norton.

Triffin, R. 1960. *Gold and the Dollar Crisis*. New Haven: Yale University Press.

Truman, E. 2006. *A Strategy for IMF Reform*. Washington, DC: Institute for International Economics.

Williamson, J., ed. 1983. *IMF Conditionality*. Washington, DC: Institute for International Economics.

# Kindleberger, Charles P. (1910–2003)

Charles P. Kindleberger was born in New York City. He received his B.A. at the University of Pennsylvania in 1932 and his Ph.D. at Columbia University in 1937. He had a distinguished career in public service (including the Federal Reserve and the Office of Strategic Services during the Second World War) before going to teach international trade at MIT. His wartime experiences directed his interests towards the interaction of countries and gave him a keen sense of how academic ideas play out among real people and governments. His scholarship was characterized by its realism and willingness to consider actual – as opposed to idealized – behaviour.

Kindleberger made his mark on the field of international trade through his textbook and through papers and books about the recovery of Europe after the Second World War. He was active in the analysis of the dollar scarcity and then the dollar glut that characterized the short life of the Bretton Woods System. He also wrote a prescient book, *Europe's Postwar Growth: The Role of Labor Supply* (1967), on the role of immigrants and guest workers from eastern and southern Europe in alleviating the labour scarcity of western Europe. Kindleberger's emphasis on the evolution of labour supply has been echoed in many subsequent studies. The legacy of these post-war policies has been evident in political and economic conflict between the children and even grandchildren of these immigrants and other residents. As Kindleberger said (1967, p. 213), the short-run benefits of labour migration are clear, but there are dangers in the long run: 'To rely heavily on foreign labor in one's economy constitutes a positive risk.'

Kindleberger made his entry into economic history with *Economic Growth in France and Britain, 1851–1950* (1964). He surveyed the extensive literature on these two countries and concluded that there was no single convincing explanation for the differences between them. He ended the book with the following famous words: 'Economic history, like all history, is absorbing, beguiling, great fun. But, for scientific problems, can it be taken seriously?' This ironic comment set the tone for Kindleberger's future work in economic history. His books and papers are distinguished by his command of the previous literature. His reasoning is informed by an intelligent, if sceptical, use of economic theory. His prose is sprightly. And his conclusions are clear, forcefully presented, and always worth debating.

Kindleberger's impact on economics and economic history comes primarily from two books first published in the 1970s. The first, *The World in Depression, 1929–1939* (1973), provided a comprehensive narrative of the Great Depression from an international perspective. Instead of seeing the Depression as a succession of national stories, Kindleberger argued persuasively that it was the result of a failure of the international economic system. The economic structure built around the gold standard had allowed the pre-war industrial economies to weather various economic shocks in the late 19th and early 20th centuries, but it proved unable to contain or offset the shocks arising in the period after the First World War.

Why so? Kindleberger argued that the inter-war economy lacked a hegemon, a dominant leader. The hegemonic power in the pre-war period was the United Kingdom, more specifically the Bank of England, which acted to contain crises wherever they started. But England was exhausted by the effort to defeat Germany in the First World War, and the Bank of England was in no shape to continue this role. Although the United States was the obvious candidate to pick up the baton, Americans were isolationist after their wartime efforts and declined to act. In the shortest summary: no longer London, not yet New York. Without a hegemon, the shocks to the world economy in the late 1920s were allowed to drag the world into the Great Depression.

The costs of encouraging immigration of foreign workers after the Second World War emerged only slowly; the costs of poor macroeconomic policies in the early 1930s became evident more quickly. Kindleberger recounted the abortive efforts of central bankers and government officials to organize some kind of cooperative solution to the economic shocks. Failing in this endeavour, the world was subjected to competing devaluations and deflations. Among the costs was extensive damage to financial institutions and to the operation of those economies that held on to the gold standard.

Kindleberger generalized his argument in *Manias, Panics, and Crashes: A History of Financial Crises* (1978). He surveyed financial crises in the past two centuries that were important enough to have macroeconomic effects. He described the various irrationalities that preceded crises, as suggested in his title, and synthesized a vast literature in a small and engaging book. He argued that irrationally optimistic expectations frequently emerge among investors in the late stages of major economic booms, differing sharply from most modern models of finance and relying on a more impressionistic theory of financial crises. When these optimistic expectations appear, investors grossly overestimate the future profitability of some promising firms. These overestimates lead unscrupulous managers to over-promote their firms vigorously and to issue bogus debt and equity with abandon. They may lead even well-meaning, sober managers to issue unsupportable amounts of debt. The more a firm's managers sincerely overestimate their firm's growth opportunities or successfully promote a Ponzi-style fraud, the more securities they try to issue. When the unrealistically high profits fail to develop as predicted, debt and stock values collapse. Markets for over-promoted financial assets may even dry up. The more severe the price decline, the more the collapsing value of previously high-flying assets spreads insolvency to creditors of both the over-expanded firms and their stockholders.

Kindleberger observed that speculation in a bubble often develops in two stages. In the first, sober stage of investment, seasoned professional investors and analysts are gradually persuaded that bubble assets offer a good chance of high returns. In the second stage, 'professional company promoters – many of them rogues interested only in quick profits – tempted a different class of investors, including ladies and clergymen'. It is of course hard for any market participant or observer to know when the bubble has progressed from the first stage to the second.

Kindleberger concluded that stability is promoted when a lender of last resort exists and follows the recommendations of Walter Bagehot over a century ago in his *Lombard Street* (1873) to lend freely at punitive rates during a crisis. This is what a hegemonic power – the United States government internationally and the Federal

Reserve domestically – should have done in the 1930s, in Kindleberger's view; it is what the International Monetary Fund should do today. His book has proved exceedingly popular with a varied audience: economists, investors and the general public alike. It was revised and expanded several times; the fourth edition was published shortly before Kindleberger's death, when he was 90 years old.

PETER TEMIN

*See Also* **bubbles; Great Depression; Great Depression (mechanisms).**

### *Selected works*

1964. *Economic Growth in France and Britain, 1851–1950*. Cambridge, MA: Harvard University Press.
1967. *Europe's Postwar Growth: The Role of Labor Supply*. Cambridge, MA: Harvard University Press.
1973. *The World in Depression, 1929–1939*, rev. edn. Berkeley: University of California Press, 1986.
1978. *Manias, Panics, and Crashes: A History of Financial Crises*. 4th edn. New York: Wiley, 2000.
1993. *A Financial History of Western Europe*, 2nd edn. New York: Oxford University Press.
2000. *Comparative Political Economy: A Retrospective*. Cambridge, MA: MIT Press.

### *Bibliography*

Bagehot, W. 1873. *Lombard Street: A Description of the Money Market*. London: H. S. King.

# laboratory financial markets

Laboratory financial markets allow human subjects to trade assets under conditions controlled by the researcher. By varying the conditions – such as the trading format, or the timing and content of private information – the researcher can make direct and sharp inferences.

Such inferences are crucial to achieve insight into the ongoing debate about the importance of behavioural anomalies in financial markets (see behavioural finance). Efficient markets and related theories provide a satisfying explanation for many of the properties of modern financial markets, but they are hard to reconcile with well documented 'market anomalies' such as home bias, the large equity premium and excessive volatility. Should financial economists force a reconciliation, or should they embrace prospect theory and other behavioural theories?

These issues are not just academic. Since the collapse of the Soviet bloc around 1990, a dominant share of the world economy has relied on financial markets to choose its economic future. If the efficient markets theory is wrong, and asset prices do not necessarily reflect all available information, then major restructuring may be in order. Perhaps the global economy would be stronger with information disclosures that cater to our behavioural idiosyncrasies, or even with non-market allocation of investment.

Laboratory asset markets inform the debate by offering evidence that complements field data. The strength of experimental methodology is that the researcher can precisely control information, public and private, and can elicit beliefs as well as track offers, transactions and allocations. Thus, in a simplified setting, researchers can systematically dissect the process of asset price formation. In conjunction with theory and field empirical work, laboratory investigations help us understand how financial markets really work.

## Early laboratory markets

Experimental economics cut its teeth on laboratory commodity markets. Reacting to Edward Chamberlin's casual classroom experiments, Vernon Smith pioneered the scientific study of markets in the laboratory. He refined the idea of *induced value and cost*: the experimenter promises to pay a subject the amount $v$ if she buys a unit, and charges another subject the amount $c$ if he sells a unit. If they transact at price $p$, she earns $v - p$ and he earns $p - c$, generating surplus of $v - c$. The payments are in cash and large enough for the subjects to take seriously.

Smith introduced *stationary repetition* – several consecutive trading periods with the same endowed values and costs but no carry-over from one period to the next, so that subjects have the opportunity to adapt to the trading environment. He also brought the *continuous double auction* (CDA) market (sometimes referred to as the double oral auction) format into the laboratory: traders can make public, committed offers to buy and to sell and can accept others' offers at any time during a trading period. Variants

of the CDA format predominate in modern financial markets, including the New York Stock Exchange (NYSE), NASDAQ, and the Chicago Mercantile Exchange.

Numerous laboratory studies, beginning with Smith (1962), show that CDA markets with only a few buyers and sellers (say, four of each) reliably produce highly efficient outcomes, where efficiency is defined as the fraction of potential surplus in the market that is captured by the buyers and sellers. Typically, over 95 per cent of total surplus is realized after a few periods of stationary repetition.

Such perishable commodity markets provide no interesting role for time or uncertainty, both important dimensions of financial assets. Laboratory financial markets should allow two-way traders who can both buy and sell, and who trade assets with a payout that is uncertain and/or carries over several periods. Experimenters at Caltech first introduced such markets in the early 1980s. For example, Plott and Sunder (1982) created a single period asset that was traded by six uninformed traders, who knew only that one of two states would occur with given probabilities independently each period, and six informed traders, who knew the realized state. Both informed and uninformed traders were distributed evenly across three types of state-contingent dividend schedules. Within a few periods, prices became highly efficient, and the trading patterns demonstrated that the market fully disseminated the private information. About the same time, several teams of researchers found very efficient asset prices in laboratory markets with assets paying individual- and state-contingent dividends over several trading periods. These and other early laboratory experiments demonstrated that futures and options contracts can speed convergence towards efficient asset prices. See Sunder (1995) for a thorough survey.

The main lesson from these studies is that financial markets can process information very efficiently. As Hayek (1945) conjectured, markets can fully aggregate and disseminate dispersed private information, and can do so quite rapidly. A few bids and asks in the CDA suffice to fully inform experienced traders, dealing appropriate assets, in moderately complex environments.

## Dissecting financial markets

These positive early results encourage us to look more deeply at how financial markets process information. The process has several logical stages. Investors and other participants acquire relevant information from diverse sources, public and private. Individual investors incorporate the information into their beliefs about future asset prices. Acting on their beliefs, investors try to buy assets they expect to appreciate relatively rapidly and to sell assets that they expect to do less well. Their buy and sell orders in turn produce observable market outcomes such as asset price and trading volume. The market outcomes provide further public information for investors, other new information arrives from time to time, and so the process continues. We now know that the process can work quite well in favourable circumstances. But even the early laboratory studies show that it is sometimes fallible. When and where might it go wrong?

Each stage of the process can be examined in the laboratory and compared with theoretical predictions. Cognitive scientists focus on the first stage, the formation of beliefs given arriving information, and have documented many biases that might distort

beliefs. Examples include overconfidence, the gambler's fallacy (believing that a coin that has come up 'heads' many times in succession is the more likely to come up 'tails') and the hot-hand fallacy (believing that basketball players who have made ten free throws in succession are especially likely to make the next). In the next stage, investors may make decision errors when they buy and sell assets, even when their beliefs are realistic. There are numerous examples, including hyperbolic (or quasi-hyperbolic) discounting, the disposition effect, and the sunk-cost fallacy.

It is often tempting to explain financial market anomalies simply by pointing to one or more of these biases and errors. But such explanations are incomplete and potentially erroneous. One problem is that there are so many documented biases and errors; indeed, a complete list seems not to exist. Given any market anomaly A, a diligent student can always find some decision error or bias B that superficially seems connected, whether or not B really causes A. Even more important, investors' biases and decision errors never translate directly into financial market imperfections. Asset prices are non-trivial functions of investors' buy and sell orders, and they provide information that affects subsequent orders and prices. These later stages of the process depend on the market format, and they can attenuate or amplify investors' biases and errors.

## Attenuating biases and errors

Three different market forces can greatly attenuate the financial market impact of erratic investors. First, it is a powerful learning experience to lose money in a financial market, or even to see other investors do better when they have no informational advantage. Friedman (1998) and later studies demonstrate that people can overcome even the strongest biases and errors in a suitable learning environment. To the extent that a bias or error leads to clearly inferior performance, an investor will learn to do better over time. Subjects in most laboratory financial markets commit fewer errors and trade more efficiently in later periods than in earlier periods, and subjects with previous experience in a particular laboratory market do better yet.

Second, the market shares of investors with inferior trading strategies tend to shrink over time, reducing their influence on market performance. Blume and Easley (1992) demonstrate theoretically that wealth redistribution eventually eliminates all but the most effective investors. Laboratory studies routinely cancel out this force via stationary repetition, but it can easily be inferred by compounding relative profits across periods.

Third, persistent costly errors and biases create profit opportunities for entrepreneurs whose efforts attenuate (or even eliminate) the market impact. For example, yellow pages and speed dials help us overcome our cognitive limitations in remembering phone numbers. Similarly, mutual funds and a host of investor advisory services allow investors to sidestep their personal biases. Such entrepreneurs can create new problems but, as noted below, those problems also can be studied in the laboratory. Arbitrage is the most direct form of such entrepreneurship. If error-prone investors create an asset price discrepancy, this will attract profit-seeking arbitrageurs whose buy and sell orders tend to make it disappear. Laboratory studies, including those of Plott and Sunder (1982), confirm the power of arbitrage.

## Amplifying biases and errors

There are also three strong forces that can amplify the market impact of errant investors. First, raw information is often gathered, analysed and released by individuals who have major personal stakes in the market reaction. Despite oversight by authorities such as the US Securities and Exchange Commission, these individuals may use their discretion to distort the market reaction. Bloomfield and O'Hara (1999) and subsequent laboratory studies confirm the possibility.

Second, professional fund managers typically are compensated (directly or indirectly, via competing job offers) for returns that rank highly relative to their peers. It is difficult to infer from field data whether such incentives have an impact, but inference is straightforward in the laboratory. James and Isaac (2000) find major distortions of laboratory asset prices when traders have rank-based performance incentives, and the distortions disappear in otherwise identical markets when traders are paid only their own realized returns.

Third, and most intriguingly, investors may go astray when they try to glean information from the trades of informed investors. Information mirages (for example, Camerer and Weigelt, 1991) can arise as follows. Uninformed trader A observes trader B attempting to buy (due to some slight cognitive bias, say) and mistakenly infers that B has favorable inside information. Then A tries to buy. Now trader C infers that A (or B) is an insider and tries to mimic their trades. Other traders follow, creating a price bubble.

Several research teams (including the author's) have occasionally observed such episodes in the laboratory. They cannot be produced consistently, because incurred losses teach traders to be cautious when they suspect the presence of better-informed traders. The lesson does not necessarily improve market efficiency, since excessive caution impedes information aggregation.

Price bubbles deserve longer discussion, as bubbles have produced important distortions in market prices. Asset prices seemed to disconnect from fundamental value in Japan in the late 1980s, in the dot.com bubble and crash of 1997–2002, and in a number of other episodes since the famous 17th and 18th century events now known as tulipmania and the South Sea bubble. Do such episodes indicate dysfunctional financial markets? Perhaps, but the field data also can be interpreted merely as unusual movements in fundamental value (Garber, 1989). By contrast, in the laboratory the experimenter can always observe (or more typically, control) the fundamental value, so bubbles can be detected and measured precisely.

Smith, Suchanek and Williams (1988) found large positive bubbles, and subsequent crashes, for long-lived laboratory assets and inexperienced traders. Figure 1 shows a representative example. The expected dividend is constant, so the fundamental value (the sum of expected remaining dividends) declines steadily over the 15 trading periods. Ask ('offer') and bid prices start low, but by the second period the transaction prices (indicated by lines connecting accepted bids and asks) rise above fundamental value. The bubble inflates rapidly until late in period 4. In period 9, prices crash below fundamental value.

Keynes's 'greater fool' theory provides a possible interpretation. Traders who themselves have no cognitive bias might be willing to buy at a price above fundamental

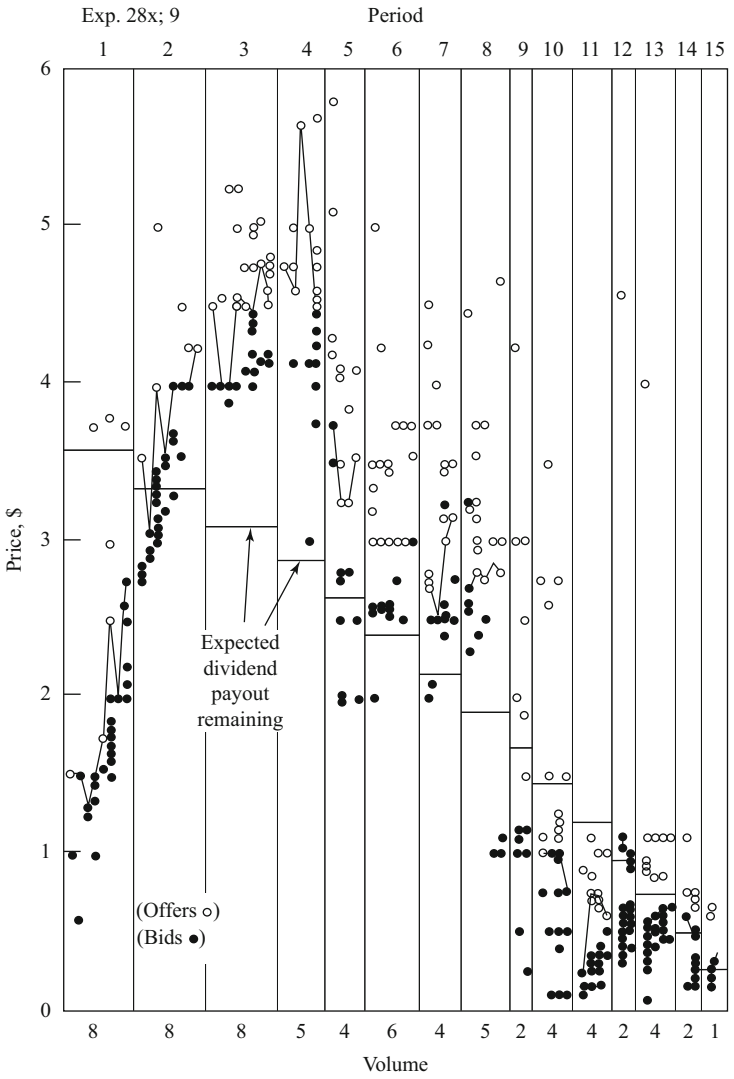



**Figure 1** A bubble and crash in the laboratory. *Source*: Smith, Suchanek and Williams (1988, Figure 9).

value because they expect to sell later at even higher prices to other traders dazzled by rising prices. Subsequent studies confirm that such dazzled traders do exist, and that bubbles are more prevalent when traders are less experienced (individually and as a group), have larger cash endowments, and have less conclusive information.

## Current frontiers: market formats, agents, and prediction markets

Which underlying biases and errors are most important? When does attenuation predominate, and when does amplification? Accumulating laboratory evidence inspires new theoretical and empirical field work as well as follow-up laboratory studies.

It is increasingly clear that answers hinge on the market format or institution – the rules that transform bids and asks into transactions. In particular, the CDA format allows all traders to observe other traders' attempts to buy and sell in real time, and thereby encourages information dissemination. The CDA format attenuates the impact of erratic traders because the closing price is not set by the most biased trader or even by a random trader. The most optimistic traders buy (or already hold) and the most pessimistic traders sell (or never held) the asset, so the closing price reflects the moderate expectations of marginal traders (see Smith, Vernon).

Other traditional formats include the call market (CM), in which bids and asks (or limit orders) are gathered and executed simultaneously at a uniform price, and the posted offer (PO), in which one side (usually sellers) simultaneously announces prices and the other side (buyers) choose transaction quantities at the given prices. Many other formats and hybrids are possible in the Internet age. Which formats are most efficient? Which can attract market share from other formats? Work so far indicates that the CM format does relatively well for thinly traded assets and the PO format works best when the posting side is more concentrated; but the questions remain far from settled.

Related new work blurs the line between computer simulations and laboratory markets. Computer algorithms for artificial agents, or bots, incorporate specified cognitive limitations, and simulations examine the market level impact (for example, Arthur et al., 1997). Gode and Sunder (1993) showed that simple perishables CDA markets are quite efficient even when populated by zero intelligence (ZI) agents, bots that are constrained not to take losses but are otherwise quite random. Current work puts ZI and more intelligent bots into the same asset markets as human traders, and compares efficiency and the distribution of surplus. Such work should help inform regulators, reformers, and entrepreneurs creating new asset markets. Early published examples of policy-oriented research includes performance assessment of (*a*) trader privileges such as price posting and access to order flow information (for example, Friedman, 1993), and (*b*) transaction taxes, price change limits and trading suspensions intended (typically ineffectively) to mitigate price bubbles and panics (for example, Coursey and Dyl, 1990).

Prediction markets, which use the information-aggregation property of markets to forecast events such as election outcomes, are gaining increased attention. The Iowa Electronic Market, designed and operated by experimental economists (Berg et al., 2008), offers various assets that pay the holder ten dollars if (and only if) a specified event occurs by a specified date. Participants self-select, are not representative of the general public, and their trades exhibit partisan bias – for example, self-styled Democrats are more likely to buy assets that pay off when the Democratic Party candidates win. Nevertheless, political event asset prices have consistently outperformed opinion polls and all other available predictors. Prediction markets are a growing presence on the Internet, for example tradesports.com, and some corporations such as HP are beginning to rely on them when making business decisions. The line between laboratory and field financial markets is beginning to blur.

DANIEL FRIEDMAN

*See Also* **behavioural finance; Smith, Vernon.**

### *Bibliography*

Arthur, W.B., Holland, J.H., LeBaron, B., Palmer, R. and Taylor, P. 1997. Asset pricing under endogenous expectations in an artificial stock market. In *The Economy as an Evolving Complex System II*, ed. W.B. Arthur, S.N. Durlauf and D.A. Lane. Reading, MA: Addison-Wesley.

Berg, J., Forsythe, R., Nelson, F. and Rietz, T. 2008. Results from a dozen years of election futures markets research. In *Handbook of Experimental Economics Results*, ed. C. Plott and V. Smith. Amsterdam: North-Holland (forthcoming).

Bloomfield, R.J. and O'Hara, M. 1999. Market transparency: who wins and who loses? *Review of Financial Studies* 12, 5–35.

Blume, L. and Easley, K. 1992. Evolution and market behavior. *Journal of Economic Theory* 58, 9–40.

Camerer, C. and Weigelt, K. 1991. Information mirages in experimental asset markets. *The Journal of Business* 64, 463–93.

Coursey, D.L. and Dyl, E.A. 1990. Price limits, trading suspension, and the adjustment of prices to new information. *Review of Futures Markets* 9, 343–60.

Friedman, D. 1993. How trading institutions affect financial market performance: some laboratory evidence. *Economic Inquiry* 31, 410–35.

Friedman, D. 1998. Monty Hall's three doors: construction and deconstruction of a choice anomaly. *American Economic Review* 88, 933–46.

Garber, P.M. 1989. Tulipmania. *Journal of Political Economy* 97, 535–60.

Gode, D.K. and Sunder, S. 1993. Allocative efficiency of markets with zero intelligence traders: market as a partial substitute for individual rationality. *Journal of Political Economy* 101, 119–37.

Hayek, F.A. 1945. The use of knowledge in society. *American Economic Review* 35, 519–30.

Holt, C.A. 1999. Y2 K bibliography of experimental economics and social science asset market experiments. Online. Available at http://people.virginia.edu/ ~cah2k/assety2k.htm, accessed 19 February 2007.

James, D. and Isaac, R.M. 2000. Asset markets: how they are affected by tournament incentives for individuals. *American Economic Review* 90, 995–1004.

Plott, C.R. and Sunder, S. 1982. Efficiency of experimental security markets with insider information: an application of rational-expectations models. *Journal of Political Economy* 90, 663–98.

Smith, V.L. 1962. An experimental study of competitive market behavior. *Journal of Political Economy* 70, 111–37.

Smith, V.L., Suchanek, G.L. and Williams, A.W. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica* 56, 1119–51.

Sunder, S. 1995. Experimental asset markets: a survey. In *The Handbook of Experimental Economics*, ed. J.H. Kagel and A.E. Roth. Princeton, NJ: Princeton University Press.

# Law, John (1671–1729)

John Law of Lauriston has been regarded by some observers as a monetary crank, by others as a precursor of modern schemes of managed money and Keynesian full-employment policies. He was the originator of the Mississippi Bubble, perhaps the greatest speculative bubble of all time.

Born in Edinburgh, the son of prosperous parents, Law was well educated in political economy. A fugitive from justice in 1694 for killing a man in a duel in England, Law travelled extensively throughout Europe, observing and gaining experience in banking, insurance and finance. He proposed a number of unsuccessful schemes to set up a national bank of issue – in Paris in 1702, Edinburgh in 1705 and Savoy in 1712 – finally attaining success in France with the establishment in 1718 of the Banque Royale.

Law's theories on money and banking are contained in *Money and Trade Considered: With a Proposal for Supplying the Nation With Money* (1705) and other works (Hamilton, 1968; Harsin, 1934). Like other 18th-century writers Law adopted a disequilibrium theory of money, viewing it as a stimulant to trade. In a state of unemployment, Law maintained that an increase in the nation's money supply would stimulate employment and output without raising prices since the demand for money would rise with the increase in output. Moreover, once full employment was attained the monetary expansion would attract factors of production from abroad, so output would continue to increase.

According to Law, a paper-money standard was preferable to one based on precious metals. Suitable candidates for the money supply included government fiat, banknotes, stocks and bonds. Since the primary function of money was as a medium of exchange, it could best be served by a commodity (paper) not subject to considerable fluctuation in value and high resource costs. Thus Law advocated the establishment of note-issuing national banks that would extend productive loans (real bills), providing sufficient currency to guarantee prosperity. Two proposals for such banks, in Paris 1702 and Edinburgh 1705, would have had the note issues based on land initially valued in terms of silver.

From 1716 to 1720 John Law had the unique opportunity to apply his theories to the French economy. In 1715, the heritage of two exhausting wars was depression and deflation. Law succeeded in convincing the Regent (the Duke of Orleans) that a bank of issue would alleviate the problem of financing the national debt. Accordingly, he established in Paris on 2 May 1716 a private bank, the Banque Générale. In its 31 months of operation, the bank was remarkably successful; its notes (convertible into specie and payable as taxes) were issued in moderation and gained national circulation. On 4 December 1718, the Banque Générale was nationalized and renamed the Banque Royale, with Law in control, and in January 1719 it began to issue notes denominated in *livres tournois*, the unit of account, replacing the previously issued *écus de banque* representing fixed amounts of specie.

Alongside the bank, in August 1717, Law established the Compagnie d'Occident after obtaining the franchise on Louisiana and the monopoly of the Canadian fur trade. This company in the succeeding 22 months acquired the tobacco monopoly, the East

India Company and the trading monopolies to Africa and China. Law changed its name in June 1719 to the Compagnie des Indes, and the following winter obtained the farm of the royal mints and of the indirect taxes. In October 1719 he refunded the national debt of 1.5 million *livres tournois*, and in January 1720 became Finance Minister.

The stock of the Compagnie des Indes, initially selling at a par value of £500, within half a year in an unprecedented speculative mania was bid up to many times its original price. The bubble burst in January 1720 after the price of the stock reached a peak of £18,000. To support the price Law made the mistake of pegging it at £9,000, thereby monetizing it and engendering a rapid expansion of notes (125 per cent in two months). In May 1720, in a desperate attempt to salvage his system Law issued a deflationary decree depreciating the stock and reducing the denomination of notes by stages. This decree led to a panic as the public, fearful of further capital losses, sold off both notes and stock. Law's dismissal by the Regent worsened the panic. He was quickly reinstated but his final attempt to restore confidence by reducing the outstanding note issue proved unsuccessful. By December 1720 the 'system' collapsed. Law fled to Belgium and payments quickly reverted to a specie basis. The collapse of the system ruined many in all walks of life and made the word 'bank' anathema in France for well over a century.

Though Law's system reduced unemployment and stimulated output, it was at the expense of doubling the price level. His system was undermined by his actions breaking the link between the note issue and specie convertibility; by retiring the national debt with bank notes convertible into stock; and by encouraging speculation in stock by declaring dividends unrelated to the company's true prospects. Monetizing the stock by pegging its price in the end destroyed the public's confidence in his system. Law was aware of many of the principles of sound money and banking, but by equating money with stock and relying on the real bills doctrine he sowed the seeds of disaster.

MICHAEL D. BORDO

### *Selected works*

1705. *Money and Trade Considered: With a Proposal for Supplying the Nation With Money*. New York: Augustus Kelley, 1966.

### *Bibliography*

Blaug, M. 1978. *Economic Theory in Retrospect*. 3rd edn. Cambridge: Cambridge University Press.

Hamilton, E.J. 1936. Prices and wages at Paris under John Law's system. *Quarterly Journal of Economics* 51, 42–70.

Hamilton, E.J. 1968. Law, John. In *International Encyclopedia of the Social Sciences*, vol. 9. New York: The Free Press.

Hamilton, E.J. 1969. The political economy of France at the time of John Law. *History of Political Economy* 1, 123–49.

Harsin, P., ed. 1934. *John Law: Oeuvres complètes*. 3 vols. Paris: Sirey.

Kindleberger, C.P. 1984. *A Financial History of Western Europe*. London: Allen & Unwin.

Neal, L. and Schubert, E. 1985. The first rational bubble: a new look at the Mississippi and South Sea schemes. Faculty Working Paper No. 1188, Bureau of Economic and Business Research, University of Illinois at Urbana-Champaign.

Rist, C.W. 1940. *A History of Monetary and Credit Theory from John Law to the Present Day*. London: Allen & Unwin.

Schumpeter, J.A. 1954. *History of Economic Analysis*. New York: Oxford University Press.

# Lehman Brothers bankruptcy, what lessons can be drawn?

What is in a name? In the case of Lehman Brothers the name has two different and distinct meanings. Prior to the autumn of 2008, Lehman Brothers referred to one of the oldest investment banks in the USA, with roots in the cotton exchange of the mid-19th century. At the time it filed for protection under Chapter 11 of the US Bankruptcy Code, Lehman Brothers Holdings International was the fourth largest US investment bank and the largest bankruptcy on record. Today Lehman Brothers, used synonymously with the Lehman Brothers bankruptcy filing, is commonly used to refer to an important episode during the 2007–2009 financial crisis. To borrow a line from Winston Churchill, the Lehman Brothers bankruptcy filing on 15 September 2008 did not represent the beginning of the end of the financial crisis, but rather marked the end of the beginning.

## Just the facts

In the 1960s police drama *Dragnet*, the main character Sergeant Joe Friday would direct witnesses to give him 'just the facts'. So what are the facts concerning the episode of the financial crisis attributed to the Lehman bankruptcy?

The Lehman Brothers bankruptcy filing occurred during a period of market turmoil which intensified in the days that followed. Financial markets continued to exhibit signs of increased stress thereafter and during the autumn of 2008. Yields in short-term markets spiked during the week following the Lehman filing. Risk spreads in short-term credit markets widened –indicating a 'flight to quality' by market participants. For example, the 3-month term LIBOR-OIS spread, an indicator of market stress (Thornton, 2009), increased around 14.75 basis points from the Friday before the Lehman bankruptcy filing to 16 September, the day after. From 16 September to 10 October the LIBOR widened by another through 263 basis points. Increased market stress was also evident in the credit default swaps (CDS) market, where the cost of buying credit protection rose sharply in the days just after the Lehman Brothers bankruptcy filing. The five-year CDX.NA.IG index (which is an index of credit default swaps written against North American investment grade companies from Markit and Bloomberg) rose 55 basis points, a 36% increase from 12 September to 17 September. The CDX.NA.IG index declined from its 17 September peak to the end of the month, but still finished September some 20 basis points higher than where it started.

The financial turbulence in the autumn of 2008 was the product of a series of events. The Lehman bankruptcy was one of nearly two dozen significant disruptive events in September 2008 alone, some unrelated to the Lehman bankruptcy filing and some related to its failure. Notable among the economically significant events is the placement of Fannie Mae and Freddie Mac in conservatorship by the Federal Housing Finance Authority, the Federal Reserve assisted rescue of AIG by the US Treasury, and the deathbed acquisition of Merrill Lynch by Bank of America Corporation. Also, notable is the

Reserve Primary Money Fund announcement that it had 'broken the buck': due to losses on its holdings of Lehman debt, the net asset value of the Fund's shares had fallen to $0.97 a share. It was only the second time since the SEC adopted rules governing money market mutual funds in 1983 that a money market fund's share value had fallen below one dollar. Runs on money market mutual funds (MMMFs) would follow.

## Interpreting the facts

While the facts about what happened and when are clear, the connections between them are not. Drawing inferences from any single event is problematic at best. Just as any single point on a plane is consistent with an infinite number of lines, a single event may not allow one to discriminate between numerous different hypotheses. Not surprisingly, there are two different interpretations of the facts associated with Lehman and they arrive at diametrically opposed positions as to causation, and the implications of it for the use of the Bankruptcy Code to handle failing financial firms.

One of the most contentious issues emanating from the Lehman Brothers episode is whether the bankruptcy process is, or with modifications could be, a suitable method for handling the failure of complex, non-bank financial firms. Opinions are sharply divided on the adequacy of US bankruptcy law to resolve complex non-bank financial firms in an orderly fashion. Bankruptcy scholars argue that the market turmoil in the aftermath of the Lehman bankruptcy had little to do with the use of bankruptcy to resolve it, and that in the face of the complexity inherent in resolving an institution the size and scope of Lehman Brothers, the bankruptcy was orderly. In other words, there was no causation running from the bankruptcy filing to the disorderly markets that followed. Proponents of this view argue that the near collapse of markets following Lehman's bankruptcy filing was the result of policy uncertainty: The US government decided to let Lehman fail when the market expected a government-assisted rescue. In fact, Lehman was not prepared for its bankruptcy filing, ostensibly because its management expected government intervention to prevent this outcome (Miller, 2010).

The other view, which one might call the official view of the Lehman episode, is that Lehman's filing for protection is articulated by the Federal Deposit Insurance Corporation (FDIC), among others, which interpreted the facts as supporting a causal relationship between the financial turmoil following Lehman's bankruptcy filing and the use of bankruptcy to resolve Lehman. Under this view, the near collapse of markets in the days following the bankruptcy filing was a direct result of a disorderly windup of Lehman's affairs. Under this interpretation of events in the autumn of 2008 the answer is clear – an orderly resolution of the insolvency of a large financial firm cannot be done in bankruptcy.

This debate is largely unsettled. Even the Dodd–Frank Wall Street Reform and Consumer Protection Act of 2010 (DFA) appears to codify both positions. Title II of DFA creates the Orderly Liquidation Authority (OLA), an administrative receivership process under the FDIC to resolve systemic financial companies. OLA is, however, an exceptional power for resolving systemic non-bank financial firms; bankruptcy remains the default. In addition, DFA mandates that systemic financial companies create and maintain 'living wills': resolution plans for dismantling them in bankruptcy.

Understanding the lessons of the episode during the financial crisis identified with the Lehman Brothers bankruptcy filing requires a careful accounting of the cluster of events that surrounded it. Moreover, no analysis would be complete without an analysis of the role of incentives and expectations in the setup and propagation of the financial crisis. Studying the entire mosaic of the Lehman Brothers episode is necessary to provide context to the period in question and proper attribution of the effects of the bankruptcy filing on the subsequent market turmoil.

As the Lehman episode represents one point in financial history it is impossible to prove or disprove any reasonable interpretation of it. It is possible to, however, to point to some lessons that can be drawn from it. These lessons concern whether the insolvency of large or complex financial companies can be adequately handled through the judicial process of bankruptcy. Moreover, an understanding the Lehman Brothers episode may point to types of reforms to the Code that may be required if bankruptcy is to be a viable option for handling large complex financial firms and a desirable alternative to *ad hoc* bailouts or to resolution under the DFA's Orderly Liquidation Authority.

## International issues

Every country's insolvency regime is inherently complicated by its jurisdictional boundaries. Systemically important financial institutions do not operate in a single country, nor do they have all of their assets located in a single jurisdiction. When Lehman filed for bankruptcy, it operated nearly 3,000 US and foreign chartered separate entities in 20 countries, and its complex legal structure was virtually unrelated to its operational structure (Cumming and Eisenbeis, 2010). This made it incredibly difficult to determine what assets were in each entity in a bankruptcy estate. Further complicating this, substantial sums were transferred between Lehman's cross-border subsidiaries on the eve of bankruptcy.

While Lehman's global presence added substantial complexity to the resolution process, it is difficult to argue that this complexity is a shortcoming of US bankruptcy law. US bankruptcy law has provisions to address cross-border insolvencies (Chapter 15), but these do not guarantee effective or efficient operation. Each country has its own insolvency regimes, and there is substantial variation in their treatment of creditors. This is an issue present whenever a global institution is resolved under any bankruptcy scheme, and to date very little has been done to address it. The United Nations Commission on International Trade Law has developed a model law on cross-border insolvency, but it has not yet been adopted by a sufficient number of jurisdictions to be meaningfully operable. In some sense the international issues raised by the failure of Lehman is immaterial to the insolvency regime debate in the USA. Nonetheless, one lesson that can be learned from the Lehman bankruptcy is that there is plenty of room for improvement in cross-border insolvency regimes.

## US bankruptcy law and complex financial institutions

Irrespective of international issues, some analysts maintain that it was Lehman's use of the bankruptcy courts that caused the market turmoil. They often point to the increased financial turmoil during the week following Lehman's bankruptcy filing as

evidence of the insufficiency of bankruptcy law to resolve complex financial firms. Others claim that it was not the use of bankruptcy, but rather policy responses inconsistent with market expectations that caused markets to panic. That is, Lehman was allowed to fail when financial markets, and even the Lehman management team, expected a government-assisted rescue. A closer look at events around that time suggests that neither view is entirely correct.

The Lehman bankruptcy occurred during a time when there were good reasons for market participants to question the solvency of a number of large financial firms. As noted above, the bankruptcy was accompanied by nearly two dozen significant disruptive events in September 2008 alone. The clustering of multiple events around the time of the bankruptcy makes it difficult to identify the causal effects of the bankruptcy on markets, let alone the effect of the use of US bankruptcy law.

While Lehman's failure triggered many problems in markets, event clustering makes it impossible to identify empirically the use of bankruptcy courts as the root of those problems. Moreover, it is impossible to separate out the impact of Lehman's bankruptcy filing from the uncertainty created by its filing.

Studies have shown that such uncertainty can have significant effects on markets. For example, in 1982 Penn Square Bank was liquidated by the FDIC, which experimented with modified payouts to resolve large bank failures (Furlong, 1984). These modified payouts created uncertainty in the minds of the large, explicitly uninsured creditors of Continental Illinois as to whether they were exposed to losses in the event Continental was closed. This uncertainty drove the run on Continental Illinois' deposits before its collapse in 1984 (Sprague, 1986).

The source of market turmoil following Lehman's failure, then, cannot conclusively be attributed either to the use of bankruptcy law to resolve the firm's insolvency or to the uncertainty created by policy actions inconsistent with market expectations.

## Bankruptcy and contagion

When a large, complex financial firm fails, the method of resolution should not be conducive to contagion. That is, the resolution process should not endanger the solvency of other firms. This is especially true in systemic crises, when the financial system is already stressed. Bankruptcy critics often argue that bankruptcy law may trigger contagion because it is designed to pay creditors strictly according to the priority of their claims. There is no consideration of their financial condition or potential market instability. Thus, contagion may spread through the use of bankruptcy if the recovery of creditors in need of liquidity is insufficient, or indirectly through CDS written on the resolved firm's debt. But the Lehman bankruptcy does not support the view that bankruptcy leads to contagion.

As mentioned above, the day after Lehman Brothers filed for bankruptcy, the Reserve Primary Money Fund announced that it had 'broken the buck': this reflected how large an impact Lehman's collapse was having.

Most analysts would concede that the Fund's 'breaking the buck' was a direct consequence of the Fund's losses on its holdings of Lehman debt, that the losses led to contagion, and that the contagion effects impacted the money market mutual fund

industry and the commercial paper market thereafter. It is harder to argue that the structure of US bankruptcy law, and not the insolvency of Lehman itself, was responsible for the losses on Lehman debt and the subsequent contagion. It may also be the case that the contagion effects were more a consequence of the money market funds' overexposure to Lehman and to a specific feature of the money funds themselves – the pegging of the share price to $1. The share-price peg creates incentives for retail customers to run on a fund when its ability to maintain the peg becomes uncertain. Customers believe it is in their best interest to run to ensure par redemption of their money-fund shares.

Lehman's bankruptcy also tested the CDS market, as there was a reported $400 billion of credit protection written against Lehman's debt. At the time of its bankruptcy, Lehman was the largest failure to be handled in the CDS market. For the purpose of settling the CDS contracts, Lehman's debt was determined to be worth 9.75 cents on the dollar at an International Swaps and Derivatives Association auction, lower than the pre-auction estimates of 12 to 15 cents. However, the settlement of credit protection written on Lehman did not have material effects on financial markets (Summe, 2009; Senior Supervisors Group, 2009).

## Bankruptcy and qualified financial contracts

Derivatives and repos are special types of contract called qualified financial contracts (QFCs), which are exempt from the trust avoidance powers of the Bankruptcy Code and the automatic stay. The trust avoidance provisions and automatic stay are designed to coordinate creditor payouts and ensure that they occur according to the priority of the claims that existed when the original agreements were made. These provisions are designed to prevent a race to grab a firm's assets on the eve of failure or after the firm fails. Instead of being stayed and handled through the bankruptcy estate, each counterparty may close out, net, and settle its QFCs before other debts are paid in bankruptcy. In a sense, QFCs are super priority claims, as they are settled before all others. The special treatment of QFCs may complicate the process of reorganising financial companies in bankruptcy by allowing counterparties to grab assets before the claim priority provisions take hold, but bankruptcy experts disagree about the effect of the QFC exemption in bankruptcy. There is even disagreement on how well Lehman's QFC book, the largest in history to be handled in bankruptcy, was dealt with.

While Lehman's reorganisation has provided additional guidance on which financial contracts are exempted from the automatic stay and how QFCs will be handled in bankruptcy, there is still disagreement on how well bankruptcy handles QFCs. Generally opinions fall into one of two schools of thought. First, there are those who argue that the QFC exemption was an obstacle to an orderly resolution in the Lehman case. In testimony before a House subcommittee in 2009, Harvey Miller, the lead bankruptcy attorney for Lehman, argued that the exemption of some 930,000 derivative counterparties from the automatic stay led to a massive destruction of value through counterparties canceling their contracts. Ayotte and Skeel (2010) and Roe (2011) argue that the safe harbour provisions of bankruptcy for QFCs create perverse incentives for counterparties. Those incentives contribute to the systemic implications of a firm's

failure, including creating a stampede for the exits, which inhibit orderly resolution under bankruptcy.

Second, there are those who argue that Lehman's derivatives portfolio was handled effectively *because of* the exemption from the automatic stay. Kimberly Anne Summe, a former managing director at Lehman, provided this interpretation of the impact of Lehman's counterparties cancelling their contracts on the value of Lehman's estate. Summe noted that only around 3% of Lehman's derivative contracts remained in the bankruptcy estate 106 days after the filing, potentially preventing the spread of distress to Lehman's counterparties by allowing them to close out quickly and re-establish their hedges before market conditions changed too dramatically (Summe, 2009). However, the benefit of allowing quick re-hedging is unclear, as is the cost of losing going-concern value (the value of the company as an ongoing entity rather than a liquidated one) due to the stay exemption.

To the extent that the Bankruptcy Code's safe harbour provisions for QFCs are a stumbling block to an orderly resolution of a systemic financial firm, a simple amendment to the Code is the logical fix. In fact, bankruptcy supporters argue for such a change in the law subjecting QFCs to a limited automatic stay, and there appears to be a case for their position. The FDIC enjoys a one-day stay on QFCs in bank receivership cases, and there is little evidence that this limited stay for FDIC receiverships has been a problem. Moreover, when a non-bank financial firm is resolved under the orderly liquidation authority established in the Dodd–Frank Act, QFCs are subject to a one-day stay. Both provisions allow for the transfer of QFCs during the stay. If this stay is priced into QFCs with depository or systemically important financial institutions and US bankruptcy law were changed to parallel the Dodd–Frank provision, markets would not likely be disrupted, and the pricing of QFCs would be identical across counterparties. It would also have the added benefit of giving the bankruptcy estate up to three days to determine what to do with a derivatives book before counterparties could close out and net, provided that the insolvent firm filed on a Friday.

## The scope of US bankruptcy law

The final material stumbling block to an orderly resolution under bankruptcy of a complex financial firm such as Lehman is the exclusion of certain types of businesses from Chapter 11 (which provides for corporate reorganisation). In the case of Lehman, the exclusion of its broker-dealer subsidiary (Lehman Brothers, Inc.) from filing for Chapter 11 complicated the resolution of Lehman Brothers Holdings International. Lehman Brothers, Inc., became the subject of a liquidation proceeding under the US Securities Investor Protection Act four days after Lehman Brothers Holdings International filed for bankruptcy, during which time the brokerage was borrowing from the Federal Reserve Bank of New York under the Primary Dealer Credit Facility.

The absence of government support likely would have complicated the sale. Because it did not have access to the special financing provisions that firms filing under Chapter 11 are entitled to, the brokerage would have lost going-concern value but for its access to the Primary Dealer Credit Facility. While the sale of Lehman's broker--dealer to Barclay's was quickly approved, without government support the sale might

not have been possible under bankruptcy law. Whether this merits a change in US bankruptcy law would have to be addressed separately for each exemption, though some argue that the prohibition of broker-dealers reorganising in bankruptcy no longer makes sense (Skeel, 2010).

## Policy implications

Lehman Brothers Holdings International is not the first, nor likely the last, systemic financial company to run aground. The case is interesting, however, because the failure occurred during the most severe financial crisis in the USA since the Great Depression. The economic and financial market climate in which Lehman failed greatly complicated any resolution method that did not involve taxpayer assistance in the form of capital infusions or blanket guarantees of creditors. Yet Lehman became the poster child for the orderly liquidation authority provisions of Title II of the 2010 Dodd–Frank Act.

Drawing inferences from Lehman about the effectiveness of bankruptcy in dealing with failing financial firms is problematic. It is difficult to use a single data point – the Lehman bankruptcy – to separate out the impact of Lehman's failure, the use of bankruptcy to resolve it, and the policy uncertainty.

Still, Lehman's bankruptcy offers guidance on how to approach future failures of large, complex financial firms. It appears that there are provisions of bankruptcy law that merit review and possible revision. In the absence of those changes, it may be the case that systemically important pieces of an insolvent firm may be more effectively resolved in an administrative proceeding such as the Orderly Liquidation Authority established under Dodd–Frank. But based on the experience with Lehman, there is no clear evidence that bankruptcy law is insufficient to handle the resolution of large, complex financial firms.

THOMAS J. FITZPATRICK IV AND JAMES B. THOMSON

*See Also* **bankruptcy, economics of; credit crunch chronology: April 2007–September 2009; deposit insurance; bankruptcy law, economics of corporate and personal; fall of AIG, the; Fannie Mae, Freddie Mac and the crisis in US mortgage finance; Federal Reserve System; finance (new developments); financial market contagion; Minsky crisis; regulatory responses to the financial crisis: an interim response.**

### *Bibliography*

Ayotte, K. and Skeel, D. A. Jr. 2010. Bankruptcy or bailouts? *The Journal of Corporation Law*, 35(3), 469–98.

Board of Governors of the Federal Reserve System. 2011. *Study on the Resolution of Financial Companies under the Bankruptcy Code*.

Cumming, C. and Eisenbeis, R. A. 2010. Resolving troubled systemically important cross-border financial institutions: is a new corporate organizational form required? *Federal Reserve Bank of New York Staff Report No. 457*.

Federal Deposit Insurance Corporation (FDIC). 2011. The orderly liquidation of Lehman Brothers Holdings Inc. under the Dodd–Frank Act. *FDIC Quarterly*, 5(2), 1–19.

Edwards, F. and Morrison, E. R. 2005. Derivatives and the bankruptcy code: why the special treatment? *Yale Journal on Regulation*, 22, 101–32.

Furlong, F. T. 1984. FDIC's modified payout plan. *Federal Reserve Bank of San Francisco Economic Letter*, 18 May.

Mengle, D. L. 2007. Credit derivatives: an overview. *Federal Reserve Bank of Atlanta Economic Review*, 92(4), 1–24.

Miller, H. R. 2009. Testimony before the Subcommittee on Commercial and Administrative Law of the House of Representatives Committee on the Judiciary, 111th Congress, 1st Session, for Hearings on 'Too Big to Fail: The Role for Bankruptcy and Antitrust Law in Financial Regulation Reform', 22 October.

Miller, H. R. 2010. Testimony before the Financial Crisis Inquiry Commission for Hearing on Examining the Causes of the Current Financial and Economic Crisis of the United States and of the Collapse of Lehman Brothers, 22 October.

Roe, M. J. 2011. The derivatives players' payment priorities as financial crisis accelerator. *Stanford Law Review*, 63, 539–88.

Senior Supervisors Group. 2009. *Observations on Management of Recent Credit Default Swap Credit Events*, Available at: http://www.sec.gov/news/press/2009/report030909.pdf.

Skeel, D. A. 2010. Bankruptcy boundary games. *Brooklyn Journal of Corporate, Financial and Commercial Law*, 4, 1–21.

Skeel, D. 2011. *The New Financial Deal: Understanding The Dodd–Frank Act And Its (Unintended) Consequences*. John Wiley & Sons, Hoboken, NJ.

Sprague, I. H. 1986. *Bailout: An Insider's Account of Bank Failures and Rescues*. Basic Books, New York.

Summe, K. A. 2009. Lessons learned from the Lehman bankruptcy. In: K. E. Scott, G. P. Schultz and J. B. Taylor (eds.) *Ending Government Bailouts as We Know Them*, Chapter 5. Stanford University Press, Stanford CA.

Thornton, D. L. 2009. What the Libor-OIS spread says. *Federal Reserve Bank of St. Louis Economic Synopses*, No. 24.

# liquidity trap

A liquidity trap is defined as a situation in which the short-term nominal interest rate is zero. In this case, many argue, increasing money in circulation has no effect on either output or prices. The liquidity trap is originally a Keynesian idea and was contrasted with the quantity theory of money, which maintains that prices and output are, roughly speaking, proportional to the money supply.

According to the Keynesian theory, money supply has its effects on prices and output through the nominal interest rate. Increasing money supply reduces the interest rate through a money demand equation. Lower interest rates stimulate output and spending. The short-term nominal interest rate, however, cannot be less than zero, based on a basic arbitrage argument: no one will lend 100 dollars unless she gets at least 100 dollars back. This is often referred to as the 'zero bound' on the short-term nominal interest rate. Hence, the Keynesian argument goes, once the money supply has been increased to a level where the short-term interest rate is zero, there will be no further effect on either output or prices, no matter by how much money supply is increased.

The ideas that underlie the liquidity trap were conceived during the Great Depression. In that period the short-term nominal interest rate was close to zero. At the beginning of 1933, for example, the short-term nominal interest rate in the United States – as measured by three-month Treasuries – was only 0.05 per cent. As the memory of the Great Depression faded and several authors challenged the liquidity trap, many economists begun to regard it as a theoretical curiosity.

The liquidity trap received much more attention again in the late 1990s with the arrival of new data. The short-term nominal interest rate in Japan collapsed to zero in the second half of the 1990s. Furthermore, the Bank of Japan (BoJ) more than doubled the monetary base through traditional and non-traditional measures to increase prices and stimulate demand. The BoJ policy of 'quantitative easing' from 2001 to 2006, for example, increased the monetary base by over 70 per cent in that period. By most accounts, however, the effect on prices was sluggish at best. (As long as five years after the beginning of quantitative easing, the changes in the CPI and the GDP deflator were still only starting to approach positive territory.)

### The modern view of the liquidity trap

The modern view of the liquidity trap is more subtle than the traditional Keynesian one. It relies on an intertemporal stochastic general equilibrium model whereby aggregate demand depends on current and expected future real interest rates rather than simply the current rate as in the old Keynesian models. In the modern framework, the liquidity trap arises when the zero bound on the short-term nominal interest rate prevents the central bank from fully accommodating sufficiently large deflationary shocks by interest rate cuts.

The aggregate demand relationship that underlies the model is usually expressed by a consumption Euler equation, derived from the maximization problem of a representative household. On the assumption that all output is consumed, that equation can be approximated as:

$$Y_t = E_t Y_{t+1} - \sigma(i_t - E_t \pi_{t+1} - r_t^e) \tag{1}$$

where $Y_t$ is the deviation of output from steady state, $i_t$ is the short-term nominal interest rate, $\pi_t$ is inflation, $E_t$ is an expectation operator and $r_t^e$ is an exogenous shock process (which can be due to host of factors). This equation says that current demand depends on expectations of future output (because spending depends on expected future income) and the real interest rate which is the difference between the nominal interest rate and expected future inflation (because lower real interest rates make spending today relatively cheaper than future spending). This equation can be forwarded to yield

$$Y_t = E_t Y_{T+1} - \sigma \sum_{s=t}^{T} E_t(i_s - \pi_{s+1} - r_s^e)$$

which illustrates that demand depends not only on the current short-term interest rate but on the entire expected path for future interest rates and expected inflation. Because long-term interest rates depend on expectations about current and future short-term rates, this equation can also be interpreted as saying that demand depends on long-term interest rates. Monetary policy works through the short-term nominal interest rate in the model, and is constrained by the fact that it cannot be set below zero,

$$i_t \geq 0. \tag{2}$$

In contrast to the static Keynesian framework, monetary policy can still be effective in this model even when the current short-term nominal interest rate is zero. In order to be effective, however, expansionary monetary policy must change the public's expectations about future interest rates at the point in time when the zero bound will no longer be binding. For example, this may be the period in which the deflationary shocks are expected to subside. Thus, successful monetary easing in a liquidity trap involves committing to maintaining lower future nominal interest rates for any given price level in the future once deflationary pressures have subsided (see, for example, Reifschneider and Williams, 2000; Jung, Teranishi and Watanabe, 2005; Eggertsson and Woodford, 2003; Adam and Billi, 2006).

This was the rationale for the BoJ's announcement in the autumn of 2003 that it promised to keep the interest rate low until deflationary pressures had subsided and CPI inflation was projected to be in positive territory. It also underlay the logic of the Federal Reserve announcement in mid-2003 that it would keep interest rates low for a 'considerable period'. At that time, there was some fear of deflation in the United States (the short-term interest rates reached one per cent in the spring of 2003, its lowest level since the Great Depression, and some analysts voiced fears of deflation).

There is a direct correspondence between the nominal interest rate and the money supply in the model reviewed above. There is an underlying demand equation for real

money balances derived from a representative household maximization problem (like the consumption Euler equation 1). This demand equation can be expressed as a relationship between the nominal interest rate and money supply

$$\frac{M_t}{P_t} \geq L(Y_t, i_t) \tag{3}$$

where $M_t$ is the nominal stock of money and $P_t$ is a price level. On the assumption that both consumption and liquidity services are normal goods, this inequality says that the demand for money increases with lower interest rates and higher output. As the interest rate declines to zero, however, the demand for money is indeterminate because at that point households do not care whether they hold money or one-period riskless government bonds. The two are perfect substitutes: a government liability that has nominal value but pays no interest rate. Another way of stating the result discussed above is that a successful monetary easing (committing to lower *future* nominal interest rate for a given price level) involves committing to higher money supply *in the future* once interest rates have become positive again (see, for example, Eggertsson, 2006a).

## Irrelevance results

According to the modern view outlined above, monetary policy will increase demand at zero interest rates only if it changes expectations about the future money supply or, equivalently, the path of future interest rates. The Keynesian liquidity trap is therefore only a true trap if the central bank cannot to stir expectations. There are several interesting conditions under which this is the case, so that monetary easing is ineffective. These 'irrelevance' results help explain why BoJ's increase in the monetary base in Japan through 'quantitative easing' in 2001–6 may have had a somewhat more limited effect on inflation and inflation expectations in that period than some proponents of the quantity theory of money expected.

Krugman (1998), for example, shows that at zero interest rates if the public expects the money supply in the future to revert to some constant value as soon as the interest rate is positive, quantitative easing will be ineffective. Any increase in the money supply in this case is expected to be reversed, and output and prices are unchanged.

Eggertsson and Woodford (2003) show that the same result applies if the public expects the central bank to follow a 'Taylor rule', which may indeed summarize behaviour of a number of central banks in industrial countries. A central bank following a Taylor rule raises interest rates in response to above-target inflation and above-trend output. Conversely, unless the zero bound is binding, the central bank reduces the interest rate if inflation is below target or output is below trend (an output gap). If the public expects the central bank to follow the Taylor rule, it anticipates an interest rate hike as soon as there are inflationary pressures in excess of the implicit inflation target. If the target is perceived to be price stability, this implies that quantitative easing has no effect, because a commitment to the Taylor rule implies that any increase in the monetary base is reversed as soon as deflationary pressures subside.

Eggertsson (2006a) demonstrates that, if a central bank is discretionary, that is, unable to commit to future policy, and minimizes a standard loss function that

depends on inflation and the output gap, it will also be unable to increase inflationary expectations at the zero bound, because it will always have an incentive to renege on an inflation promise or extended 'quantitative easing' in order to achieve low *ex post* inflation. This deflation bias has the same implication as the previous two irrelevance propositions, namely, that the public will expect any increase in the monetary base to be reversed as soon as deflationary pressures subside. The deflation bias can be illustrated by the aid of a few additional equations, as illustrated in the next section.

### The deflation bias and the optimal commitment

The deflation bias can be illustrated by completing the model that gave rise to (1), (2) and (3). In the model prices are not flexible because firms reset their price at random intervals. This gives rise to an aggregate supply equation which is often referred to as the 'New Keynesian' Phillips curve. It can be derived from the Euler equation of the firm's maximization problem (see, for example, Woodford, 2003)

$$\pi_t = \kappa(Y_t - Y_t^n) + \beta E_t \pi_{t+1} \tag{4}$$

where $Y_t^n$ is the natural rate of output (in deviation from steady state), which is the 'hypothetical' output produced if prices were perfectly flexible, $\beta$ is the discount factor of the household in the model and the parameter $\kappa > 0$ is a function of preferences and technology parameters. This equation implies that inflation can increase output above its natural level because not all firms reset their prices instantaneously.

If the government's objective is to maximize the utility of the representative household, it can be approximated by

$$\sum_{t=0}^{\infty} \beta^t \{\pi_t^2 + \lambda_y (Y_t - Y_t^e)^2\} \tag{5}$$

where the term $Y_t^e$ is the target level of output. It is also referred to as the 'efficient level' or 'first-best level' of output. The standard 'inflation bias' first illustrated by Kydland and Prescott (1977) arises when the natural level of output is lower than the efficient level of output, that is, $Y_t^n < Y_t^e$.

Eggertsson (2006a) shows that there is also a deflation bias under certain circumstances. While the inflation bias is a steady state phenomenon, the deflation bias arises to temporary shocks. Consider the implied solution for the nominal interest rate when there is an inflation bias of $\overline{\pi}$. It is

$$i_t = \overline{\pi} + r_t^e.$$

This equation cannot be satisfied in the presence of sufficiently large deflationary shocks, that is, a negative $r_t^e$. In particular if $r_t^e < -\overline{\pi}$ this solution would imply a negative nominal interest rate. It can be shown (Eggertsson, 2006a) that a discretionary policymaker will in this case set the nominal interest rate to zero but set inflation equal to the 'inflation bias' solution $\overline{\pi}$ as soon as the deflationary pressures have subsided (that is, when the shock is $r_t^e \geq -\overline{\pi}_t$). If the disturbance $r_t^e$ is low enough, the zero bound frustrates the central bank's ability to achieve its 'inflation target' $\overline{\pi}$ which can in turn

lead to excessive deflation. (While deflation and zero interest rates are due to real shocks in the literature discussed above, an alternative way of modelling the liquidity trap is that it is the result of self-fulfilling deflationary expectations; see, for example, Benhabib, Schmitt-Grohe and Uribe, 2001.)

To illustrate this consider the following experiment. Suppose the term $r_t^e$ is unexpectedly negative in period 0 ($r_t^e = r_L < 0$) and then reverts back to its steady state value $\overline{r} > 0$ with a fixed probability $\alpha$ in every period. For simplicity assume that $\overline{\pi} = 0$. Then it is easy to verify from eqs. (1), (4), the behaviour of the central bank described above and the assumed process for $r_t^e$ that the solution for output and inflation is given by (see Eggertsson, 2006a, for details)

$$\pi_t = \frac{1}{\alpha(1-\beta(1-\alpha)) - \sigma\kappa(1-\alpha)} \kappa\sigma r_L^e \text{ if } r_t^e = r_L^e \text{ and } \pi_t = 0 \text{ otherwise} \quad (6)$$

$$Y_t = \frac{1-\beta(1-\alpha)}{\alpha(1-\beta(1-\alpha)) - \sigma\kappa(1-\alpha)} \sigma r_L^e \text{ if } r_t^e = r_L^e \text{ and } Y_t = 0 \text{ otherwise} \quad (7)$$

Figure 1 shows the solution in a calibrated example for numerical values of the model taken from Eggertsson and Woodford (2003). (Under this calibration $\alpha = 0.1$, $\kappa = 0.02$, $\beta = 0.99$ and $r_L = -\frac{0.02}{4}$ but the model is calibrated in quarterly frequencies.) The dashed line shows the solution under the contingency that the natural rate of

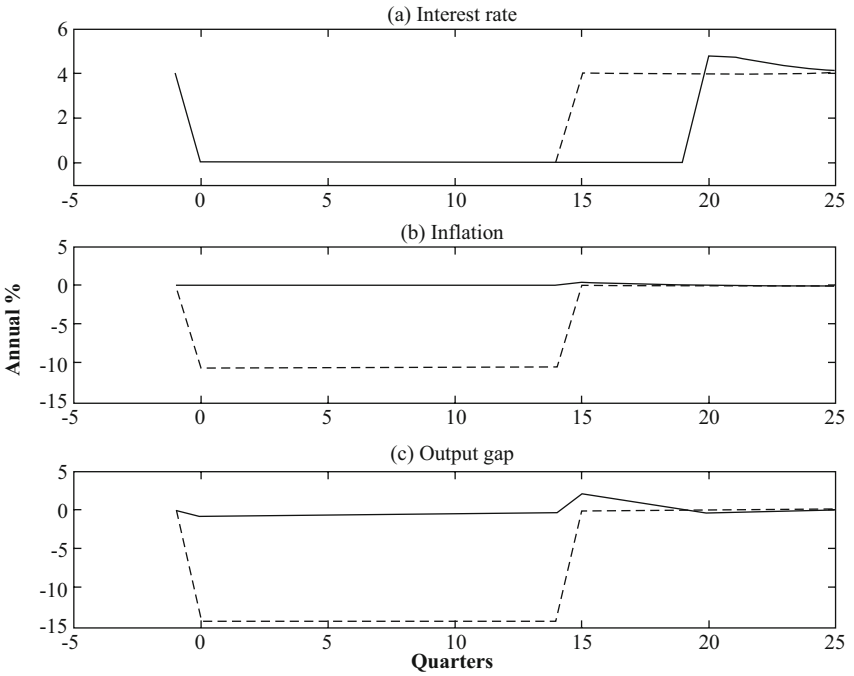



**Figure 1** Response of the nominal interest rate, inflation and the output gap to a shock that lasts for 15 quarters. *Note*: The dashed line shows the solution under policy discretion, the solid line the solution under the optimal policy commitment.

interest reverts to positive level in 15 periods. The inability of the central bank to set negative nominal interest rate results in a 14 per cent output collapse and 10 per cent annual deflation. The fact that in each quarter there is a 90 per cent chance of the exogenous disturbance to remaining negative for the next quarter creates the expectation of future deflation and a continued output depression, which creates even further depression and deflation. Even if the central bank lowers the short-term nominal interest rate to zero, the real rate of interest is positive, because the private sector expects deflation. The same results applies when there is an inflation bias, that is, $\overline{\pi} > 0$, but in this case the disturbance $r_t^e$ needs to be correspondingly more negative to lead to an output collapse.

The solution illustrated in Figure 1 is what Eggertsson (2006a) calls the deflation bias of monetary policy under discretion. The reason why this solution indicates a deflation bias is that the deflation and depression can largely be avoided by the correct *commitment* to optimal policy. The solid line shows the solution in the case that the central bank can commit to optimal future policy. In this case the deflation and the output contraction are largely avoided. In the optimal solution the central bank commits to keeping the nominal interest at zero for a considerable period beyond what is implied by the discretionary solution; that is, interest rates are kept at zero even if the deflationary shock $r_t^e$ has subsided. Similarly, the central bank allows for an output boom once the deflationary shock subsides and accommodates mild inflation. Such commitment stimulates demand and reduces deflation through several channels. The expectation of future inflation lowers the real interest rate, even if the nominal interest rate cannot be reduced further, thus stimulating spending. Similarly, a commitment to lower future nominal interest rate (once the deflationary pressures have subsided) stimulates demand for the same reason. Finally, the expectation of higher future income, as manifested by the expected output boom, stimulates current spending, in accordance with the permanent income hypothesis (see Eggertsson and Woodford, 2003, for the derivation underlying this figures. The optimal commitment is also derived in Jung, Teranishi and Watanabe, 2005, and Adam and Billi, 2006, for alternative processes for the deflationary disturbance).

The discretionary solution indicates that this optimal commitment, however desirable, is not feasible if the central bank cannot commit to future policy. The discretionary policymaker is cursed by the deflation bias. To understand the logic of this curse, observe that the government's objective (5) involves minimizing deviations of inflation and output from their targets. Both these targets can be achieved at time $t = 15$ when the optimal commitment implies targeting positive inflation and generating an output boom. Hence the central bank has an incentive to renege on its previous commitment and achieve zero inflation and keep output at its optimal target. The private sector anticipates this, so that the solution under discretion is the one given in (6) and (7); this is the deflation bias of discretionary policy.

## Shaping expectations

The lesson of the irrelevance results is that monetary policy is ineffective if it cannot stir expectations. The previous section illustrated, however, that shaping expectations in

the correct way can be very important for minimizing the output contraction and deflation associated with deflationary shocks. This, however, may be difficult for a government that is expected to behave in a discretionary manner. How can the correct set of expectations be generated?

Perhaps the simplest solution is for the government to make clear *announcements* about its future policy through the appropriate 'policy rule'. This was the lesson of the 'rules vs. discretion' literature started by Kydland and Prescott (1977) to solve the inflation bias, and the same logic applies here even if the nature of the 'dynamic inconsistency' that gives rise to the deflation bias is different from the standard one. To the extent that announcements about future policy are believed, they can have a very big effect. There is a large literature on the different policy rules that minimize the distortions associated with deflationary shocks. One example is found in both Eggertsson and Woodford (2003) and Wolman (2005). They show that, if the government follows a form of price level targeting, the optimal commitment solution can be closely or even completely replicated, depending on the sophistication of the targeting regime. Under the proposed policy rule the central bank commits to keep the interest rate at zero until a particular price level is hit, which happens well after the deflationary shocks have subsided.

If the central bank, and the government as a whole, has a very low level of credibility, a mere announcement of future policy intentions through a new 'policy rule' may not be sufficient. This is especially true in a deflationary environment, for at least three reasons. First, the deflation bias implies that the government has an incentive to promise to deliver future expansion and higher inflation, and then to renege on this promise. Second, the deflationary shocks that give rise to this commitment problem are rare, and it is therefore harder for a central bank to build up a reputation for dealing with them well. Third, this problem is even further aggravated at zero interest rates because then the central bank cannot take any direct actions (that is, cutting interest rate) to show its new commitment to reflation. This has led many authors to consider other policy options for the government as a whole that make a reflation credible, that is, make the optimal commitment described in the previous section 'incentive compatible'.

Perhaps the most straightforward way to make a reflation credible is for the government to issue debt, for example by deficit spending. It is well known in the literature that government debt creates an inflationary incentive (see, for example, Calvo, 1978). Suppose the government promises future inflation and in addition prints one dollar of debt. If the government later reneges on its promised inflation, the real value of this one dollar of debt will increase by the same amount. Then the government will need to raise taxes to compensate for the increase in the real debt. To the extent that taxation is costly, it will no longer be in the interest of the government to renege on its promises to inflate the price level, even after deflationary pressures have subsided in the example above. This commitment device is explored in Eggertsson (2006a), which shows that this is an effective tool to battle deflation.

Jeanne and Svensson (2007) and Eggertsson (2006a) show that foreign exchange interventions also have this effect, for very similar reasons. The reason is that foreign exchange interventions change the balance sheet of the government so that a policy of reflation is incentive compatible. The reason is that, if the government prints nominal

liabilities (such as government bonds or money) and purchases foreign exchange, it will incur balance-sheet losses if it reneges on an inflation promise because this would imply an exchange rate appreciation and thus a portfolio loss.

There are many other tools in the arsenal of the government to battle deflation. Real government spending, that is, government purchases of real goods and services, can also be effective to this end (Eggertsson, 2005). Perhaps the most surprising one is that policies that temporarily reduce the natural level of output, $Y_t^n$, can be shown to increase equilibrium output (Eggertsson, 2006b). The reason is that policies that suppress the natural level of output create actual and expected reflation in the price level and this effect is strong enough to generate recovery because of the impact on real interest rates.

## Conclusion: the Great Depression and the liquidity trap

As mentioned in the introduction, the old literature on the liquidity trap was motivated by the Great Depression. The modern literature on the liquidity trap not only sheds light on recent events in Japan and the United States (as discussed above) but also provides new insights into the US recovery from the Great Depression. This article has reviewed theoretical results that indicate that a policy of reflation can induce a substantial increase in output when there are deflationary shocks (compare the solid line and the dashed line in Figure 1: moving from one equilibrium to the other implies a substantial increase in output). Interestingly, Franklin Delano Roosevelt (FDR) announced a policy of reflating the price level in 1933 to its pre-Depression level when he became President in 1933. To achieve reflation FDR not only announced an explicit objective of reflation but also implemented several policies which made this objective credible. These policies include all those reviewed in the previous section, such as massive deficit spending, higher real government spending, foreign exchange interventions, and even policies that reduced the natural level of output (the National Industrial Recovery Act and the Agricultural Adjustment Act: see Eggertsson, 2006b, for discussion). As discussed in Eggertsson (2005; 2006b) these policies may greatly have contributed to the end of the depression. Output increased by 39 per cent during 1933–7, with the turning point occurring immediately after FDR's inauguration, when he announced the policy objective of reflation. In 1937, however, the administration moved away from reflation and the stimulative policies that supported it – prematurely declaring victory over the depression – which helps explaining the downturn in 1937–8, when monthly industrial production fell by 30 per cent in less than a year. The recovery resumed once the administration recommitted to reflation (see Eggertsson and Puglsey, 2006). The modern analysis of the liquidity trap indicates that, while zero short-term interest rates made static changes in the money supply irrelevant during this period, expectations about the future evolution of the money supply and the interest rate were key factors determining aggregate demand. Thus, recent research indicates that monetary policy was far from being ineffective during the Great Depression, but it worked mainly through expectations.

GAUTI B. EGGERTSSON

*See Also* **expectations; inflation expectations; optimal fiscal and monetary policy (with commitment); optimal fiscal and monetary policy (without commitment).**

### *Bibliography*

Adam, K. and Billi, R. 2006. Optimal monetary policy under commitment with a zero bound on nominal interest rates. *Journal of Money, Credit and Banking* (forthcoming).

Benhabib, J., Schmitt-Grohe, S. and Uribe, M. 2001. Monetary policy and multiple equilibria. *American Economic Review* 91, 167–86.

Calvo, G. 1978. On the time consistency of optimal policy in a monetary economy. *Econometrica* 46, 1411–28.

Eggertsson, G. 2005. Great expectations and the end of the depression. Staff Report No. 234. Federal Reserve Bank of New York.

Eggertsson, G. 2006a. The deflation bias and committing to being irresponsible. *Journal of Money, Credit and Banking* 38, 283–322.

Eggertsson, G. 2006b. Was the New Deal contractionary? Working paper, Federal Reserve Bank of New York.

Eggertsson, G. and Pugsley, B. 2006. The mistake of 1937: a general equilibrium analysis. *Monetary and Economic Studies* 24(SI), 151–90.

Eggertsson, G. and Woodford, M. 2003. The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity* 2003(1), 212–19.

Jeanne, O. and Svensson, L. 2007. Credible commitment to optimal escape from a liquidity trap: the role of the balance sheet of an independent central bank. *American Economic Review* 97, 474–90.

Jung, T., Teranishi, Y. and Watanabe, T. 2005. Zero bound on nominal interest rates and optimal monetary policy. *Journal of Money, Credit and Banking* 37, 813–36.

Krugman, P. 1998. It's baaack! Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity* 1998(2), 137–87.

Kydland, F. and Prescott, E. 1977. Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 85, 473–91.

Reifschneider, D. and Williams, J. 2000. Three lessons for monetary policy in a low inflation era. *Journal of Money, Credit and Banking* 32, 936–66.

Wolman, A. 2005. Real implications of the zero bound on nominal interest rates. *Journal of Money, Credit and Banking* 37, 273–96.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton: Princeton University Press.

# Minsky crisis

## Introduction

Stability is destabilizing. Those three words capture in a concise manner the insight that underlies Minsky's analysis of the transformation of the economy over the entire post-war period. The basic thesis is that the dynamic forces of the capitalist economy are explosive so that they must be contained by institutional ceilings and floors – part of the 'safety net'. However, to the extent that the constraints successfully achieve some semblance of stability, that will change behaviour in such a manner that the ceiling will be breached in an unsustainable speculative euphoria. If the inevitable crash is cushioned by the institutional floors, the risky behaviour that caused the boom will be rewarded. Another boom will build, and its crash will again test the safety net. Over time, the crises become increasingly frequent and severe until finally 'it' (a great depression with a debt deflation) becomes possible.

While Minsky's 'financial instability hypothesis' is fundamentally pessimistic, it is not meant to be fatalistic (Minsky, 1975, 1982, 1986) According to Minsky, policy must adapt as the economy is transformed. The problem with the stabilizing institutions that had been put in place in the early post-war period is that they no longer served the economy well by the 1980s, as they had not kept up with the evolution of financial institutions and practices. Further, they had been purposely degraded and even in some cases dismantled, often on the erroneous belief that 'free' markets are self-regulating. Indeed, that became the clarion call of most of the economics profession after the early 1970s, based on the rise of 'new' classical economics with its rational agents and instantaneously clearing markets and the 'efficient markets hypothesis' that proclaimed prices fully reflect all information about 'fundamentals'. Hence, not only had firms learned how to circumvent regulations and other constraints, but policymakers had removed regulations and substituted 'self-regulation' in place of government oversight.

From his earliest writings in the late 1950s to his final papers written before his death in 1996, Minsky always analyzed the financial innovations of profit-seeking firms that were designed to subvert New Deal constraints. For example, he was one of the first economists to recognize how the development of the federal funds market had already reduced the Fed's ability to use reserves to constrain bank lending, while at the same time 'stretching' liquidity because banks would have fewer safe and liquid assets should they need to unwind balance sheets (Minsky 1957). And much later, in a remarkably prescient piece in 1987, Minsky had foreseen the development of securitization (to move interest rate risk off bank balance sheets while reducing capital requirements) that would later be behind the global financial crash of 2007 (published as Minsky, 2008) At the same time, Minsky continually formulated and advocated policy to deal with these new developments. Unfortunately, his warnings were largely ignored by the profession and by policymakers – until it was too late.

## Minsky's theory of the business cycle

In the introduction I focused on long-term transformations because too often Minsky's analysis is interpreted as a theory of the business cycle. There have even been some analyses that attempted to 'prove' Minsky wrong by applying his theory to data from one business cycle. Further, the global crisis that began in 2007 has been called the 'Minsky moment' or a 'Minsky crisis'. As I will discuss, I agree that this crisis does fit with Minsky's theory, but I object to analyses that begin with, say, 2004 – attributing the causes of the crisis to changes that occurred over a handful of years that preceded the collapse. Rather, I argue that we should find the causes of the crisis in the transformation that began in 1951. We will not understand the crisis if we begin with a US real estate boom fueled by lending to subprime borrowers. That will be the topic of the next section.

Now, Minsky *did* have a theory of the business cycle (see Papadimitriou and Wray (1998) for a summary of Minsky's approach). He called it 'an investment theory of the cycle and a financial theory of investment'. He borrowed the first part of that from Keynes: investment is unstable and tends to be the driver of the cycle (through its multiplier impact). Minsky's contribution was the financial theory of investment, with his book *John Maynard Keynes* (1975) providing the detailed exposition. In brief, investment is financed with a combination of internal and external (borrowed) funds. Over an expansion, success generates a greater willingness to borrow, which commits a rising portion of expected gross profits (Minsky called it gross capital income) to servicing debt. This exposes the firm to greater risk because if income flows turn out to be less than expected, or if finance costs rise, firms might not be able to meet those debt payment commitments. There is nothing inevitable about that, however, because Minsky incorporated the profits equation of Michal Kalecki in his analysis: at the aggregate level total profits equal investment plus the government's deficit plus net exports plus consumption out of profits and less saving out of wages (Minsky, 1986). The important point is that all else being equal, higher investment generates higher profits at the aggregate level. This can actually make the system even more unstable, because if profits continually exceed expectations, making it easy to service debt, then firms will borrow even more.

This then leads to Minsky's famous categorization of financial positions: a hedge unit can meet payment commitments out of income flow; a speculative unit can only pay interest but must roll over principal; and a Ponzi unit cannot even make the interest payments so must 'capitalize' them (borrowing to pay interest). (In his classification of 'Ponzi finance', Minsky borrowed the name of a famous fraudster, Charles Ponzi, who ran a 'pyramid' scheme – in more recent times, Bernie Madoff ran another pyramid that failed spectacularly.) Over a 'run of good times', firms (and households) are encouraged to move from hedge to speculative finance, and the economy as a whole transitions from one in which hedge finance dominates to one with a greater weight of speculative finance. Eventually some important units find they cannot pay interest, driving them to Ponzi finance. Honest bankers do not like to lend to Ponzi units because their outstanding debt grows continually unless income flows eventually rise. When the bank stops lending, the Ponzi unit collapses. Following Irving Fisher, Minsky then described a 'debt deflation' process: collapse by one borrower can bring down his

creditors, who default on their own debts, generating a snowball of defaults. Uncertainty and pessimism rise, investment collapses and through the multiplier income and consumption also fall, and we are on our way to a recession.

But Minsky did not mean to imply that all financial crises lead to recessions, nor that all recessions result from the transition to speculative and Ponzi finance. The Federal government in the post-war period was big – 20–25% of the economy versus only 3% on the verge of the Great Depression. This meant that government itself could be both stabilizing and destabilizing. Countercyclical movement of its budget from surplus in a boom to deficit in a slump would stabilize income and profits (recall from the Kalecki accounting identity above that government deficits add to profits). A rising deficit could potentially offset the effects of falling investment, and, indeed, over the post-war period that helped to cushion every recession. However, it is also possible for the government to cause a downturn by cutting spending – as it did in the demobilization from the Second World War. And if the budget is excessively biased toward surplus when the economy grows, it will generate 'fiscal drag' that removes household income and profits of firms – causing a recession. For that reason, a recession could occur well before the private sector is dominated by speculative and Ponzi positions. (Note that an economy that moves toward current account deficits when it grows robustly – such as the USA – will suffer an additional 'headwind' that sucks income and profits from domestic households and firms.)

In addition to the 'big government', the post-war period also had what Minsky called the 'big bank' – the Federal Reserve. The Fed plays a number of roles: it sets interest rates, it regulates and supervises banks, and it acts as lender of last resort. Generally, it moves interest rates in a procyclical manner (raising them in expansion and lowering them in recession), which is believed by many orthodox economists to be stabilizing. Like many heterodox economists, Minsky doubted that spending is very interest-sensitive: in a boom, raising rates by a moderate amount will not curb enthusiasm, and in a bust, even very low interest rates cannot overcome pessimism. In addition, Minsky emphasized the impact of interest rates on financial fragility: raising rates in a boom would increase finance costs and hasten the transition to speculative and Ponzi financial positions, hence, to the extent that tight monetary policy 'works', it does so by inducing a financial crisis. Thus, Minsky rejected the notion that the Fed can use interest rates to 'fine tune' the economy.

But lender of last resort policy was viewed by Minsky as essential – it would stop a bank run and would help to put a floor to asset prices, attenuating the debt deflation process discussed above. If the Fed lends to a troubled financial institution, it does not have to sell assets to try to cover demands by creditors for redemption. For example, if depositors are demanding cash withdrawal, in the absence of a lender of last resort the bank would have to sell assets to raise the cash required; this is normally difficult for assets such as loans, and nearly impossible to do in a crisis. So the Fed lends the reserves to cover withdrawals.

In sum, the intervention of the big bank and the big government helps to prevent a financial crisis from turning into a deep downturn. The big government's deficit puts a floor to falling income and profits, and the big bank's lending relieves pressure in financial markets (Minsky, 1986). A financial crisis can even occur without setting off a

recession – a good example was the 1987 stock market crash, in which the Fed quickly intervened with the promise that it would lend reserves to market participants to stop necessitous selling of stocks to cover positions. No recession followed the crash – unlike the October 1929 crash, in which margin calls forced sales of stocks. And the big government deficits kept profits flowing in 1987, again unlike 1929 when the government's budget was far too small to make up for collapsing investment.

Unfortunately, most Fed policy over the post-war period involved reducing regulation and supervision, promoting the natural transition to financial fragility. From Minsky's perspective, this was a dangerous combination. While the big bank and the big government reduced the fall-out of crisis, the move to 'self-regulation' by financial institutions and markets made riskier behaviour possible. As the fear of failure was attenuated by a government safety net, perceived risk was lowered. Chairman Ben Bernanke (2004) proclaimed the onset of 'the great moderation' – a new era of stability. As Minsky argued, though, 'stability is destabilizing'. In his view, if the government is going to provide a safety net to prop up and 'validate' risky behaviour, then the other side of the coin must be *greater* oversight and regulation, not less. With rapid financial innovation, reduced regulatory oversight, and less fear of a debt deflation process, financial fragility would build until a collapse.

## Money manager capitalism and the crisis

Beginning in 2007, the world faced the worst economic crisis since the 1930s. References to Keynesian theory and policy became commonplace, with only truly committed free marketeers arguing against massive government spending to cushion the collapse and re-regulation to prevent future crises. All sorts of explanations were proffered for the causes of the crisis: lax regulation and oversight, rising inequality that encouraged households to borrow to support spending, greed and irrational exuberance, and excessive global liquidity – spurred by easy money policy in the USA and by US current account deficits that flooded the world with too many dollars. While each of these explanations does capture some aspect of the crisis, none of them fully recognizes the systemic nature of the global crisis.

Unfortunately, Minsky died in 1996, but after the crash, his work enjoyed unprecedented interest, with many calling this the 'Minsky Moment' or 'Minsky Crisis'. (Cassidy, 2008; Chancellor, 2007; McCulley, 2007; Whalen 2007) I argued above that we should not view this as a 'moment' that can be traced to recent developments. Rather, as Minsky had been arguing for nearly fifty years, what we have seen is a slow transformation of the global financial system toward what Minsky called 'money manager capitalism' that finally collapsed in 2007. Hence I call it the 'Minsky half-century' (Wray, 2009).

It is essential to recognize that we have had a long series of crises in the USA and abroad, and the trend has been toward more severe and more frequent crises: muni bonds in the mid-1960s; real estate investment trusts in the early 1970s; developing country debt in the early 1980s; commercial real estate, junk bonds and the thrift crisis in the USA (with banking crises in many other nations) in the 1980s; stock market crashes in 1987 and again in 2000 with the dot-com bust; the Japanese meltdown from

the early 1980s; Long Term Capital Management, the Russian default and Asian debt crises in the late 1990s; and so on. Until the current crisis, each of these was resolved (some more painfully than others – impacts were particularly severe and long-lasting in the developing world) with some combination of central bank or international institution (IMF, World Bank) intervention plus a fiscal rescue (often taking the form of US Treasury spending of last resort to prop up the US economy to maintain imports that helped to restore rest of world growth).

According to Minsky, the problem is money manager capitalism – the economic system characterized by highly leveraged funds seeking maximum returns in an environment that systematically under-prices risk (Wray, 2009). There are a number of reasons for this. For example, there was the belief in the Greenspan 'put' (the Chairman would always intervene to bail out financial markets if problems developed) and the Bernanke 'great moderation' – both of which lowered perceived risk. Since the last depression and debt deflation had occurred so long ago, few market participants had any memory of it; indeed, many of those in markets did not even remember the savings and loan crisis of the 1980s! Many of the models that were used to price assets were based on a very short time horizon (five years or less; sometimes this was necessitated by the fact that the financial instruments did not exist previous to that), a period that was unusually quiescent. Further, the rise of 'shadow banks' (financial institutions that often had lower costs and less regulation) led to a competitive reduction of risk spreads (pushing interest rates on riskier assets down relative to those on safe assets). Credit ratings agencies played an important role, providing high ratings to assets that proved to be very much riskier than indicated. All of this was made worse by a general 'euphoric' belief that prices of assets (such as real estate and commodities) could only go up. Finally, there was an explosion of various types of derivatives that appeared to reduce risk by shifting it to institutions better able to absorb losses. Perhaps the best example was the use of credit default swaps that were used as insurance in case of default; but when the crisis began, it turned out that all the risk came back in the form of counterparty risk (AIG, the seller of the 'insurance', could not cover the losses). While we cannot go into all the details here, it was even worse than that because credit default swaps were also used as pure bets on failure (the bettor would win if the assets went bad), and prices of these instruments were used as indicators of the probability of default (rising credit default swap prices could induce credit raters to lower ratings, which then triggered pay-offs on the bets even as they raised borrowing costs for the debtors) (see Wray, 2009).

In sum, contrary to efficient markets theory, markets generate perverse incentives for excess risk, punishing the timid with low returns (Cassidy, 2009). Any money manager who tried to swim against the stream by avoiding excessive leverage and complex and hard-to-value assets found it hard to retain clients. Those playing along were rewarded with high returns because highly leveraged funding drives up prices for the underlying assets – whether they are dot-com stocks, Las Vegas homes, or corn futures. It all works – until it doesn't. We now know from internal emails that many financial market participants knew that risk was under-priced, but adopted an 'I'll be gone, you'll be gone' strategy – take the risk, get the millions of dollars in compensation now, and retire when the whole thing collapses.

Many have accurately described the phenomenon as 'financialization' – growing debt that leverages income flows and wealth. At the 2007 peak, total debt in the US reached a record 5 times GDP (versus 3 times GDP in 1929), with most of that private debt of households and firms. From 1996 until 2007 the US private sector spent more than its income (running deficits that increased debt) every year except during the recession that followed the dot-com bust in 2000. Financial institution debt also grew spectacularly over the two decades preceding the crisis, totaling more than GDP. Exotic financial instruments exploded – outstanding credit default swaps (bets on default by households, firms, and even countries) reached over $60 trillion, and total financial derivatives (including interest rate swaps, and exchange rate swaps) reached perhaps $600 *trillion* – many times world GDP.

Some accounts blame subprime mortgages (home loans made to riskier borrowers, typically low income households) for the global financial collapse – but that is too simple. The total value of riskier mortgage loans made in the USA during the real estate boom could not have totalled more than a trillion or two dollars (big numbers but small relative to the total volume of financial instruments). The USA was not the only country that experienced a speculative boom in real estate – Ireland, Spain and some countries in eastern Europe also had them. Then there was also speculation in commodities markets –leading to the biggest boom in history, followed by the inevitable crash – that involved about a half trillion dollars of managed money (mostly US pension funds) placing bets in commodities futures markets (Wray, 2008). Global stock markets also enjoyed a renewed speculative hysteria. Big banks like Goldman Sachs speculated against US state governments, as well as countries like Greece. (For example, Goldman Sachs encouraged clients to bet against the debt issued by at least 11 US states – while collecting fees from those states for helping them to place debt. A common technique was to pool risky debt into securities, sell these to investors, then 'short' the securities using credit default swaps to bet on failure. The demand for CDSs for shorting purposes would lead to credit downgrades that raised finance costs and hastened default. The most famous shorter of mortgage debt is John Paulson, whose hedge fund asked Goldman Sachs to create toxic synthetic collateralized debt obligations (CDOs) that it could bet against. According to the US Securities and Exchange Commission, Goldman allowed Paulson's firm to increase the probability of success by picking particularly risky MBSs to include in the CDOs. Goldman arranged a total of 25 such deals, named Abacus, totaling about $11 billion. Out of 500 CDOs analyzed by UBS, only two did worse than Goldman's Abacus. Just how toxic were these CDOs? Only five months after creating one of these Abacus CDOs, the ratings of 84% of the underlying mortgages had been downgraded. By betting against them, Goldman and Paulson won – Paulson pocketed $1 billion on the Abacus deals (he made a total of $5.7 billion shorting mortgage-based instruments in a span of two years) and Goldman earned fees for arranging the deals. According to the SEC Goldman's customers actually met with Paulson as the deals were assembled – but Goldman never informed them that Paulson was the shorter of the CDOs they were buying!)

On top of all this speculative fervor there was also fraud – which appears to have become normal business practice in all of the big financial institutions. It will be years, perhaps decades, before we will unravel all of the contributing factors, including the

financial instruments and practices as well as the questionable activities by market players and government officials that led to the collapse. (The *Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States* (commissioned by the US Congress and President Obama) concluded that the crisis was both foreseeable and preventable. It blamed the 'captains of finance' (heads of the biggest banks) and the 'public stewards' (officials charged with regulating the banks) for the systemic breakdown in accountability and ethics that led to the crisis. Former bank regulator William Black (who blew the whistle on Charles Keating, the convicted felon who ran Lincoln Savings, the biggest thrift to fail as a result of the 1980s crisis, and the patron of five US Senators known as the 'Keating Five') is more blunt: the biggest banks in America were run as 'control frauds' designed to enrich top management while defrauding customers and shareholders. By his reckoning, thousands of individuals committed go-to-jail fraud. Only time will tell whether they will be brought to justice.)

This much we do know: the entire financial system had evolved in a manner that made 'it' – an economic collapse and debt deflation – possible. Riskier practices had been permitted by regulators, and encouraged by rewards and incentives. Lack of oversight and prosecution led to a dramatic failure of corporate governance and risk management at most big institutions (see the *Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*). The combination of big government and big bank interventions plus bail-outs of 'too big to fail' institutions in crisis after crisis since the 1960s let risk grow on trend. The absence of depressions allowed financial wealth to grow over the entire post-war period – including personal savings and pension funds. All of these funds needed to earn returns. As a result, the financial sector grew relative to GDP – as a percentage of value added, it grew from 10% to 20%, and its share of corporate profits quadrupled from about 10% to 40% from 1960 to 2007 (Nersisyan and Wray, 2010). It simply became too large relative to the size of the economy's production and income. The crash was the market's attempt to downsize finance – just as the crash in 1929 permanently reduced the role played by finance, and allowed for the robust growth of the post-war period. Beginning in summer 2007, a series of runs on financial institutions began that would have snowballed without unprecedented intervention by governments around the world. Typically these took the form of a refusal by markets to 'refinance' banks. Recall from above that debt of financial institutions had grown tremendously, as they borrowed mostly short-term to finance positions in financial assets. Often this took the form of overnight borrowing plus very short-term commercial paper on the basis of high-quality collateral. As the crisis unfolded, borrowers had to pledge more and more collateral, and pay higher and higher interest rates to borrow. By fall of 2007, the 'haircut' (a 10% haircut means the bank can borrow 90 cents against each dollar of good collateral) was so large that many financial institutions could no longer borrow enough to finance their positions in assets – meaning they had to sell assets into a market that now feared risk. Such 'fire sales' would lead to what Irving Fisher and Minsky called a 'debt deflation'. At the same time, worried shareholders began to dump bank stocks. Without prompt rescue by governments, the 'market' would have operated in a manner that would have led to failure of most institutions. US Treasury Secretary Timothy

Geithner later said that 'none of [the biggest banks] would have survived a situation in which we had let that fire try to burn itself out' and Fed Chairman Ben Bernanke said 'As a scholar of the Great Depression, I honestly believe that September and October of 2008 was the worst financial crisis in global history... out of maybe the 13, 13 of the most important financial institutions in the United States, 12 were at risk of failure within a period of a week or two' (*Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*, p. 354).

It is important to include as contributing factors the erosion of New Deal institutions that had enhanced economic stability, including most importantly the creation of a high-consumption, high-employment and high-wage society. As Minsky (1986, 1996) argued, the USA emerged from the Second World War with powerful labour unions that were able to obtain good and growing wages, which fueled growth of domestic consumption out of income. According to Minsky, debt loads were extremely low in the private sector – with debts having been paid down or wiped out by bankruptcy in the Great Depression – and with lots of safe government bonds held as assets. In combination with a strengthened government safety net (Social Security for the aged, welfare and unemployment compensation for those without jobs, the GI bill for soldiers returning home, low interest rate loans for students) this meant that consumption comprised a relatively larger part of GDP. For Minsky, consumption out of income is a very stable component – unlike investment, which is unstable. Minsky argued that investment-led growth is more unstable than growth led by a combination of consumption out of income plus government spending because the second model does not lead to worsening private sector balance sheets.

However, over the course of the past four decades, union power declined. Minsky frequently claimed that the most significant action taken during the Reagan administration was the busting of the air traffic controllers' union (which, he claimed, sent a message to all of labour). Median real wages stopped growing, consumer debt grew on trend (and then exploded after 1995), and the generosity of the safety net was reduced. Further, over the whole period, policy increasingly favoured investment and saving over consumption – with favourable tax treatment of savings and investment, and with public subsidies of business investment. Federal government also stopped growing (relative to the size of the economy) and its spending shifted away from public infrastructure investment. Inequality grew on trend, so that it actually surpassed the 1929 record inequality. President Bush even celebrated the creation of the 'ownership society' – ironically, with concentration of ownership of financial assets at the very top (Wray, 2005). The only asset that was widely owned was the home, which then became the basis for a speculative real estate bubble that produced financial assets traded around the world. The global financial collapse and deep recession in the USA after 2007 then generated widespread foreclosures (13 million by 2012) – with families kicked out of their homes, owing lots of debt, and with real estate prices collapsing so that vulture hedge funds could buy up blocks of houses at pennies on the dollar. By 2010 the home ownership rate in the USA had returned to the pre-boom level.

The 1929 crash ended what Minsky and Rudolf Hilferding designated the finance capitalism stage (Wray, 2009) Perhaps the global financial crisis of 2007 will prove to be the end of this stage of capitalism – the money manager phase. Of course, it is too

early to even speculate on the form capitalism will take in the future. In the final section I will look at the policy response that could help to reformulate global capitalism along Minskian lines.

## Minskian policy in the aftermath of the collapse of money manager capitalism

Minsky (1986) argued that the Great Depression represented a failure of the small-government, *laissez faire* economic model, while the New Deal promoted a Big Government/Big Bank highly successful model for financial capitalism. Following Minsky, we might say that the current crisis represents a failure of the Big Government/Neoconservative (or, outside the USA, what is called neo-liberal) model that promotes deregulation, reduced supervision and oversight, privatization, and consolidation of market power. It replaced the New Deal reforms with self-supervision of markets, with greater reliance on 'personal responsibility' as safety nets were reduced, and with monetary and fiscal policy that is biased against maintenance of full employment and adequate growth to generate rising living standards for most Americans. Even before the crisis, the USA faced record inequality, a healthcare crisis, and high rates of incarceration, among other problems facing the lower and middle classes (Wray 2000, 2005). All of these trends are important as they increase insecurity and the potential for instability, as Minsky described in one of his last published pieces (Minsky 1996).

We must return to a more sensible model, with enhanced oversight of financial institutions and with a financial structure that promotes stability rather than speculation. We need policy that promotes rising wages for the bottom half so that borrowing is less necessary to achieve middle class living standards. We need policy that promotes employment, rather than transfer payments – or worse, incarceration – for those left behind. Monetary policy must be turned away from using rate hikes to pre-empt inflation and toward a proper role: stabilizing interest rates, direct credit controls on bank lending to prevent runaway speculation, and stronger bank supervision. (A central bank could, for example, increase margin requirements on lending to speculators, raise required down payments for bank real estate lending, and set limits on bank lending for specified purposes in a euphoric boom.)

Minsky insisted that 'the creation of new economic institutions which constrain the impact of uncertainty is necessary', arguing that the 'aim of policy is to assure that the economic prerequisites for sustaining the civil and civilized standards of an open liberal society exist. If amplified uncertainty and extremes in income maldistribution and social inequalities attenuate the economic underpinnings of democracy, then the market behavior that creates these conditions has to be constrained' (Minsky, 1996, pp. 14, 15). It is time to take finance back from the clutches of Wall Street's casino.

Minsky had long called for an 'employer of last resort' program to provide jobs to those unable to find them in the private sector. In a sense this would be a counterpart to the central bank's 'lender of last resort' program. In the jobs program, government would offer a perfectly elastic supply of jobs at a basic program wage. Anyone willing to work at that wage would be guaranteed a job. Workers would be 'taken as they are' – whatever their level of education or training – and jobs would be designed for their

skill level. Training would be a part of every job – to improve skills and to make workers more employable outside the program. The work would provide useful services and public infrastructure, improving living standards. While Minsky is best known for his work on financial instability, his proposal for the employer of last resort program received almost as much of his attention, especially in the 1960s and 1970s. Interested readers are referred to the growing body of work on use of job guarantee programs as part of long-term development strategy (Bhaduri, 2005; Felipe *et al.*, 2009; Hirway, 2006; Minsky, 1965; Mitchell and Wray, 2005; Tcherneva and Wray, 2007; Wray, 2007). Note that this would help to achieve Minsky's goal of a high-employment economy with decent wages to finance consumption. Minsky always saw the job guarantee as a stabilizing force – and not something that is desirable only for humanitarian reasons.

The global crisis offers both grave risks as well as opportunities. Global employment and output collapsed faster than at any time since the Great Depression. Hunger and violence grew after the financial crisis – even in developed nations. The 1930s offer examples of possible responses – on the one hand, nationalism and repression (Nazi Germany), on the other a New Deal and progressive policy. From a Minskian perspective, finance played an outsized role in the run-up to the crisis, both in the developed nations, where policy promoted managed money, and in the developing nations, which were encouraged to open to international capital. Households and firms in developed nations were buried under mountains of debt even as incomes for wage earners stagnated. Developing nations were similarly swamped with external debt service commitments, while the promised benefits of Neoliberal policies often never arrived.

Minsky would probably argue that it is time to put global finance back in its proper place as a tool to achieving sustainable development, much as the USA did in the aftermath of the Great Depression. This means substantial downsizing and careful re-regulation. Government must play a bigger role, which in turn requires a new economic paradigm that recognizes the possibility of simultaneously achieving social justice, full employment, and price and currency stability through appropriate policy.

L. RANDALL WRAY

*See Also* **banking crisis; credit crunch chronology; European Central Bank and monetary policy in the Euro area; euro zone crisis 2010.**

### *Bibliography*

Bernanke, B. S. 2004. *The Great Moderation*. Speech given at the meetings of the Eastern Economics Association, Washington DC, 20 February. Available at http://www.federalreserve.gov/Boarddocs/Speeches/2004/20040220/default/htm. Accessed 12 May 2009.

Bhaduri, A. 2005. *Development With Dignity: a Case for Full Employment*. National Book Trust, India.

Cassidy, J. 2008. The Minsky moment. *The New Yorker*, 4 February. http://www.newyorker.com/. Accessed 29 January 2008.

Cassidy, J. 2009. *How Markets Fail: the Logic of Economic Calamities*. Picador, New York.

Chancellor, E. 2007. Ponzi Nation. *Institutional Investor*, 7 February.

Felipe, J., Mitchell, W. and Wray, L. R. 2009. A reinterpretation of Pakistan's 'economic crisis' and options for policymakers. Manuscript, Asian Development Bank.

Hirway, I. 2006. Enhancing livelihood security through the National Employment Guarantee Act: toward effective implementation of the Act. *The Levy Economics Institute Working Paper No. 437*. http://www.levy.org/.

McCulley, P. 2007. The plankton theory meets Minsky. *Global Central Bank Focus*, March. PIMCO Bonds: http://www.pimco.com/LeftNav/Featured + Market + Commentary/FF/1999-2001/FF_01_2001.htm. Accessed 8 March 2007.

Minsky, H. P. 1957. Central banking and money market changes. *Quarterly Journal of Economics*, 71(2), 171.

Minsky, H. P. 1965. The role of employment policy. In: *Poverty in America* (ed. M. S. Gordon). Chandler Publishing Company, San Francisco, CA.

Minsky, H. P. 1975. *John Maynard Keynes*. Columbia University Press, New York.

Minsky, H. P. 1982. *Can it Happen Again?* M. E. Sharpe, Armonk, NY.

Minsky, H. P. 1986. *Stabilizing an Unstable Economy*. Yale University Press, New Haven and London.

Minsky, H. P. 1996. *Uncertainty and the Institutional Structure of Capitalist Economies*. The Levy Economics Institute of Bard College, Working Paper No. 155.

Minsky, H. P. 2008 (1987). *Securitization*. Levy Economics Institute of Bard College, Policy Note No. 2, 12 May.

Mitchell, W. F. and Wray, L. R.2005. In defense of employer of last resort: a response to Malcolm Sawyer. *Journal of Economic Issues*, 39(1), 235–245.

Nersisyan, Y. and Wray, L. R. 2010. Transformation of the financial system: financialization, concentration, and the shift to shadow banking. In: *Minsky, Crisis and Development* (eds. D. Tavasci and J. Toporowski), pp. 32–49. Palgrave Macmillan, Basingstoke.

Papadimitriou, D. B. and Wray, L. R. 1998. The economic contributions of Hyman Minsky: varieties of capitalism and institutional reform. *Review of Political Economy*, 10(2), 199–225.

Tcherneva, P. R. and Wray, L. R. 2007. Public employment and women: the impact of Argentina's Jefes Program on female heads of poor households. Levy Economics Institute *Working Paper* No. 519. http://www.levyinstitute.org/publications/?docid = 965.

Whalen, C. 2007. The U.S. credit crunch of 2007: a Minsky moment. Levy Economics Institute *Public Policy Brief*, No. 92. http://www.levy.org/.

Wray, L. R. 2000. A new economic reality: penal Keynesianism. *Challenge*, September–October, 31–59.

Wray, L. R. 2005. The ownership society: social security is only the beginning. Levy Economics Institute *Public Policy Brief*, No. 82. http://www.levy.org/.

Wray, L. R. 2007. The employer of last resort programme: could it work for developing countries? *Economic and Labour Market Papers*, 2007/5, International Labour Office, Geneva.

Wray, L. R. 2008. The commodities market bubble: money manager capitalism and the financialization of commodities. Levy Economics Institute *Public Policy Brief*, No. 96. http://www.levy.org/.

Wray, L. R. 2009. The rise and fall of money manager capitalism: a Minskian approach. *Cambridge Journal of Economics*, 33(4), 807–828.

## New Deal

Franklin Roosevelt's New Deal created the most dramatic peacetime expansion of government in American economic history.

When Franklin D. Roosevelt became president in March 1933, real output had fallen 30 per cent from its 1929 peak and the unemployment rate exceeded 25 per cent. Within his first hundred days in office Roosevelt and the Democratic Congress established an incredible array of programmes, a virtual 'alphabet soup' of acronyms. More programmes were added under the First New Deal until 1935, when the Supreme Court declared the National Recovery Administration's (NRA) codes of 'fair' competition for industry and the Agricultural Adjustment Administration (AAA) farm programme unconstitutional. A Second New Deal re-established the farm programme in the name of soil conservation, strengthened the role of unions in collective bargaining, and established the basic structure of most of America's current social insurance and public assistance programmes.

After Roosevelt took office, the federal government, often in conjunction with state and local governments, built a huge number of roads, dams, sanitation facilities, schools, public housing projects, and other public works. The federal government expanded regulation of banking, finance, labour, and a host of other markets, insured and refinanced housing loans, and made extensive loans to numerous private and public entities. In the decades following the 1930s, several waves of historians have provided narratives and interpretations of the New Deal and introductions to their work can be found in collections edited by Dubofksy (1992), Braeman, Bremner and Brody (1975), and Hamby (1969). The recent trends in New Deal studies include a series of studies by economists and economic historians (Fishback et al., 2007; Bordo, Goldin and White, 1998).

Searching for an overarching theme for the programmes is a daunting task. The doubling of annual federal spending between the Hoover (1929–32) and Roosevelt years tempts many to describe the New Deal as Keynesian expansionary policy. But the Roosevelt administration ran relatively small budget deficits, as federal tax collections also more than doubled. In a brief meeting and a letter to the *New York Times* Keynes had encouraged Roosevelt to follow an expansionary policy, but the levels of government spending and the small budget deficits pale in comparison with the fall in output to be counteracted (Barber, 1996; Brown, 1956; Peppers, 1973; Romer, 1992).

One goal appeared to have been to raise prices and wages, as the establishment of the NRA allowed each industry to establish cartel-like codes that stifled price and quality competition, labour policies promoted unionization and high wages, and farm policies offered price guarantees while cutting output. Ultimately, Roosevelt and his advisors were pragmatists faced with terrible economic problems of nearly every kind. They established agencies and programmes meant to try to solve nearly each and every one. At times the programmes operated at cross-purposes. Higher farm and industry prices worsened the plight of the unemployed and other consumers. The pressure to

raise wages exacerbated the unemployment problem, and the NRA codes limited output growth. The administration made constant adjustments in policies, creating a climate of uncertainty about the regulatory environment that left businesses wary of making new investments (Higgs, 1997).

## New Deal monetary, banking, and international policy

Building on the seminal work by Friedman and Schwartz (1963), many economists argue that monetary policy contributed significantly to the harsh decline in the economy between 1929 and 1933. The Federal Reserve took seriously its international responsibilities in maintaining the gold standard and thus failed to respond sufficiently to three major waves of bank failures in a timely fashion. Many states had begun declaring 'holidays' that closed state banks to stave off bank runs. Roosevelt took office in the midst of the third wave of failures and declared a Bank Holiday that closed all national banks. Two-thirds of the banks were declared sound and reopened within the week. The troubled banks were reorganized and the Reconstruction Finance Corporation (RFC) subscribed to their new stock issues, reassuring the public about the solvency of the banking system (Smiley, 2002; Mason, 2001).

In 1933 Roosevelt also announced that the United States was leaving the gold standard, prohibited gold exports, and devalued the dollar to $35 per ounce of gold. In response, the United States received a substantial flow of gold that stimulated the money supply, and economic growth resumed. Japan, Britain, France and several other leading nations experienced similar resumptions of economic growth when they broke free of their 'golden fetters' (Eichengreen, 1992; Temin, 1989; Temin and Wigmore, 1990). Gold inflows continued for the rest of the 1930s as Europe moved towards war. By choosing not to offset the gold inflows, Roosevelt and the Federal Reserve allowed the money supply to expand (Romer, 1992). The Federal Reserve took a misstep, however, when it used its newly awarded control over reserve requirements to double them in three steps between 1935 and 1937. The goal was to prevent a potentially inflationary rise in lending by soaking up the substantial excess reserves that banks were holding at the time. The banks responded by increasing their reserves and keeping the same cushion because they did not trust the Federal Reserve to provide adequate liquidity if a bank run occurred. The money supply fell and contributed to a sharp rise in unemployment and drop in real GDP in 1937–8 (Friedman and Schwartz, 1963; Romer, 1992). There is some disagreement about the impact of the monetary policies. Real business cycle economists argue that monetary and investment changes played much smaller roles than productivity shocks and high-wage labour policies in accounting for the fluctuations during the 1930s (Chari, Kehoe and McGratton, 2005).

The decision to leave the gold standard was accompanied by efforts to expand world trade beginning in 1934 with the Reciprocal Trade Agreement Act (RTA). The Smoot–Hawley Tariff Act of 1930 had helped touch off a series of protectionist responses by other countries that had caused total imports for a group of 75 countries to fall to one-third of their 1929 level. The RTA freed the Roosevelt administration to sign a series of tariff reduction agreements with Canada, several South American countries, Britain and

key European trading partners. Consequently, American imports rose from a 20-year low in 1932–3 to an all-time high by 1940 (Irwin, 1998; Kindleberger, 1986).

Meanwhile, the Banking (Glass–Steagall) Act of June 1933 enacted an additional set of banking policies. Despite the checkered history experienced by state deposit insurance programmes (Calomiris and White, 2000), the act created the Federal Deposit Insurance Corporation (FDIC) to insure commercial bank deposits of up to $10,000. Insurance for savings and loans followed within the year. The Banking Act also established regulations, eliminated in the late 1970s, that prevented commercial banks from investing more than ten per cent of their assets in stocks and paying interest on deposits (Regulation Q). To increase the capital available for housing loans, the Home Owners' Loan Corporation (HOLC) provided funds to refinance troubled mortgages between 1933 and 1936, and the Federal Housing Administration (FHA) began offering insurance of mortgages and home improvement loans. Both agencies aided in the spread of the modern long-term, amortized mortgage loan that replaced short-term loans in which repayment of only interest over the course of the loan was followed by a balloon payment of the principal when it fell due.

## The Reconstruction Finance Corporation (RFC): New Deal lender

Established by President Herbert Hoover in 1932, the RFC was an off-budget government corporation that maintained control of the funds repaid on its earlier loans. The RFC offered the Roosevelt administration flexibility because they could start funding programmes without constantly seeking new appropriations from Congress. In consequence, the RFC became the lender during the starting phase of nearly every major New Deal grant and lending programme. In addition, the RFC provided loans to large numbers of financial institutions of all types, railroads, farmers and local governments (Olson, 1998). The RFC loans to private business met with mixed success. The liquidity loans to failing banks in 1932 had not prevented many bankruptcies because the RFC loans were given first priority over depositors and other lenders in case of failure; therefore, banks were prevented from selling their most liquid assets to meet depositor demands for cash. The RFC's purchases of preferred stock in banks reorganized after the Bank Holiday of 1933 exposed the RFC funds to more risk but led to more success at preventing failures (Mason, 2001). RFC lending to railroads succeeded in preventing several railroad bankruptcies. However, the spared railroads continued to underinvest in maintenance and capital improvements. In contrast, railroads forced into bankruptcy had to make such investments to attract enough capital to reopen for business (Mason and Schiffman, 2004).

## Emergency relief and public works programmes

Unprecedented unemployment rates ranging from 10 to 25 per cent through the 1930s were the New Deal's greatest challenge. Prior to the New Deal, aid to the poor and labour policies had been the purview of state and local governments. Claiming unemployment to be a national emergency, Roosevelt and Congress raised the federal share of relief spending as high as 79 per cent while nearly quadrupling relief spending even as unemployment rates fell by the mid-1930s. The Federal Emergency Relief

Administration (FERA, 1933–5), the Civil Works Administration (CWA, winter of 1933–4), and the Works Progress Administration (WPA, 1935–42) offered work relief jobs to households whose incomes fell below a target budget for necessities. The Civilian Conservation Corps (CCC) offered conservation jobs in the nation's hinterlands to youths whose earnings were shared with their parents. The FERA also handed out direct relief until 1935, when the responsibility for 'unemployables' was returned to state and local governments, and the federal government began offering matching grants for public assistance for children, the blind, and the elderly.

Harry Hopkins, who headed the FERA, CWA and the WPA, preferred work relief because it 'provided a man with something to do, put money in his pocket, and kept his self-respect' (Adams, 1977, p. 53). To give people incentive to leave work relief for private jobs, WPA monthly earnings averaged 40 to 50 per cent of full-time private earnings, and the WPA assured people that they would be reaccepted should the private job end. Even so, a significant percentage of workers stayed on work relief jobs for periods as long as a year and in some cases several years (Margo, 1993).

Roughly one-fourth of New Deal grant spending went to the Public Works Administration (PWA), Public Buildings Administration (PBA), the Public Roads Administration (PRA), and the Tennessee Valley Authority (TVA). The planning stages on these large-scale projects were longer, the wages were higher, and there was more freedom to hire already employed workers. The relief and public works programmes grants were designed to provide employment, build public projects, and stimulate the economy.

At one level the relief and public works programmes were very successful. Millions of Americans obtained work relief jobs to tide them over, and most of the original public works, many renovated since, are still in place today. To understand the true impact of the New Deal, areas with different amounts of spending need to be compared to get a sense of how their economies would have performed without the New Deal. Since the mid-1990s economists have been using the substantial variation in spending across local areas to make such comparisons while working to control for the feedbacks caused by administrators using New Deal programmes to respond to economic problems. At the local level the benefits of the projects were likely to be stronger when the general share of goods produced in the area for local consumption was higher, the projects hired the unemployed without crowding out private or state and local government employment, and expansions did not raise incomes enough to generate federal income tax payments.

Although cross-sectional studies show little effect of relief jobs on private employment, analysis of panel data can control for unmeasured factors using the information across time for a cross section of areas. The panel studies suggest that an additional relief job reduced private employment by up to half a job (Wallis and Benjamin, 1981; 1989; Fleck, 1999a). A new relief job also raised 'measured' unemployment by one person because many discouraged workers, who had been out of the labour force and thus not counted as unemployed, were defined as re-entering the labour force as unemployed workers when they accepted relief jobs (Darby, 1976; Fleck 1999a).

The impact of public works and relief programmes had more clearly beneficial effects on other measures of socio-economic welfare. Cross-sectional studies of US

counties suggest that an added dollar of public works and relief spending per person raised per capita income by roughly 85 cents and stimulated in-migration (Fishback, Horrace and Kantor, 2005; 2006). Panel studies of more than 100 major cities between 1929 and 1940 show that increased relief spending stimulated birth rates, reduced property crime, and reduced infant deaths and deaths from suicide and several diseases. The relief costs per death prevented in today's dollars are within the range of modern market values of life, and the costs are lower than the costs per death prevented of many modern safety programmes (Fishback, Haines and Kantor, 2007; Johnson, Kantor and Fishback, 2006).

## Farm programmes

To raise the incomes of farmers, who had struggled through over a decade of hard times, the New Deal established the structure of the modern US farm programmes. The Agricultural Adjustment Administration (AAA) paid farmers to take land out of production. In 1935 in *United States v. Butler* the Supreme Court struck down the output processing tax that had originally funded the payments. The AAA payments were quickly reinstituted (minus the processing tax) under the Soil Conservation and Domestic Allotment Act (1935). The Commodity Credit Corporation (CCC) insured that farmers were paid higher prices by making loans that could be repaid with the crop itself if market prices fell below a target price. The Farm Credit Administration (FCA) reorganized and expanded farm lending, ultimately becoming involved in more than half of all farm mortgages and a large share of production loans. Meanwhile, the Rural Electrification Administration (REA) provided subsidized loans to give farmers access to electricity, while the Farm Security Administration (FSA) developed programmes to aid low-income farmers.

Efforts to determine the AAA's impact on limiting farm output have been confounded because a series of major climatic disasters in the 1930s served to cut output anyway. There is evidence that farmers stopped planting their least productive land and raised the inputs used on the remaining land. The AAA clearly aided large farmers but possibly at the expense of farm workers and tenants (Alston and Ferrie, 1999; Whatley, 1983). Cross-county studies show that increases in AAA payments in counties led to no increases in retail sales, were associated with higher infant mortality in the South, and stimulated net outmigration (Fishback, Horrace and Kantor, 2005; 2006; Fishback, Haines and Kantor, 2001; Alston and Ferrie, 1999; Whatley, 1983). On the positive side, the AAA soil conservation programmes encouraged a move to larger farms and practices that cut soil erosion, so that the Great Plains avoided a recurrence of the Dust Bowl when the same drought and wind conditions arose later (Hansen and Libecap, 2004).

## The political economic geography of New Deal spending

New Deal grant spending across states and counties varied enormously, as some western states received several times more per head than some southern states. Roosevelt in a radio 'fireside chat' vowed that the New Deal would promote 'Relief, Recovery, and Reform'. Critics argued that Roosevelt used the monies primarily to aid his re-election

efforts. The distribution process for many programmes was opaque, so New Deal scholars have turned to econometric analysis that simultaneously tests the importance of the stated motives and presidential politicking. Politicking was clearly part of the process in the distribution of total funds and at the programme level. Nearly every study finds that more grants went to swing states and areas with higher political turnout, while some find rewards for loyal Democratic areas as well as districts represented by powerful congressmen. The Roosevelt administration was innovative in targeting radio owners in their push to win elections (Wright, 1974; Wallis, 1998; Fleck, 1999b; Stromberg, 2004; Couch and Shughart, 1998).

Winning elections required more than just manipulation of spending to hit specific political targets. The Roosevelt administration also enhanced its future re-election prospects by following its stated aims. Many studies find evidence that the Roosevelt administration promoted recovery and relief by spending more in areas with higher unemployment and larger declines in income from 1929 to 1933. Few find signs that the total spending was reform-oriented, but specific relief programmes did target areas with long-term poverty. State governments influenced the distribution by the intensity of their lobbying and their spending in matching grant programmes, while the presence of federal land in a state also drew substantial public works grants. Specific programmes typically followed stated goals. There were so many programmes that nearly everybody could find one that benefited them, ranging from relief for the unemployed and poor to loans and AAA grants for large farmers. The HOLC and FHA housing programmes benefited carefully vetted home owners who were perceived as having lower risk of default (Fishback, Wallis and Kantor, 2003). There were constant charges of corruption, but the WPA actively battled corruption at the state and local levels by establishing an internal investigative agency. When the federal government increased its control of the distribution of funds within states in the switch from the FERA to the WPA, the distribution of funds within states more closely mirrored the relief, recovery and reform goals (Wallis, Fishback and Kantor, 2006).

## Industrial and labour policies

To combat 'destructive competition', low prices and low wages, the National Recovery Administration (NRA) was created to allow industries to establish their own codes for minimum prices, quality standards, trade practices, and labour relations (Bellush, 1975). The NRA appeared to be sponsoring a series of industry cartels, as large firms tended to dominate the code-writing process in most industries. Wholesale prices jumped 23 per cent in two years, although consumer prices were much slower to rise. Simulations of the economy with and without the NRA imply that it served to slow economic recovery (Cole and Ohanian, 2004). The internal problems of cartels were also present, as industries with diverse firms had trouble coming to agreement and a number of firms routinely violated the codes (Alexander and Libecap, 2000). The NRA ended in 1935 when the Supreme Court declared it unconstitutional in the Schechter Poultry case, and few mourned its passing.

The National Labor Relations (Wagner) Act of 1935 expanded the right of workers to collective bargaining through their own representatives beyond the protections

originally offered in the 1933 act that created the NRA. Employers were required to bargain with unions when a majority of workers voted for union representation, and employer-sponsored unions were banned. The National Labor Relations Board (NLRB) was established to oversee union elections and the collective bargaining process. As a result, unionization expanded rapidly through a mixture of strikes and elections. In the long run the NLRB policies regularized the union recognition and bargaining process, and the incidence of violent strikes has diminished sharply since (Freeman, 1998).

The emphasis on raising wages continued when the Fair Labor Standards Act (FSLA) of 1938 set a national minimum wage, overtime requirements, and child labour restrictions. Workers in agriculture or not employed in interstate commerce were exempted. Congressional support for the act was centred in states outside the South with high-wage industries, more unionization, and more advocates for teenage workers. As a result, the first minimum wage was binding only for low-wage industries in the South, where employers in some southern industries responded by reducing employment, and others switched to labour-saving technologies or limited their business to intra-state commerce to avoid federal regulation (Seltzer, 1995, 1997; Fleck, 2004).

## The Social Security Act of 1935

The legislative centerpiece of the Second New Deal was the Social Security Act (SSA) of 1935, which established the modern structure of public assistance and social insurance programmes. The public assistance grants set some federal guidelines and offered matching grants that gave the states latitude in setting benefits. The new Aid to Dependent Children (ADC), Aid to the Blind (AB), and Old-Age Assistance (OAA) programmes replaced similar state programmes in more than half of the states, and provided coverage for the first time in the remaining states.

State unemployment insurance programmes funded by employer contributions with administrative costs paid by the federal government were established as a long-term alternative to providing emergency work relief. The states retained control over benefits offered. Each designed its own experience-rating system that required employers who laid off more workers to pay higher premiums, a feature not commonly found in other countries' unemployment insurance systems. The experience rating helped reduce seasonal unemployment fluctuations (Baicker, Goldin and Katz, 1998).

Social security is most associated with the federal old-age retirement system. In the debates over social security, Roosevelt pressed for an actuarially sound system where the individual's retirement benefits were based purely on his and his employer's own contributions. He was not convinced the old-age pensions were necessary and sought to ensure that future generations would not be saddled with the costs. Others pressed for a subsidized system that provided adequate payments to all who contributed. The plan adopted in 1935 was a hybrid, but the inadequacies of the hybrid system had become apparent by 1939, and the current pay-as-you-go structure was created. A worker and his employer pay taxes into an administrative trust fund that pays benefits to current retirees and serves as a commitment by the federal government to collect enough taxes to pay the worker his own social security pension when he reaches retirement age. The initial taxes were one per cent of wages each for workers and

employers, and the initial benefits paid in 1940 were roughly 25 per cent of the average earnings of workers contributing to the system. Average pension payments are now roughly 40 per cent of the contributing workers' average earnings, and the increase in average lifespans has caused rapid increases in the ratio of retirees to workers. In consequence, the tax rates had risen to over 5.3 per cent each for worker and employer by 2000, with expectations that relative benefits will have to be cut or taxes raised in the future to sustain the system (Schieber and Shoven, 1999).

## Conclusion

The New Deal was a response to the Great Depression, a major peacetime crisis sandwiched between two world wars. All three crises contributed to short-run rapid expansions of the federal government. When each ended, the government's role retracted somewhat but never to the level that would likely have occurred without the crisis (Higgs, 1987). In the span of six years the Roosevelt administration built an incredible array of public works and established a series of regulations, government insurance, and public assistance programmes that are still in place today. The New Deal arguably did more to expand the role of government in the United States than the more evolutionary changes that have occurred since the end of the Second World War.

PRICE V. FISHBACK

*See Also* **Great Depression.**

### *Bibliography*

Adams, H.H. 1977. *Harry Hopkins: A Biography*. New York: G.P. Putnam Sons.

Alexander, B. and Libecap, G. 2000. The effect of cost heterogeneity in the success and failure of the New Deal's agricultural and industrial programs. *Explorations in Economic History* 37, 370–400.

Alston, L. and Ferrie, J. 1999. *Southern Paternalism and the American Welfare State*. New York: Cambridge University Press.

Baicker, K., Goldin, C. and Katz, L. 1998. A distinctive system: origins and impact of U.S. unemployment compensation. In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, ed. M. Bordo, C. Goldin, and E.N. White. Chicago: Chicago University Press.

Barber, W.J. 1996. *Designs within Disorder: Franklin D. Roosevelt, the Economists, and the Shaping of American Economic Policy, 1933–1945*. New York: Cambridge University Press.

Bellush, B. 1975. *The Failure of the NRA*. New York: Norton.

Bordo, M., Goldin, C. and White, E.N., eds. 1998. *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*. Chicago: University of Chicago Press.

Braeman, J., Bremner, R.H. and Brody, D., eds. 1975. *The New Deal*. Columbus: Ohio State University Press.

Brown, E.C. 1956. Fiscal policy in the 'thirties: a reappraisal. *American Economic Review* 46, 857–79.

Calomiris, C. and White, E.N. 2000. The origins of federal deposit insurance. In *U.S. Bank Deregulation in Historical Perspective*, ed. C. Calomiris. New York: Cambridge University Press.

Chari, V.V., Kehoe, P. and McGrattan, E.R. 2005. Spectral methods in business cycle accounting. *Revista de Economia* 12, 5–18.

Cole, H. and Ohanian, L. 2004. New Deal policies and the persistence of the Great Depression: a general equilibrium analysis. *Journal of Political Economy* 112, 779–816.

Couch, J. and Shughart, W., III 1998. *The Political Economy of the New Deal*. New York: Edward Elgar.

Darby, M.R. 1976. Three and a half million US employees have been mislaid: or, an explanation of unemployment, 1934–1941. *Journal of Political Economy* 84, 1–16.
Dubofksy, M., ed. 1992. *The New Deal: Conflicting Interpretations and Shifting Perspectives*. New York: Garland Publishers.
Eichengreen, B. 1992. *Golden fetters: the gold standard and the depression, 1919–1939*. New York: Oxford University Press.
Fishback, P., Haines, H. and Kantor, S. 2001. The impact of the New Deal on black and white infant mortality in the South. *Explorations in Economic History* 38, 93–122.
Fishback, P., Haines, H. and Kantor, S. 2007. Births, deaths, and New Deal relief during the Great Depression. *Review of Economics and Statistics* 89, 1–14.
Fishback, P., Higgs, R., Libecap, G., Wallis, J.W., Engerman, S., Hummel, J., LaCroix, S., Margo, R., McGuire, R., Sylla, R., Alston, L., Ferrie, J., Guglielmo, M., Pasour, E.C., Rucker, R. and Troesken, W. 2007. *Government and the American Economy: A New History*. Chicago: University of Chicago Press.
Fishback, P., Horrace, W. and Kantor, S. 2005. The impact of New Deal expenditures on local economic activity: an examination of retail sales, 1929–1939. *Journal of Economic History* 65, 36–71.
Fishback, P., Horrace, W. and Kantor, S. 2006. Do federal programs affect internal migration? The impact of New Deal expenditures on mobility during the Great Depression. *Explorations in Economic History* 43, 179–222.
Fishback, P., Wallis, J.J. and Kantor, S. 2003. Can the New Deal's three R's be rehabilitated? A program-by-program, county-by-county analysis. *Explorations in Economic History* 40, 278–307.
Fleck, R. 1999a. The marginal effect of New Deal relief work on county-level unemployment statistics. *Journal of Economic History* 59, 659–87.
Fleck, R. 1999b. The value of the vote: a model and test of the effects of turnout on distributive policy. *Economic Inquiry* 37, 609–23.
Fleck, R. 2004. Democratic opposition to the Fair Labor Standards Act of 1938: reply to Seltzer. *Journal of Economic History* 62, 231–5.
Freeman, R. 1998. Spurts in union growth: defining moments and social processes. In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, ed. M. Bordo, C. Goldin and E.N. White. Chicago: University of Chicago Press.
Friedman, M. and Schwartz, A. 1963. *A Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.
Hamby, A., ed. 1969. *The New Deal: Analysis and Interpretation*. New York: Weybright and Talley.
Hansen, Z. and Libecap, G. 2004. Small farms, externalities, and the dust bowl of the 1930s. J*ournal of Political Economy* 112, 665–94.
Higgs, R. 1987. *Crisis and Leviathan: Critical Episodes in the Growth of American Government*. New York: Oxford University Press.
Higgs, R. 1997. Regime uncertainty: why the Great Depression lasted so long and why prosperity resumed after the war. *Independent Review* 1, 561–90.
Irwin, D. 1998. Changes in U.S. tariffs: the role of import prices and commercial policies. *American Economic Review* 88, 1015–26.
Johnson, R., Kantor, S. and Fishback, P. 2006. Striking at the roots of crime: the impact of social welfare spending on crime during the Great Depression. Working Paper No. 12825, Cambridge, MA: NBER.
Kindleberger, C. 1986. *The World in Depression, 1929–1939*. Rev. edn. Berkeley: University of California Press.
Margo, R. 1993. Employment and unemployment in the 1930s. *Journal of Economic Perspectives* 7 (2), 41–59.
Mason, J. 2001. Do lenders of last resort policies matter? The effects of the reconstruction finance corporation assistance to banks during the Great Depression. *Journal of Financial Services Research* 20, 77–95.

Mason, J. and Schiffman, D. 2004. Too-big-to-fail, government bailouts, and managerial incentives: the case of Reconstruction Finance Corporation assistance to the railroad industry during the Great Depression. In *Too-Big-To Fail: Policies and Practices in Government Bailouts*, ed. B.E. Gup. Westport, CT: Greenwood Press.

Olson, J.S. 1998. *Saving Capitalism: The Reconstruction Finance Corporation and the New Deal*. Princeton: Princeton University Press.

Peppers, L. 1973. Full employment surplus analysis and structural change: the 1930s. *Explorations in Economic History* 10, 197–210.

Romer, C.D. 1992. What ended the Great Depression? *Journal of Economic History* 52, 757–84.

Schieber, S.J. and Shoven, J.B. 1999. *The Real Deal: The History and Future of Social Security*. New Haven: Yale University Press.

Seltzer, A.J. 1995. The political economy of the Fair Labor Standards Act. *Journal of Political Economy* 103, 1302–42.

Seltzer, A.J. 1997. The effects of the Fair Labor Standards Act of 1938 on the southern seamless hosiery and lumber industries. *Journal of Economic History* 57, 396–415.

Smiley, G. 2002. *Rethinking the Great Depression: A New View of Its Causes and Consequences*. Chicago: Ivan R. Dee.

Stromberg, D. 2004. Radio's impact on public spending. *Quarterly Journal of Economics* 119, 189–221.

Temin, P. 1989. *Lessons from the Great Depression*. Cambridge, MA: MIT Press.

Temin, P. and Wigmore, B. 1990. The end of one big deflation. *Explorations in Economic History* 27, 483–502.

Wallis, J.J. 1998. The political economy of New Deal spending revisited, again: with and without Nevada. *Explorations in Economic History* 35, 140–70.

Wallis, J.J. and Benjamin, D.K. 1981. Public relief and private employment in the Great Depression. *Journal of Economic History* 41, 97–102.

Wallis, J.J. and Benjamin, D.K. 1989. Private employment and public relief during the Great Depression. Working paper, Department of Economics, University of Maryland.

Wallis, J.J., Fishback, P. and Kantor, S. 2006. Politics, relief, and reform: Roosevelt's efforts to control corruption and manipulation during the New Deal. In *Corruption and Reform*, ed. E. Glaeser and C. Goldin. Chicago: University of Chicago Press.

Whatley, W.C. 1983. Labor for the picking: the New Deal in the South. *Journal of Economic History* 43, 905–29.

Wright, G. 1974. The political economy of New Deal spending: an econometric analysis. *Review of Economics and Statistics* 56, 30–8.

# quantitative easing by the major western central banks during the global financial crisis

### Short-term interest rates had been the primary policy tool of central banks prior to the financial crisis

Since the end of the post-Second World War Bretton Woods System of quasi-fixed exchange rates in the early 1970s and the advent of floating exchange rates across the developed world, the primary policy tool of the major central banks for managing aggregate demand and controlling inflation has been short-term interest rates. In the USA the main policy rate is called the Federal Funds Rate (FFR), whilst in the UK, Eurozone and Japan it is the Bank Rate, Main Refinancing Rate (MRR) and the Uncollateralised Overnight Call Rate (UOCR) respectively. Changes in short-term interest rates impact the economy via their influence on other types of interest rates (e.g. mortgage rates, auto loan rates and business loan rates), which affect the spending/saving decisions of private sector agents, and through influencing expectations about the future path for the economy.

The aftermath of the bursting of the late 1990s technology bubble illustrated the potency of interest rates as a tool to manage aggregate demand. Fearing a repeat of the deflationary spiral that had beset Japan a decade earlier after the bursting of its real estate and stock market bubbles, the US Federal Reserve (Fed) slashed the FFR from 6.5% in May 2000 to a historic low of 1% in June 2003 (Figure 1) and it remained at extremely low levels until the latter part of 2005. A combination of extremely accommodative monetary policy and significant fiscal easing fuelled a sharp upswing in the US economy amid a booming housing market and robust gains in consumer spending. Aggressive monetary easing was also pursued by other central banks. The Bank of England (BoE) cut the Bank Rate to 3.5% in July 2003, which was its lowest since 1955, helping to fuel a doubling of national house prices between the ends of 2000 and 2005. Meanwhile, the European Central Bank (ECB) cut its MRR from a peak of 4.75% in October 2000 to a record low of 2% in June 2003, where it remained for over two years until December 2005. The easy monetary policy did not have much impact on the German economy, which was undergoing a painful restructuring in an attempt to regain competitiveness lost during the previous decade amid reunification of East and West Germany. In contrast, negative real interest rates in the peripheral economies such as Spain and Ireland (where inflation was higher than the ECB's MRR) fuelled major housing market booms and what initially appeared to be a virtuous cycle of rapid increases in wages, wealth and consumer spending.

The US housing market began to weaken materially in 2007 amid stretched affordability conditions, and as rising interest rates forced homeowners, particularly at the sub-prime level, to default *en masse* on their mortgages. Major ripple effects from this began to be felt widely across financial markets. Private interbank lending markets

became dysfunctional as financial institutions began to refrain from lending due to concerns about their potential exposures to huge losses on housing-related investments and amid fears about the creditworthiness of their counterparts. This was reflected in a surge in the LIBOR-OIS spread (which is a measure of the willingness of banks to lend to each other) from an average of around 10 basis points in the years leading up to the financial crisis to over 350 basis points in October 2008 (Figure 2).

Amid growing fears of a recession the Fed began to accelerate the pace of its interest rate cuts, and similar aggressive reductions were made by the BoE and the ECB. By the time of the major intensification of the financial crisis with the collapse of the investment bank Lehman Brothers on 15 September 2008, US interest rates were already at just 2%, and before year-end they had fallen to a record low of just 0–0.25% (Figure 1). This was viewed as the effective lower bound for the FFR which would not impair the effective functioning of money markets (an important source of short-term finance for the corporate sector). In the UK, the BoE followed suit, slashing the Bank rate to just 0.5% by March 2009, its lowest level in over 300 years. This was thought to be the lower bound for rates in the UK, as the BoE's Monetary Policy Committee (MPC)

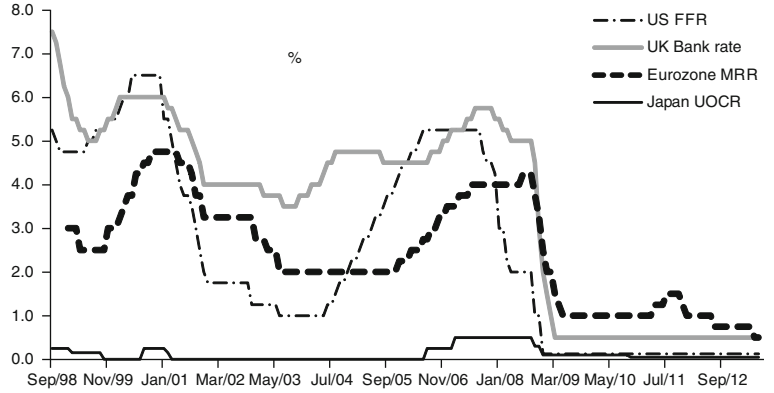



**Figure 1** Central bank short-term interest rates, % (source: Haver Analytics).

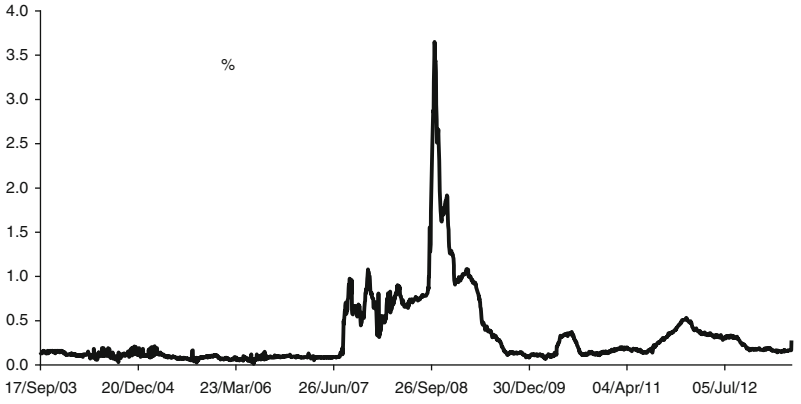



**Figure 2** US 3-month LIBOR-OIS spread, % (source: Haver Analytics).

feared that additional reductions could further weaken the profitability of a number of financial institutions whose primary assets (e.g. mortgages) were linked to the Bank Rate (Bank of England, 2009). The MPC was also concerned about the impact that further cuts would have on the functioning of money markets. The ECB cut interest rates aggressively too, with the MRR reaching just 1% in May 2009 (this was never viewed as a floor, however, and it would ultimately fall to 0.75% in July 2012 as the Eurozone sovereign debt crisis intensified, and then 0.5% in May 2013, as the economy continued to languish in recession).

Despite the extremely aggressive monetary easing by central banks, banks and other financial institutions continued to refuse to lend to each other, and hence the private markets for credit, the key lubricant in the wheels of the advanced economies over the last several decades, became almost completely dysfunctional. Faced with the likelihood that financial institutions would be forced into continued fire sales of their marketable assets in order to meet their own liquidity needs, which would likely fuel a self-reinforcing downward spiral, the major central banks stepped in by significantly expanding their traditional roles as 'lenders of last resort'. This involved central banks using their balance sheets to lend emergency funds to financial institutions (and non-financial institutions in certain cases) for much longer periods of time and against a much broader range of collateral (e.g. collateral that had become illiquid in financial markets at the time) than was typically the case in normal times. The Fed was at the forefront of this, with schemes such as the 'Term Auction Facility', 'Term Asset-Backed Securities Loan Facility' (TALF) and the 'Commercial Paper Funding Facility' (Bernanke, 2009).

With the US and UK economies contracting extremely sharply in 4Q 2008 and 1Q 2009 and amid fears of a deflationary spiral increasing, traditional central bank policy rule models (e.g. Taylor Rules, which prescribe a rule for policy rates based on the output gap and inflation in an economy; see Taylor, 1993) were indicating that significant further monetary easing was necessary (i.e. theoretically highly negative interest rates would be optimal). But with central bank policy rates at their effective lower bounds, the authorities needed to consider other more unconventional monetary policies in order to further reduce borrowing costs, increase liquidity and boost aggregate demand across their economies (or at least prevent it from falling further).

## Quantitative easing begins in the US and the UK

The main template for the course on which the Fed and the BoE were about to embark was provided by Japan, which had begun a program of quantitative easing (QE) earlier in the decade. From March 2001 to early 2006, after years of extremely low interest rates had failed to restore lending by Japanese banks, the Bank of Japan (BoJ) reduced its main interest rate to zero and changed the operational target of monetary policy from the Uncollateralized Overnight Call Rate (UOCR) to the absolute level or quantity of current account balances (bank reserves plus deposits of non-bank financial institutions) held at the central bank. The rationale for setting targets for current account balances held at the central bank by financial institutions was that, if the banks held large stocks of reserves (thus facing little prospect of a liquidity shortage), then they would feel more confident about resuming lending to the private sector.

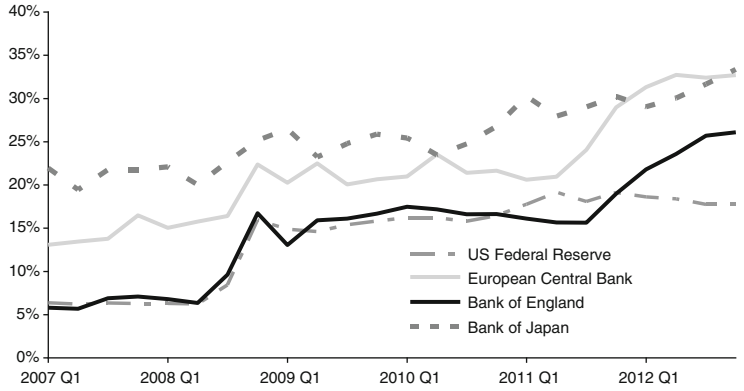



**Figure 3** Central bank balance sheets, % of GDP (source: Morgan Stanley Research, Haver Analytics).

The Fed began its own version of QE in November 2008, dubbed Large Scale Asset Purchases (LSAPs), with the announcement that it would purchase $100bn of agency debt (debt issued by the Government Sponsored Enterprises (GSEs) such as Fannie Mae and Freddie Mac) and $500bn of Mortgage Backed Securities (MBS) backed by the GSEs. In addition, in March 2009 the Federal Open Market Committee (the Fed's main monetary policy setting body) voted to increase the size of its balance sheet by buying an additional $100bn of agency debt, $750bn of MBS backed by the GSEs and $300bn of longer-term Treasury securities. By the time QE1 ended in early 2010, the Fed's balance sheet had expanded to 16% of GDP (Figure 3).

The BoE began QE with £75bn of purchases of long-term government bonds (commonly referred to as gilts in the UK) in March 2009 via its newly created Asset Purchase Facility (APF), and it also bought small quantities of commercial paper and corporate bonds to help provide liquidity in those markets which had become increasingly dysfunctional. By the time QE1 had ended in January 2010, the BoE had purchased around £200bn of government bonds, equal to around 14% of GDP (Figure 3) and 20% of all outstanding government debt at the time.

The ECB did not embark on QE, but it did institute a number of lending programs that significantly increased the size of its balance sheet (Figure 3). In October 2008 it announced that borrowing from its weekly Main Refinancing Operation (MRO) would be allowed on a fixed rate tender procedure with full allotment, which effectively meant that financial institutions could borrow unlimited funds from the ECB at the specified interest rate. Crucially, it also expanded its provision of credit at longer maturities through longer-term refinancing operations (LTROs) over periods of three, six and twelve months on a fixed rate tender procedure with full allotment. This guaranteed the availability of liquidity over the medium term for financial institutions, reducing the risk that they would need to reduce their provision of credit to the private sector in order to raise funds to meet redemption payments on their own bank bonds.

The ECB did directly purchase some assets via its Covered Bond Purchase program (CBPP) and its Securities Markets Programme (SMP) introduced in June 2009 and May 2010 respectively. However, as Giannone *et al.* (2012) note, these operations were

small when compared with the size of the ECB's balance sheet and the main aim of the purchases was to support an improvement in functioning in what had become some dysfunctional asset markets. Moreover, they were not intended to expand overall liquidity, since any increases were offset by special liquidity absorbing operations, though these were perhaps more presentational than real.

Antolin Diaz (2013) notes two primary reasons why the ECB did not undertake large-scale asset purchases like the Fed and BoE. The first is the different structure of financial markets in the various regions. Bank-based financing of households and non-financial corporates is much more important in the Eurozone, whereas in the USA market-based financing and securitisation are more common. For example, more than 70% of the stock of external financing of the non-financial corporate sector in the Eurozone comes from banks, with less than 30% from financial markets and non-bank funding sources. Almost the opposite is true in the USA. Second, the Fed (subject to details discussed later), BoE and BoJ have the ability to undertake purchases of a single riskless asset (e.g. the bonds of their own sovereign government), whilst the existence of 17 sovereign governments in the Eurozone is an important institutional barrier to broad-based operations in government bond markets. Cour-Thimann and Winkler (2012) suggest that the Treaty on the Functioning of the European Union, including the prohibition of monetary financing by the central bank (Article 123), is also likely to have been a factor in why the ECB eschewed direct asset purchases.

Figure 4, from the BoE's 3Q 2011 *Quarterly Bulletin* provides a stylised guide to the main transmission channels through which QE was expected to boost economic growth and inflation (also see Krishnamurthy and Vissing-Jorgenson, 2011). The main channels include Confidence, Policy signalling, Portfolio rebalancing, Market liquidity and Money.

**Confidence:** By using a new and potentially powerful policy tool in a relatively aggressive manner, it was hoped that it would increase expectations amongst private sector agents (households, non-financial corporations and financial institutions) that it would halt the downward momentum in the economy and eventually generate a recovery in activity and asset prices. Policymakers were in effect trying to boost (or arrest the decline in) what Keynes famously referred to as the 'animal spirits' of the private sector (Keynes, 1936). Hence, by changing the spending behaviour of households and businesses, growth would likely be stronger than in the counterfactual with no asset purchases.

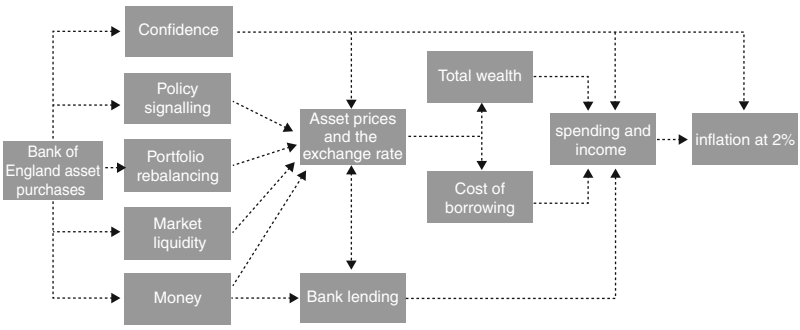



**Figure 4** Main QE transmission channels (source: Bank of England Quarterly Bulletin, 3Q 2011).

**Policy signalling:** Asset purchases are likely to lead investors to expect short-term policy rates to remain lower for a longer period than previously expected. This primarily occurs through two channels. First, the decision to ease policy provides new information about the committee's current assessment of the economic outlook and its policy bias. Second, Clouse *et al.* (2000) suggest that asset purchases might be a way of signalling a commitment to keep short-term policy rates low in the future, as the increase in the size of the bank's long-term bond holdings would leave it vulnerable to capital losses in the event that it were to raise policy rates sooner or more aggressively than financial markets expected. For economists of the New Keynesian School policy signalling is likely to be the most powerful channel of QE in boosting the economy (Woodford, 2012). In addition, the more aggressive monetary easing should boost inflation expectations and reduce the probability associated with a tail-risk scenario such as a Japanese-style deflationary cycle. The combination of delayed expectations of the timing of the first tightening in short-term interest rates and increased inflationary expectations should ensure that the projected path of real short-term interest rates is lower than prior to the onset of QE. This should encourage the private sector to bring forward future spending and also boost asset prices.

**Portfolio rebalancing:** This is one of the key transmission mechanisms cited by both the BoE and the Fed. The BoE aims to purchase government bonds from the non-bank private sector such as pension funds, insurance companies and hedge funds. The BoE gains an asset in the form of a government bond and a liability in the form of the electronically created money it has deposited in the sellers account which adds to outstanding reserve balances. Given that the seller is unlikely to view the cash as a perfect substitute for the asset sold (given differences in risk characteristics e.g. duration, yield etc.) they are likely to want to use the cash to rebalance their portfolios by purchasing assets that are better substitutes such as corporate bonds and equities (Brunner and Meltzer, 1973; Friedman and Schwartz, 1982). This should boost the prices of these other assets and reduce the yield on them (given that the yield moves inversely to prices). Higher asset prices increase wealth across the economy, whilst the lower yields on these assets reduce the cost for corporations of raising external financing via corporate debt and equity markets. According to Joyce *et al.* (2011), 'While policy signalling effects expected policy rates, portfolio balance effects work by reducing the spreads of longer-term interest rates over expected policy rates (term premia) and the required return on risky assets relative to risk-free assets (risk premia) more generally'.

**Market liquidity:** When financial markets have become dysfunctional, central bank asset purchases can improve the functioning of these markets by increasing liquidity through actively encouraging trading. This is where the central bank adopts the role of market maker of last resort. Asset prices thereby increase via lower premia for liquidity.

**Money:** By purchasing assets from the non-bank private sector with newly created money, the BoE injects monetary base (MB) into the financial system, and these extra funds should end up in deposits at the commercial banks. The extra liquidity at banks should fuel an increase in bank lending and a rise in the broader measures of money supply such as M4.

The Fed and the BoE's asset purchase programs differed from that undertaken by the BoJ in the previous decade in a number of ways. First, operationally the Fed and

the BoE had been targeting the asset side of their balance sheets by purchasing various assets, and the large increase in bank reserves was a by-product of this. In contrast, the BoJ actually targeted the liability side of its balance sheet by setting targets for current account balances held at the central bank by financial institutions, with the rationale being that, if the banks held large stocks of reserves, they would feel more confident about starting to lend again. Second, most of the government bonds purchased by the BoJ were from banks and were short-dated in nature. In contrast, the Fed and BoE have focused on purchasing longer-dated assets (from non-banks), which, given their longer duration, are likely to be less close substitutes for money, since the portfolio rebalancing channel is seen as one of the key conduits for QE to boost the economy. Meanwhile, one difference between the Fed and the BoE was that the former was said to have assumed credit risk by buying agency securities and MBS securities, whilst the BoE stuck primarily to government bonds. That said, given that the GSEs were taken into government conservatorship by the US Treasury in late 2008, a number of commentators have suggested that the distinction from government bonds is now somewhat less marked.

## The impact of QE1

Following the introduction of QE1 in the USA and various other countries (together with other crisis management tools and fiscal stimulus policies) there was a broad-based recovery in global financial markets. The key US stock market index, the S&P 500, troughed at 676.5 on 9 March 2009 (Figure 5), which was approximately 60% below its pre-crisis peak. It then went on to increase by almost 75% over the following 12 months until the end of March 2010. Other equity markets also rebounded strongly during this period, with gains of 60%, 67% and 58% in the UK FTSE 100, German DAX 30 and French CAC 40 respectively. Moreover, the LIBOR-OIS spread fell back sharply amid some thawing in the private bank lending markets, and the yields on both investment and sub-investment grade corporate bonds fell back sharply as investors became increasingly confident that the swift and aggressive action of policymakers had averted a collapse of the financial system and a repeat of the Great Depression. The VIX, a measure of stock market volatility for the S&P 500, fell back from a peak of


**Figure 5** Standard & Poors 500 stock market index (source: Haver Analytics).

over 80 in November 2008 to 17 at the end of March 2010, reflecting a significant improvement in risk appetite amongst investors.

Reflecting the improving sentiment in global financial markets and the enactment of large fiscal stimulus packages in countries such as the USA and China, business confidence began to recover strongly from the spring of 2009 and the National Bureau of Economic Research's (NBER) Business Cycle Dating Committee judged that the US economy exited its 18-month long recession in June 2009. The US economy then went on to grow at an annualised pace of over 2% per annum in 2H 2009 and 1H 2010 (Figure 6). Whilst this was a much more subdued pace of recovery than had historically been the case from recessions, it was a marked turnaround from the rapid pace of contraction in prior quarters. Meanwhile, the UK and Eurozone economies both exited recession in 3Q 2009, whilst the Chinese economy was growing again at an annual rate of over 10% by the second half of 2009 and through 2010.

Neither politicians nor policymakers were under any illusions that an extremely difficult road lay ahead given the huge increase in governments' fiscal deficits and debt burdens in the aftermath of the financial crisis (Reinhart and Rogoff, 2009) and amid a compelling need for economies such as the UK, USA and Spain to rebalance away from consumer and housing-dependent growth models to ones with a greater contribution from exports and business investment. Nevertheless, optimism was beginning to build that developed economies had avoided a worst-case scenario, were staging something of a recovery and might be able to avoid the 'lost decade' of growth that Japan had experienced in the aftermath of its own financial crisis.

Indeed, it was felt that central bankers had learnt the lessons from the mistakes of their predecessors in the USA in the 1930s and Japan in the 1990s and had acted very swiftly and aggressively. Critics of the Fed's actions during the Great Depression (see Friedman and Schwartz, 1963) suggest that it should have done more to prevent bank failures through its role as lender of last resort, which would have helped maintain public confidence in the banking system and prevented such a large fall in the money multiplier. In addition, they suggest that it should have injected significantly more monetary base into the banking system to offset the contractionary impact on broader measures of money supply from the sharp decline in the money multiplier. Meanwhile, the Japanese authorities were criticised for acting too slowly and timidly in response to

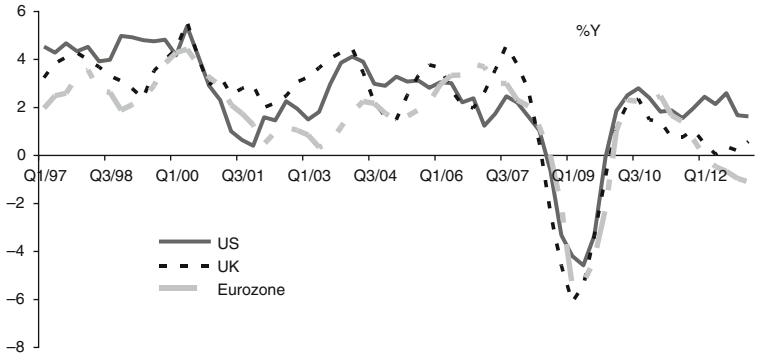



**Figure 6** GDP growth in the major Western economies, % change y/y (source: Haver Analytics).

the bursting of their real estate and stock market bubbles in 1989–90 (Bernanke, 2000). It took over five years for interest rates to be cut to 0.5% and the BoJ did not embark on quantitative easing until over a decade after the bursting of the bubble. Wadhwani (2013) also notes that the Japanese authorities made the mistake of allowing the Japanese yen to strengthen sharply after the bubble burst, something the US and UK authorities managed to avoid (sterling had actually depreciated by around 25% on a trade-weighted basis leading up to the crisis).

There have been numerous empirical studies of the impact that QE1 had on GDP and inflation in the US and UK economies. Most have generally been split into two main stages. In the first stage, event studies and econometric analysis of yield curves have been used to estimate the impact that QE had on long-term government interest rates. Examples of such exercises for the UK are Joyce *et al.* (2011) and Caglar *et al.* (2011) in Chadha *et al.* (2012), whilst for the USA we have Bomfim and Meyer (2010), D'Amico and King (2010), Gagnon *et al.* (2011) and Yellen (2011). The general conclusion of these studies was that QE1 reduced long-term government interest rates by up to 1 percentage point in both the USA and the UK.

In the second stage, one then plugs the estimated decline in riskless government bond yields into a macroeconomic model of the economy – perhaps a Vector Autoregressive Model (VAR) or a Dynamic Stochastic General Equilibrium Model (DSGE) model – and then estimates the impact on GDP and inflation. Using this approach for the UK, Kapetanios *et al.* (2012) estimated that QE1 may have boosted GDP and inflation at its peak by around 1.5% and 1.25% respectively. For the USA, Chung *et al.* (2011) estimated that QE1 and QE2 will have boosted GDP by around 3% by 2H 2012 and lowered the unemployment rate by around 1.5 percentage points. They also estimate that inflation is around 1 percentage point higher than it would have been without asset purchases. Using simulations, they suggest that the policy stimulus provided by the asset purchases is equivalent to a cut in the Federal Funds rate of around 3 percentage points.

Whilst all such calculations are tentative, the general conclusions seem to suggest that the early rounds of QE had a very significant positive impact on the US and UK economies. Whilst less analysis has been done of the non-standard measures applied by the ECB, the balance of evidence generally points to a very favourable impact too. Giannone *et al.* (2011) conclude that 'the introduction of non-standard measures has supported the availability of monetary liquidity to the non-bank private sector and flow of bank loans to households and, especially, corporations – resulting in an outcome that largely mimics what would have been anticipated in the face of the observed fall in economic activity were the financial sector to be functioning normally'. Lenza *et al.* (2010) estimate that compared to a counterfactual of no non-standard measures by the ECB, inflation would have been around 0.5 percentage points lower at the beginning of 2010 and industrial production would have been almost three percentage points lower by mid-2010.

Goodhart and Ashworth (2012) suggest that by focusing on sovereign government bond yields, existing studies of QE1 may have actually underestimated the boost it provided to GDP growth. The use of sovereign bond yields would appear to derive from the assumption of DSGE models that there is no default risk, so the only interest rate that matters is the 'riskless' rate on government debt. Given that the financial crisis was

primarily driven by increasing default risk, the latter assumption is, however, inappropriate. Using the decline in non-investment grade corporate bond spreads recorded by event study analysis of QE1 in the UK and substituting this into a VAR model of the economy instead of similar maturity government bond yields, the authors find that the boost to UK GDP may have been larger than previous studies suggest. By sharply reducing risk spreads and calming risk perceptions in financial markets the boost to GDP may have been much greater than typically estimated.

Another key positive effect of QE in the USA and UK has been in helping to limit the deterioration in the near-term debt dynamics (see Equation 1). QE helps the debt dynamics in two main ways (in addition to boosting nominal GDP growth). First, the coupon payments on the assets purchased by the Fed and the BoE are significantly greater than the interest payments on the new central bank reserves created to fund these purchases. These net interest receipts are then remitted to the government on a regular basis. Second, by reducing interest rates on government bonds across the curve, it reduces the interest payable on new bond issues from the government. Hence, over the longer-term horizon, based on the most plausible future paths for interest rates, the government finances are likely to have received a decent boost on net from the central bank asset purchase programs.

$$\Delta\,\text{Debt/GDP}_t = \text{Debt/GDP}_{t-1} \times \left[\frac{(\text{Ave int rate} - \text{Nom GDP growth})}{(1 + \text{Nom GDP growth})}\right] - \frac{\text{Primary Balance}}{\text{GDP}} \quad (1)$$

Debt/GDP = government debt-to-GDP ratio, Ave int rate = Average interest rate on government debt, Nom GDP = Nominal GDP, Primary Balance = government borrowing/lending excluding interest payments

One area where QE1 was less successful than had generally been expected *ex ante* was in boosting the broader measures of the monetary aggregates and the provision of credit to the private sector (although central bankers suggest *ex post* that preventing a major contraction in the broad money supply and credit was in and of itself a major success). Goodhart (2010) notes 'that the standard approach in monetary economics to explaining the supply of money, and the provision of bank credit to the private sector, has been the money multiplier approach [see Equation 2], whereby the central bank sets the high-powered monetary base (H), and then the stock of money (M) is a multiple of that. But when the authorities in the major developed countries attempted to use this relationship to expand the money stock and bank lending by force feeding banks with base money, the prior relationship collapsed'. As Table 1 highlights, despite the huge percentage increases in base money across the major developed economies, measures of broad money and bank lending were broadly unchanged.

$$M = H \bullet \frac{\left(1 + \frac{C}{D}\right)}{\left(\frac{R}{D} + \frac{C}{D}\right)} \leftarrow \text{Money Multiplier} \quad (2)$$

$M$ = Broad money supply, $H$ = Monetary base, $C/D$ = Currency/Deposits ratio, $R/D$ = Bank reserves/Deposits ratio

**Table 1** *Tiny multipliers (private non-financial corporations and the household sector), % ch (source: Goodhart, 2010).*

| June 2008–2009 (for Japan: March 2001–2006) | Change in bank reserves held at central bank | Change in broad money | Change in bank lending to private sector |
|---|---|---|---|
| UK | 371 | 2 | 1 |
| USA | 1,853 | 9 | 4 |
| Eurozone | 122 | 4 | 2 |
| Japan | 103 | 8 | – 17 |

Goodhart and Ashworth (2012), focusing specifically on the UK, suggest there was a major and a minor reason behind the sharp decline in the money multiplier. The major cause was the decision by banks to keep a significant ratio of their customer deposits at their own central bank in the form of reserves, i.e. the $R/D$ ratio soared. The likely cause of this was the freezing of the private interbank lending market, which deprived financial institutions of a potential source of funds, and also that it was a safe use of excess funds, with a small but guaranteed return, amid heightened counterparty risks. In addition, the increased credit risks and capital requirements associated with lending to the private sector deterred banks from doing so, except at enhanced spreads and amid more stringent collateral terms. Meanwhile, the risks and capital requirements of lending to the public sector were effectively zero, even though the rate on deposits at the BoE was very modest at just 0.5%. The minor reason was that during the time QE1 was put in place, a combination of toughened regulatory requirements and a concern with self-preservation in a crisis was encouraging banks to raise new capital, both long-term debt and equity. This was purchased by the drawing down of cash deposits held at the central banks, typically by pension funds and insurance companies.

## QE2, QE3 and the LTROs

After something of a recovery in economic growth over 2H 2009 and 1H 2010, in part due to QE1, the central banks discontinued this policy with asset purchases ending in early 2010 in both the USA and the UK. However, after some initial optimism that a self-sustaining recovery might have been in the offing across the developed world, growth began to weaken again in the second half of 2010 as the boost from the traditional inventory replenishment cycle and fiscal stimulus packages began to fade.

## Further steps in the USA

With the US unemployment rate extremely elevated at almost 10% and a low and declining inflation rate beginning to stoke concerns about the risks of deflation (Dudley, 2010), the Fed embarked on a second round of Large Scale Asset purchases (LSAP 2) in November 2010. This consisted of $600bn of government bond purchases through mid-2011. The Fed did not announce a further round of QE when LSAP 2 ended, but with problems in European sovereign debt markets beginning to spill over to the wider global economy and political gridlock in Washington raising the risk of fiscal tightening at year-end, it announced a Maturity Extension Program (MEP) at its

September 2011 meeting, whereby the Fed would sell its holdings of shorter-term bonds and use the proceeds to purchase $400bn of longer-term bonds through mid-2012. This policy, commonly referred to as Operation Twist, was first attempted in 1961 (Swanson, 2011; Modigliani and Sutch, 1966). The key stimulus from MEP came through the portfolio rebalancing channel as occurs in outright QE (with a key aim being to reduce mortgage rates). In contrast to QE, however, MEP does not increase the size of the Fed's balance sheet. This could help insulate the Fed from growing criticisms in Conservative circles that it was 'money printing' and 'debasing the US dollar'. The Fed announced an extension of its MEP program in June 2012 until the end of that year.

The balance of academic evidence suggests that LSAP 2 reduced long-term government interest rates in the USA by up to 0.45 percentage points. Moreover, the positive impact of the Fed's actions was widely felt across financial markets, with the S&P 500 and market implied inflation expectations increasing quite sharply over the following months (Figures 5 and 7). As Bernanke (2012) notes, the boost to the stock price is a particularly important channel because stock values affect both consumption and investment decisions. Meanwhile, via what Krishnamurthy and Vissing-Jorgenson (2011) characterise as the 'Inflation Channel', the boost to inflation expectations helped reduce the risk that potentially self-fulfilling deflationary expectations may begin to become embedded amongst investors.

Meanwhile, the Fed announced LSAP 3 in September 2012 with $40bn per month in purchases of mortgage-backed securities (MBS) on an open-ended basis. This was expanded to $85bn per month on an open-ended basis in December, with an additional $45bn per month of government bond purchases. There were a number of reasons for making the purchases open-ended. First, it would send a signal to financial markets and the private sector that the Fed was committed to doing whatever it takes to fuel a sustainable economic recovery. Second, a number of FOMC officials felt that an ongoing program of asset purchases, whose size could be adjusted in response to the economic data, was preferable to previous rounds of QE which involved rather abrupt

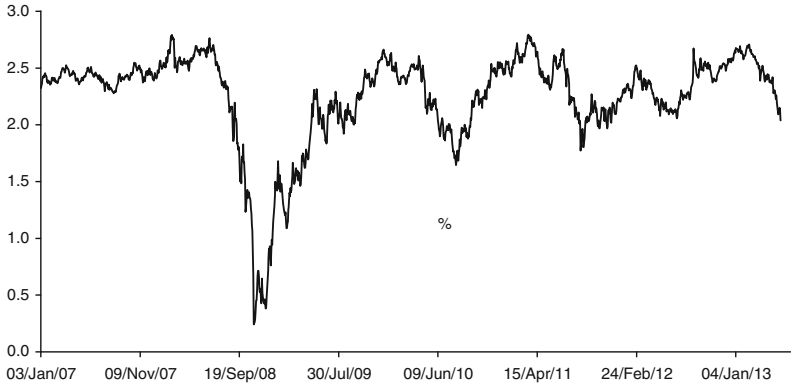



**Figure 7** US 10-year inflation expectations amongst investors (zero-coupon bond yield), % (source: Haver Analytics).

withdrawals of stimulus as the programs ended, only to be followed by renewed asset purchases once economic growth began to ebb.

### Further steps in the UK

In the UK, with the economy on the verge of a 'double-dip' recession in the Autumn of 2011, the MPC embarked on QE2 announcing £75bn of government bond purchases at its October meeting, which was followed by a further £50bn in February 2012. After a brief pause at the May and June meetings, the MPC sanctioned the beginning of QE3 with £50bn of asset purchases at its July meeting. This took the size of its stock of asset purchases to £375bn, which was approximately 25% of GDP. Ashworth and Goodhart (2011) suggested that whilst QE2 in the UK was on balance likely to be beneficial, its efficacy was likely to be less than QE1. Sovereign bond yields were already at very low levels and given the weakening economic outlook and failure of QE1 to generate a sustainable recovery, the ability of asset purchases to positively influence the behaviour of investors and households was likely to have diminished somewhat. In Goodhart and Ashworth (2012), the authors find evidence of significant diminishing returns from QE2 versus QE1. Replicating the event study analysis carried out by the BoE on QE1 for QE2, they find that gilt yields were little changed during QE2. In addition, they find little evidence that QE2 has fuelled much portfolio rebalancing into equities, suggesting the boost to household wealth from QE2 was likely to have been very modest.

### Further steps in the Eurozone

With significant tensions in peripheral sovereign bond and bank funding markets threatening to cause a major credit crunch across the Eurozone, the ECB launched two three-year long-term refinancing operations (LTROs) in December 2011 and February 2012. What was particularly striking about these was the fact that the ECB was making the loans available for significantly longer than has typically been the case for central bank loans and that they were unlimited in size: financial institutions could borrow as much as they liked at just 1% (the prevailing ECB MRR rate at the time, although the rate on the loans tracks movements in the MRR) subject to collateral constraints. Moreover, the range of collateral eligible to be used was further broadened to include bank loans to the private sector. Gross lending over the two LTROs was over €1 trillion, with net lending around half of that as banks scaled back their use of other liquidity providing operations. This latter is equivalent to around 5% of GDP.

With wholesale bank funding markets still largely closed across the region and financial institutions needing to redeem a significant amount of their maturing funding needs over the upcoming year, the provision of liquidity for an extended period through the LTROs reduced the risk of a major disorderly deleveraging of bank's balance sheets which would have entailed self-reinforcing sales of government bond holdings and a major 'credit crunch' across the region. In addition, financial institutions used much of the cheap funding to purchase higher yielding government bonds, which helped reduce sovereign interest rates and boosted the asset values of the sovereign debt held by the banks themselves. Cour-Thimann and Winkler (2012) suggest that the LTROs 'triggered crowding in phenomena in debt issued by non-bank sectors. The

reduction in the net supply of bank debt contributed to improvements in government bond markets but also allowed, via substitution effects in portfolios of investors, a rebound in non-financial corporate bond issuance'.

## Why hasn't inflation accelerated during QE?

When the Fed and BoE first introduced QE in 2008-2009, a number of commentators characterised it as 'money printing', involving the direct monetary financing of government deficits and debt by the central banks. The comparison was made with the German Weimar Republic in the 1920s and Zimbabwe in recent years, where, in cooperation with their governments, the central banks financed huge increases in government spending, eventually resulting in hyperinflation.

Whilst there are a number of similarities and the dividing lines can become increasingly blurred (Turner, 2013), it is important to clarify the key differences between the *current* QE programs in the USA and UK and the direct monetary financing of government deficits, as has typically occurred in the past. These sometimes subtle differences have major implications for the likely inflationary consequences of the two different policies.

First, the current QE programs are only intended to be temporary. Both the Fed and the BoE have clearly stated their intention to sell back the assets they have accumulated at some future data as and when the economic conditions allow, meaning that the increase in the monetary base over recent years is likely to be reversed. With monetary financing of deficits the increase in the monetary base has typically been permanent, although it does not necessarily have to be. Second, with QE the decision to purchase assets by central banks has been made independently of governments and has been taken with the intention of boosting aggregate demand and so preventing inflation *falling* materially below its target/objective. As Miles (2012) notes, 'The decision of the MPC to embark on asset purchases on an enormous scale was not done because it had abandoned the inflation target, it was done *because of* the inflation target'. In an attempt to underline the separation from the fiscal authorities, bond purchases during QE have been made via the secondary market from private sector agents, rather than in the primary market from governments. In contrast, with monetary financing of deficits, asset purchases by the central bank have typically been done in cooperation with (or by direction from) the government and with little sensitivity for their inflationary consequences. Third, monetary financing of deficits usually occur alongside rapid increases in government spending. The later rounds of QE have come at a time when governments have been attempting to reduce their structural budget deficits. Fourth, given the differences outlined, the risk of a major unhinging of inflationary expectations amongst households, firms and investors is materially less likely, but certainly not insignificant, under QE than OMF.

Nevertheless the risks remain of a potential future sharp acceleration of inflation due to QE. The 1970s highlighted how excessive policy accommodation of an oil price shock via less unconventional policy tools (e.g. short-term interest rates) helped fuel a wage–price spiral and a sharp acceleration in consumer price inflation. There appear to

be three key reasons why QE has thus far not fuelled a sharp jump in inflationary pressures:

1. Despite the massive increase in the size of the monetary base in the US and UK economies, this has not fuelled much of a rise in broader aggregates of money supply and lending to the private sector (Table 1). The unwillingness of financial institutions to lend to the private sector has meant that the monetary base has remained at the central bank in the form of excess reserves. Hence, via the money and credit channel, the boost to aggregate demand has been significantly less than typically would have been the case with a normally functioning banking system. If financial institutions should, in future, shift towards greater credit expansion, central banks are confident that they have the tools to mitigate this.
2. Spare capacity in the developed economies has been exerting significant downward pressures on inflation. There is a widespread debate in the academic literature about the magnitude of spare capacity in economies such as the USA and UK, with measures of spare capacity such as the output gap being notoriously difficult to measure. Supply optimists highlight that output in the USA and UK remains significantly below a continuation of its pre-crisis trend, whilst the more pessimistic suggest that the financial crisis has destroyed previously productive resources and that much of the increase in unemployment may be structural in nature. In truth, it is difficult to say with great certainty, but a reasonable assessment of the evidence suggests that a decent amount of spare capacity exists, particularly in labour markets. Unemployment rates of 7.6% and 7.8% in the USA and UK respectively remain well above what most people would consider their Non-Accelerating Inflation Rate of Unemployment (NAIRU) (Ball and Mankiw, 2002). This labour market slack and heightened uncertainty about job prospects for those in employment has meant that workers have had little bargaining power to push for wage increases. Such a wage–price spiral has typically been a factor in pushing up inflation in the past. In fact, average earnings growth has been extremely subdued in the USA and UK in recent years (Figure 8). Moreover, with final demand very weak, firms have been reluctant to try and raise their prices.

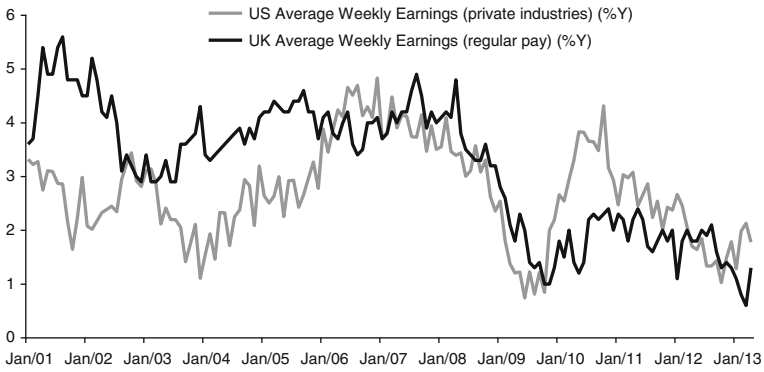



**Figure 8** Average earnings growth US and UK, % change y/y (source: Haver Analytics).

3. Inflationary expectations have remained well anchored around the central bank's target levels. A combination of downward pressure on wages due to slack in labour markets, modest pricing power amongst firms and central banks that (whilst increasingly exhibiting more flexibility) have remained firmly committed to hitting their inflation targets has helped keep inflation expectations from becoming unmoored in an upward direction in recent years. Indeed, it must be remembered that much of the unconventional policy easing has actually been put in place to prevent the risk of a sharp downward move in inflation expectations, in the hope of averting a Japanese style deflationary cycle.

The risk of acceleration in inflation might increase materially once a more sustainable recovery in economic growth begins to take hold in the USA and UK. First, as the credit risk associated with lending to the private sector declines, banks are likely to start using some of their excess reserves to extend loans. This should fuel an increase in the money multiplier, broader measures of money supply and ultimately aggregate demand. Second, as labour markets begin to improve workers will have greater confidence to push for higher wage settlements and firms may attempt to restore profit margins, which are likely to have contracted over recent years given the weakness of final demand. Hence the onus will clearly fall on central banks to use all of the tools at their disposal in a timely fashion and begin to 'remove the punchbowl' once confidence builds that enduring recoveries are in place. The huge increase in their balance sheets over recent years, however, will make this task problematical.

## Worries about other potential negative side effects from quantitative easing have grown

As the longevity and magnitude of QE and other unconventional monetary policies have increased over recent years, fears have grown about other potential unintended negative side effects of such policies. Most prominent amongst these have been concerns about the risks to financial stability. In particular, in an environment where interest rates are extremely low across the yield curve, incentives increase for investors to engage in an unsafe 'reach for yield' either through excessive use of leverage or through other forms of risk-taking (Bernanke, 2012, 2013; Stein, 2013). In its June 2013 Financial Stability Report, the BoE suggested there was some evidence of this occurring in certain market segments (e.g. high yield credit instruments in the USA).

Meanwhile, former BoJ Governor Masaaki Shirakawa (2012) suggests that a prolonged period of very accommodative monetary policy can reduce the incentives towards balance sheet repair (by banks, households and firms) and may induce investment projects that are only profitable at very low interest rates. The potential inefficient allocation of resources could have a negative impact on productivity and the economy's growth potential. He also suggests that beyond a certain threshold further monetary easing can squeeze the margins of financial intermediaries and discourage credit provision. This is due to the maturity transformation role played by financial intermediaries, which borrow funds in the short-term and make loans over the longer-term. By reducing interest rates on longer-term government bonds, QE can reduce the net interest

margins of financial institutions. The Bank for International Settlements (2012) suggests that there are signs of this happening already, 'as the more recent flattening of the yield curve in the US and UK has gone hand in hand with a drop in banks' net interest margin'.

Goodhart and Ashworth (2012) also highlight the negative impact that low long-term government bond yields have had on the solvency of defined-benefit pension funds (also see National Association of Pension Funds, 2012), which could potentially lead to corporate profits being diverted away from business investment to plugging these pension fund deficits. The BoE found some evidence of this occurring in a survey of companies (see BoE *Agents' Summary of Business Conditions*, June 2013).

## Conclusions

After rapidly cutting short-term interest rates to their effective lower bounds during the financial crisis of 2008–09 and with traditional policy rule models indicating that significant further monetary easing was necessary, central banks in the USA and UK turned to quantitative easing in order to sustain aggregate demand and avoid a Japanese style deflationary spiral. The European Central Bank eschewed large-scale asset purchases, but instituted a number of major lending programmes that significantly increased the size of its balance sheet.

Existing studies suggest that the initial rounds of QE provided a significant boost to both growth and inflation in the USA and UK, whilst the evidence on the impact of the ECB's actions has also been favourable. Another key positive of QE in the USA and UK has been in helping to limit the deterioration in near-term debt dynamics, whilst the ECB's lending programs also indirectly reduced governments' borrowing costs. One area of disappointment, however, was the inability of unconventional policies to generate a material revival in bank lending. Moreover, there has been some evidence of diminishing returns in later rounds of QE.

Meanwhile, fears that QE would fuel a sharp acceleration in inflation have so far proven misplaced, although central banks will need to be increasingly alert to the risks once more sustainable economic recoveries begin to emerge. Policymakers must also pay significant attention to the financial stability risks associated with the long period of extremely accommodative monetary policy.

## Acknowledgements

Unless otherwise indicated, the views in this paper are those of the author and not of Morgan Stanley. The author is grateful to Charles Goodhart for his comments.

***This article is not an offer to buy or sell any security/instruments or to participate in a trading strategy. For important current disclosures that pertain to Morgan Stanley, please refer to the disclosures regarding the issuer(s) that are the subject of this article on Morgan Stanley's disclosure website. https://www.morganstanley.com/researchdisclosures.***

J. ASHWORTH

*See Also* **banking crises; Credit Crunch Chronology: April 2007–September 2009; European Central Bank and monetary policy in the Euro area; euro zone crisis 2010; regulatory responses to the financial crisis: an interim assessment.**

### *Bibliography*

Antolin Diaz, J. 2013. Understanding the ECB's monetary policy. *Fulcrum Research Papers*, Fulcrum Asset Management.

Ashworth, J. and Goodhart, C. A. E. 2011. *QE2 Likely to Set Sail, but it's No Panacea*. Morgan Stanley Research. Available on request from authors.

Ball, L. and Mankiw, N. G. 2002. The NAIRU in theory and practice. *Journal of Economic Perspectives*, 16(4), 115–36.

Bank for International Settlements. 2012. *BIS Annual Report 2011/12*. Available at: http://www.bis.org/publ/arpdf/ar2012e.htm

Bank of England *Agents' Summary of Business Conditions*. June 2013. Available at: http://www.bankofengland.co.uk/publications/Pages/agentssummary/agsum13jun.aspx

Bank of England *Financial Stability Report*. June 2013. Available at:http://www.bankofengland.co.uk/publications/Pages/fsr/2013/fsr33.aspx

Bank of England MPC minutes. March 2009.

Bank of England *Quarterly Bulletin*. 2011 3Q. The United Kingdom's quantitative easing policy: design, operation and impact. Available at: http://www.bankofengland.co.uk/publications/Pages/quarterlybulletin/a11.aspx

Bernanke, B. S. 2000. Japan's slump: a case of self-induced paralysis? Paper presented at American Economic Association meetings; also in *Japan's Financial Crisis and Its Parallels to U.S. Experience* (eds. A. Posen and R. Mikitani). Institute for International Economics.

Bernanke, B. 2009. Reflections on a year of crisis. Speech delivered at *Financial Stability and Macroeconomic Policy*, a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.

Bernanke, B. S. 2012. Monetary policy since the onset of the crisis. Remarks at the *Federal Reserve Board of Kansas City Economic Symposium*, Jackson Hole, Wyoming.

Bernanke, B. S. 2013. Monitoring the financial system. At the *49th Annual Conference on Bank Structure and Competition*, sponsored by the Central Bank of Chicago, Chicago.

Brunner, K. and Meltzer, A. H. 1973. Mr. Hicks and the monetarists. *Economica*, 40, 43–9.

Bomfim, A. N. and Meyer, L. H. 2010. Quantifying the effects of Fed asset purchases on treasury yields. *Monetary Policy Insights: Fixed Income Focus*.

Caglar, E., Chadha, S., Meaning, J., Warren, J. and Waters, A.2011. *Non-conventional Monetary Policies: QE and the DSGE Literature*. Available at: http://www.kent.ac.uk/economics/documents/research/papers/2011/1110.pdf

Chadha, J. S. and Holly, S. 2012. *Interest Rates, Prices and Liquidity: Lessons from the Financial Crisis*. Cambridge University Press, Cambridge.

Chung, H., Laforte, J.-P., Reifschneider, D. and Williams, J. C. 2011. *Have We Underestimated the Likelihood and Severity of Zero Lower Bound Events?* Federal Reserve Bank of San Francisco, WP No. 2011-01. Available at: http://www.frbsf.org/publications/economics/papers/2011/wp11-01bk.pdf

Clouse, J., Henderson, D., Orphanides, A., Small, D. H. and Tinsley, P. A. 2000. Monetary policy when the nominal short-term interest rate is zero. *BE Press Journal of Macroeconomics: Topics in Macroeconomics*, 3(1), article 12.

Cour-Thimann, P. and Winkler, B. 2012. The ECB's non-standard monetary policy measures: the role of institutional factors and financial structure. *Oxford Review of Economic Policy*, 28(4), 765–803.

D'Amico, S. and King, T. B. 2010. *Flow and Stock Effects of Large-scale Treasury Purchases*. Federal Reserve Board, WP No. 2010-52. Available at: http://www.federalreserve.gov/pubs/feds/2010/201052/201052pap.pdf

Dudley, B. 2010. The outlook, policy choices and our mandate. *Remarks at the Society of American Business Editors and Writers Fall Conference*, City University of New York, Graduate School of Journalism, New York City. Available at: http://www.newyorkfed.org/newsevents/speeches/2010/dud101001.html

Friedman, M. and Schwartz, A. J. 1963. *A Monetary History of the United States, 1867–1960*. Princeton University Press, Princeton.

Friedman, M. and Schwartz, A. J. 1982. *Monetary Trends in the United States and the United Kingdom: Their relation to Income, Prices and Interest Rates, 1867–1975*. University of Chicago Press, Chicago.

Gagnon, J., Raskin, M., Remache, J. and Sack, B. 2011. The financial market effects of the Federal Reserve's large scale asset purchases. *International Journal of Central Banking*, 7(1), 3–43.

Giannone, G., Lenza, M., Pill, H. and Reichlin, L. 2011. Non-standard monetary policy measures and monetary developments. In: *Interest Rates, Prices and Liquidity: Lessons from the Financial Crisis* (eds. J. S. Chadha and S. Holly). Cambridge University Press, Cambridge.

Giannone, G., Lenza, M., Pill, H. and Reichlin, L. 2012. The ECB and the interbank market. *Discussion Paper No. 8844*, CEPR, February.

Goodhart, C. A. E. 2010. Money, credit and bank behaviour: need for a new approach. *National Institute Economic Review*, 2010 214: F73.

Goodhart, C. A. E. and Ashworth, J. P. 2012. QE: a successful start may be running into diminishing returns. *Oxford Review of Economic Policy*, 28(4), 640–70.

Joyce, M., Tong, M. and Woods, R. 2011. The United Kingdom's quantitative easing policy: design, operation and impact. *Bank of England Quarterly Bulletin*, 51(3), 200–12. Available at: http://www.bankofengland.co.uk/publications/Pages/quarterlybulletin/a11.aspx

Kapetanios, G., Mumtaz, H., Stevens, I. and Theodoridis, K. 2012. Assessing the economy-wide effects of quantitative easing. *Bank of England WP No.443*. Available at: http://www.bankofengland.co.uk/publications/Documents/workingpapers/wp443.pdf

Keynes, J. M. 1936. *The General Theory of Employment, Interest and Money* (1964 reprint). Harcourt, Brace, New York.

Krishnamurthy, A. and Vissing-Jorgensen, A. 2011. The effects of quantitative easing on interest rates: channels and implications for policy. *NBER WP No. 17555*. Available at: http://www.nber.org/papers/w17555.ack

Lenza, M., Pill, H. and Reichlin, L. 2010. Monetary policy in exceptional times. *Economic Policy*, 62, 295–339.

Miles, D. 2012. Winding and unwinding extraordinary monetary policy. Speech at *RBS Scottish Economic Society Annual Lecture*, Edinburgh. Available at: http://www.bankofengland.co.uk/publications/Pages/speeches/2012/598.aspx

Modigliani, F. and Sutch, R. 1966. Innovations in interest rate policy. *American Economic Review*, 56, 178–97.

National Association of Pension Funds. 2012. *Exceptional Times, Exceptional Measures? Economic Developments and the Impact on Pension Schemes and Members*. Available at: http://www.napf.co.uk/PolicyandResearch/DocumentLibrary/0221_Exceptional_times_exceptional_measures_economic_developments_March_2012.aspx

Reinhart, C. and Rogoff, K. 2009. *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton.

Shirakawa, M. 2012. Central banking: before, during, and after the crisis. Remarks at a *Conference Sponsored by the Federal Reserve Board and the International Journal of Central Banking*. Available at: http://www.bis.org/review/r120329b.pdf

Stein, J. 2013. Overheating in credit markets: origins, measurement, and policy responses. Remarks at *Restoring Household Financial Stability after the Great Recession: Why Household Balance Sheets Matter*, a research symposium sponsored by the Federal Reserve Bank of St Louis. Available at: http://www.federalreserve.gov/newsevents/speech/stein20130207a.htm

Swanson, E. T. 2011. Let's twist again: a high-frequency event-study analysis of Operation Twist and its implications for QE2. *Brookings Papers on Economic Activity*, Spring, 151–88.

Taylor, J. 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39, 194–214.

Turner, A. 2013. Debt, money and mephistopheles: how do we get out of this mess? Speech at Cass Business School, London.

Woodford, M. 2012. Methods of Policy Accommodation at the Interest Rate Lower Bound. Paper presented to the Jackson Hole Symposium, August/September.

Wadhwani, S. 2013. The great stagnation: what can policymakers do? Speech at the *Annual Peston Lecture*, London.

Yellen, J. 2011. The Federal Reserve's asset purchase program. Speech at the *Brimmer Policy Forum, Allied Social Science Associations Annual Meeting*, Denver, Colorado, 8 January.

# Run on Northern Rock, the

## Introduction

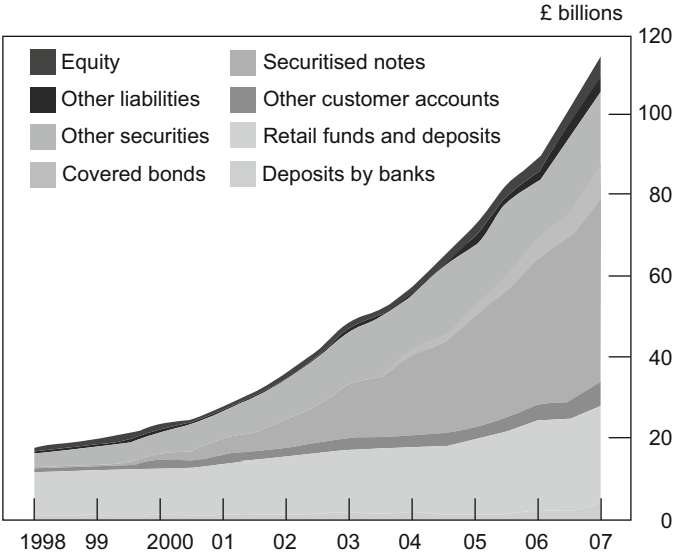
On the morning of Friday 14 September 2007, queues of depositors began to form inside, and then outside, the (relatively) few branches of Northern Rock (only nine in the London area, for example). This was the first substantial run in the UK by retail depositors since the 19th century. Northern Rock had been a building society until 1997, with a large local presence in the north-east (headquartered in Gosforth, Newcastle upon Tyne), but otherwise then not widely known and subject to relatively strict Building Society requirements. In that year it demutualized, became a bank and later embarked on a massive program of expansion, under its incoming Chief Executive, Adam Applegarth.

Northern Rock, however, stuck to its traditional area of expertise, financing household mortgages, after demutualization; but it improved its terms to borrowers, on relative spreads and down-payment, in order to take an ever larger share of the domestic mortgage market. The rate of expansion of its lending then vastly exceeded the rate of growth of its retail deposits, leading to a sharp rise in its loan to deposit ratio. It financed this gap in three main ways. First, having originated such loans, it pooled them together and transferred these to a securitization vehicle, a special-purpose entity termed Granite. Second, it issued covered bonds, secured against both the mortgages and the bank. Thus, if due payments on the bond were not made, such creditors were covered by having the right to seize, and then perhaps sell, the underlying mortgages. If that, in turn, was insufficient to repay them, they had a further claim against the assets of the bank, in this case Northern Rock. Finally, it borrowed in wholesale markets (see Figure 1).

> 'Mr Applegarth outlined the overall funding of Northern Rock:
>
> > 50% was securitisation, which had an average life of three and a half years; 10% was covered bonds, which had an average life of about seven years; and of our wholesale borrowings, which is 25%, half of that had a duration longer than one year and the other half was less than one year's duration.' House of Commons Treasury Committee Report on 'The Run on the Rock', p. 13.

So Northern Rock was largely reliant on continued access to (short-term) wholesale markets for funding, especially during the intervals in which it was originating and warehousing mortgages before transferring them in a pool as an additional tranche to its associated SPV, Granite. It was due to launch an exceptionally big securitization in August/September, and therefore was unusually heavily reliant on short-term money-market funding when the crisis broke. When the wholesale funding markets, especially but not only in Europe, began to break down on 9 August 2007, Northern Rock soon realized that it was in trouble, thus:

> 'The then Chairman and Chief Executive of Northern Rock first discussed these problems with each other on Friday 10 August.' (Treasury Select Committee, *ibid*, p. 35.)




*Sources*: Northern Rock Interim Annual Reports; Bank of England *Financial Stability Report*, October, 2007, p. 10.

**Figure 1** Northern Rock: balance sheet growth and liability structure – June 1998–June 2007. *Source*: Bank of England *Financial Stability Report*, October, 2007, p. 10.

For a number of reasons, detailed in the next section, none of the proposed remedies for Northern Rock's illiquidity proved feasible. One factor in August was the uncertainty over how long the securitization markets would be closed. It was hoped (not just by Northern Rock) that this would be temporary, and Northern Rock and its advisers were still hoping to launch an issue in early September. Its illiquidity worsened rapidly from chronic to critical, especially since the transfer of a further tranche of mortgages to Granite could not possibly succeed in those circumstances. So the stark alternatives soon became to allow Northern Rock to go bankrupt or to provide it with a massive loan from the Bank of England.

For a variety of reasons, detailed in the next section, it was felt that the existence, occasion and extent of such a loan would have to be publicly revealed. The loan itself was finally settled on the evening of Thursday 13 September. At a conference dinner at the Bank that evening, senior Bank officials kept on mysteriously getting up and leaving and then returning; so it was obvious something was occurring, but none of the others present knew what.

The idea was to announce the loan early on Monday 17 September. The announcement of the loan could then have had two effects on the retail depositors (and general public). The first could have been, 'The Bank of England is now in support; we can stop worrying', but the second was 'We did not realize how bad the situation was. It must be bad if the Bank has to help so much'.

The premise on which the deposit insurance scheme in the UK had been established was that a retail deposit run was unthinkable in the UK. So the insurance was aimed to

limit moral hazard, by being a full 100% only up to £2000, partial from there on (90%) up to a cap of £35,000, and none thereafter. The aim of the scheme was to make it (politically) easier to allow bank failures (by bailing out 'widows and orphans'), *not* to stop runs. So, naturally, in so far as retail depositors realized that they had any insurance at all (doubtful), it was often still worth their while to run, so long as there was any significant probability of Northern Rock failing. Moreover the time lag before *any* such insurance payment might be received could be weeks, if not months.

If the announcement of Bank of England support for Northern Rock had been made as planned on Monday morning, it would have been accompanied by a battery of anodyne Press comment, from the Bank and Treasury, about 'no need to worry', 'everything in hand', etc. But, even before the final terms of the loan were settled on the Thursday evening, Robert Peston of the BBC reported at 8:30 p.m. that Northern Rock had asked for and would receive emergency financial support from the Bank of England. Naturally, given the juicy nature of this leak, it was not in the interest of the media to play this down, either immediately or subsequently in the next few days. We shall probably never know how and why this leak occurred, though there are a range of (unsubstantiated) rumours about this.

Be that as it may, nothing had been made ready for the announcement. Not only were Press departments unprepared, but Northern Rock was unable to draft in additional bank tellers or expand its website to meet the expected flood of withdrawals and urgent enquiries by worried depositors. So the website crashed, and Internet depositors may have feared that such a shutdown was deliberate, to block withdrawals, and many then went to do so at the branches. The limited branch staff, in their often small branch offices, could not process withdrawals quickly enough to prevent queues snaking out of the door and around the block. These were, of course, captured on television, which served to feed the run.

In the face of this unforeseen event, the authorities seemed paralyzed on Friday and over the weekend, with no effective response. It was not until Monday that the Chancellor of the Exchequer, Alistair Darling, appointed to the position as recently as late June, announced that the government would provide a full guarantee for all Northern Rock deposits. This soon led to queries as to whether the government could discriminate in this way on behalf of just one bank, and it soon became accepted that the government would soon have to legislate, *de jure*, to provide 100% deposit insurance (at least up to some upper limit) for all bank deposits, and that it was already in that position, *de facto*. These measures did stop the run, and prevented immediate contagion, for example to Alliance and Leicester, and Bradford & Bingley.

Clearly, the Northern Rock episode forced a change in the UK's deposit insurance regime. But the effects went much further. The event underlined how limited were the options that the authorities had to hand for dealing with a failing bank, and underscored the need for a new Special Resolution Law for financial institutions, later passed in 2009. The need for such a Special Resolution Regime had been perceived beforehand, but not pursued with sufficient vigour. If such an SRR had been in place, the episode could have been handled much more smoothly. While the main blame for the collapse fell on the Board and management of Northern Rock for having embarked on such a risky business plan, questions were asked as to how the supervisor, the Financial

Services Authority (FSA), had allowed it to get that way. The resulting study, by the internal audit division of the FSA, 'The supervision of Northern Rock: a lessons learned review', (March 2008), was open and candid, and exposed a litany of shortcomings. The reputation not only of the FSA, but also of its vaunted, 'Principles based' and 'Light touch' approach, was damaged. Meanwhile, the failure of the authorities, working together in the Tripartite Committee of FSA, Bank and Treasury, to come up with a better solution before the run occurred, led to queries about 'Who was in charge?', and whether the whole UK regulatory and supervisory structure needed reconsideration (as did occur later when the Conservative/Liberal coalition came into power in May 2010).

The fallout from the Northern Rock run extended yet further into the political orbit. It formed part of the background to the decision by Gordon Brown, who had become Prime Minister in June, following the resignation of Tony Blair in the aftermath of the Iraq affair, *not* to hold a general election in October 2007 to reinforce his political position.

So the run on Northern Rock had important ramifications. In the next section we will discuss in more detail what led up to it, and particularly why it was not prevented. Then we cover what has happened since, and we finish with some brief conclusions.

## The antecedents of the run

With the inestimable benefit of hindsight, there is no question that Northern Rock had embarked on an extremely risky business strategy. The Treasury Committee Report described that strategy as being 'high-risk, reckless'. So why did no one stop them? Let us consider the following potential gatekeepers: (1) the Board, (2) shareholders, and (3) the supervisor (FSA).

The reason why the Board was content with its strategy is set out in the Treasury Committee Report, pp. 15/16:

> 'Two aspects of this worldwide liquidity squeeze appeared to surprise Northern Rock, and overcome the attempts highlighted above to combat the tightening in credit markets. One was the absence of a so-called "flight to quality". Dr Ridley told us that:
>
> > What we did not expect was that there would be no flight to quality in that process [of a tightening in credit markets]. In other words, we expected that as markets became tighter and as pricing for risk changed that low-risk prime UK mortgages (and we have below half the industry average of arrears on our mortgage book) and such a low-risk book would remain easier to fund than sub-prime mortgages elsewhere. That is why we were very determined to keep the credit quality of our book high, in order to be able to attract funding.
>
> 'Mr Applegarth told us that Northern Rock had wrongly "believed that high-quality assets and transparency [were] the way to maintain liquidity". Sir Derek Wanless [Chair of the Risk Committee of the Board] told us that Northern Rock's "first line of defence [was] good credit quality".

'Secondly, Northern Rock had not foreseen all its funding markets closing simultaneously, as happened after 9 August. Dr Ridley explained:

> We deliberately diversified our funding platform so that we would have ... three different types of funding and indeed a diversified programme within the wholesale funding, and geographically we had programmes in the United States, Europe, the Far East, Canada and Australia. That was deliberately so that if one market closed we would still have access to others. The idea that all markets would close simultaneously was unforeseen by any major authority.

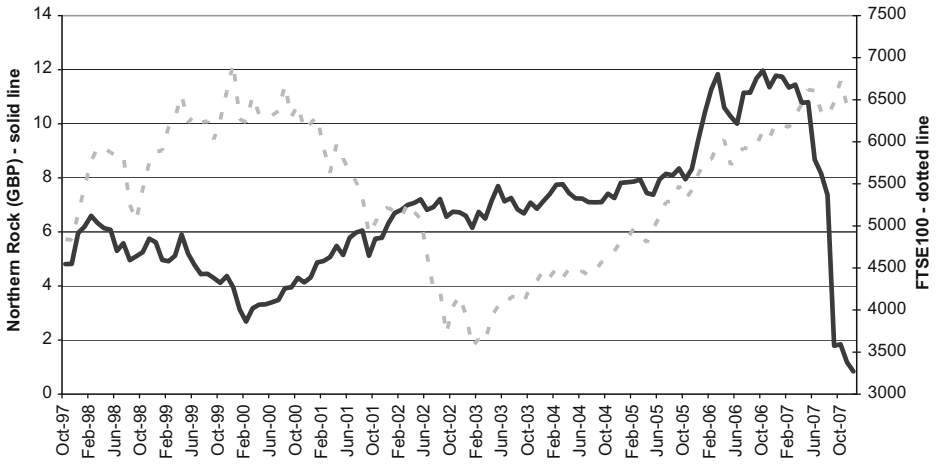
'The idea of all markets closing to Northern Rock was repeatedly characterised to us by Northern Rock officials as "unforeseeable".'

Thus there was a failure to realize that, in really stressed conditions, correlations go to unity and previously acceptable diversification ceases to be effective.

It would generally be an error to expect much risk control from shareholders. Given their limited liability, they often favour and reward risk-taking. Indeed, Northern Rock, with its expansionary policy and risky strategy, had been a favourite of the London Stock Exchange. It was not until relatively late in the day that its share value declined sharply (Figure 2).

So, finally, we come to the supervisor (FSA). Here there is considerable information from the FSA's published internal audit on 'lessons learned' (March 2008). First, there was a particular and unusual degree of maladministration in the case of Northern Rock, as noted in Figure 3 taken from page 9 of the FSA's audit.

Even more important was that the FSA was excessively focused on the implementation of the Basel II recommendations at this juncture, so much so that other aspects and dimensions of risk were largely ignored, such as excessive growth, very high leverage and poor liquidity. Largely because mortgage lending had a low Basel II risk weight,




*Source*: Compustat Global and Yahoo Finance

**Figure 2** Northern Rock share price and FTSE index, 1997–2007 (monthly).

*Northern Rock is in the highlighted box in each population*

**Supervisory Period** – data taken from IRM

| Firms with regulatory periods of: | | |
|---|---|---|
| 18–24 months | 24 | 63% |
| 25–30 | 10 | 26% |
| 36 months | 4 | 11% |

(As at 1 August 2007, high impact firms only: MRGD-26 and WIBD-12)

**Turnover of HoDs experienced by MRGD firms**

| Number of HoDs (1 Jan 05 – 9 Aug 07) | 1 HoD | 2 HoDs | 3 HoDs |
|---|---|---|---|
| Number of high impact firms | 18 | 6 | 2 |
| | 69% | 23% | 8% |

**C&C meetings** – estimates made by staff based on various sources for high impact firms

| | 1 Jan – 9 Aug 2007 | 2006 | 2005 | Period Total |
|---|---|---|---|---|
| Average for WIBD firms | 13 | 24 | 18 | 55 |
| Average for MRGD firms | 22 | 29 | 23 | 74 |
| Average for MRGD excl 5 largest banks | 17 | 22 | 19 | 58 |
| Average for 5 largest retail banks | 43 | 59 | 41 | 143 |
| Northern Rock | 7* | 1 | 0 | 8 |

* Out of which five meetings were held on one day and two were by telephone

**Risk mitigation programmes (RMP)** – data taken from IRM

| Number of firms with RMP | 37 |
|---|---|
| Number of firms without RMP | 1 |

*Source*: The Financial Services Authority (FSA) Internal Audit Division. The supervision of Northern Rock: a lessons learned review, March 2008, p. 9.

**Figure 3** The supervision of Northern Rock: lessons learned.

and the securitization of mortgages into Granite lowered the risk weighting even further, Northern Rock passed its Basel II test with flying colours in June 2007, at a time when, by US leveraged standards, it would have been considered 'critically undercapitalized', with a core equity to total assets ratio of under 2% (over 50 to 1!). This is a leading example of the folly of putting all one's supervisory concerns into the single Basel II basket.

As the Treasury Committee Report notes (p. 25),

> 'The adoption of an advanced approach requires a waiver from the Financial Services Authority. On 29 June 2007, Northern Rock was told by the FSA that its application for a Basel II waiver had been approved.
>
> Due to this approval, Northern Rock felt able to announce on 25 July 2007 an increase in its interim dividend of 30.3%. This was because the waiver and other asset realizations meant that Northern Rock had an "anticipated regulatory capital surplus over the next 3 to 4 years". Mr Applegarth explained how Northern Rock had achieved this waiver. The company had come to the end of a two and a half year process, during which period Northern Rock had

undergone several stress tests, a matter we consider further later in this chapter. As well as this, in order to obtain a Basel II waiver Northern Rock had to "show that [Northern Rock could] dynamically manage scorecards from new lending all the way through to arrears and possessions and put that information back into [Northern Rock's] front end score cards". Mr Applegarth explained that the waiver had led to a dividend increase because:

> when you get your Basel II approval, the relative risk weighting of certain assets in your balance sheet changes. So what we had, because of the quality of the loan book, was you saw our risk weighting for residential mortgages come down from 50% to 15%. That clearly required less capital behind it, so that links to why we were able to increase the dividend.'

Hardly more than a month afterwards, on 9 August 2007, the financial crisis began, and wholesale markets began to dry up. A generalized shortage of liquidity ensued, with Northern Rock particularly at risk. In previous decades, at least until the 1970s, banks might have weathered this by selling, or borrowing against the collateral of, their government debt holdings. But banks, including Northern Rock, had shifted from asset liquidity (government debt) to cheaper funding liquidity from wholesale markets. They had no government debt holdings left to fall back upon. Under these circumstances,

> 'In August 2007, the Bank of England was approached by banks arguing that the Bank of England should provide additional liquidity, at no penalty rate. The FSA had transmitted the banks' request to the Bank of England, but refused to state to us whether it had supported the banks in requesting this additional liquidity, on the grounds that conversations between Tripartite members ought to remain private. On 12 September 2007, in advance of his oral evidence on 20 September, the Governor of the Bank of England wrote a letter to the Chairman of this Committee. In that letter, the Governor pointed out that he did not agree with the suggestions for additional measures that others believed the Bank of England should undertake: lending at longer maturities, removing the penalty rate or increasing the range of collateral against which the Bank would be prepared to lend. In the letter, he gave three reasons for his position. First, he stated that "the banking system as a whole is strong enough to withstand the impact of taking onto the balance sheet the assets of conduits and other vehicles". Second, "the private sector will gradually re-establish valuations of most asset backed securities, thus allowing liquidity in those markets to build up". Third, there would be a risk of "moral hazard". In essence, this "moral hazard" argument is that, should the central bank act, and effectively provide extra liquidity at different maturities against weaker collateral, markets would, especially if the liquidity were provided at little or no penalty, take it as a signal that the central bank would always rescue them should they take excessive risk and get into difficulties. Such a signal would lead to ever more risk taking, and the next crisis would consequently be greater than it would otherwise have been.' (Treasury Committee Report, pp. 38/39.)

The Governor's position was clearly contentious. Moreover, the ECB, though for historical reasons rather than from taking a different stance on moral hazard, was

prepared to lend against a wider range of assets. It may be that, had Northern Rock been within the euro zone, it could have survived this first storm. But now, with further securitization into Granite now impossible, and liquidity running out, Northern Rock had either to sell itself to another (bigger) bank, or apply for a large, formal support facility from the Bank of England.

One major high street retail bank, widely believed to be Lloyds, showed considerable interest, but the bank, having discovered the Rock's liquidity problems, wanted a loan from the Bank. 'The Chancellor of the Exchequer stated clearly that the financial support requested was in the form of a loan, which "could have been as much as £30 billion... to be given at commercial rates by the Bank of England"' (Treasury Committee Report, p. 51). There were two main problems with this. First, if such favourable terms were going to be offered to one bank, they would have to be offered to all banks, which would have had to be done publicly and would have taken a long time. Second, would it be consistent with the European Union prohibition on state aid to commercial banks? So such a liquidity facility was denied, and that ended any prior hope of merging Northern Rock with a stronger partner. In any case, such a merger would have had to be agreed by the shareholders, and that would have meant exposing their liquidity problems.

It is not clear whether liquidation was ever seriously considered. If Northern Rock had been allowed to fail, there would almost certainly have been immediate contagion to the other weaker mortgage lenders in the UK, notably Bradford & Bingley. In any case,

> 'It appears that a decision in principle that Northern Rock would be granted a support facility should neither securitisation or a takeover prove possible was taken at a meeting between the Chancellor of the Exchequer, the Chairman of the FSA and the Governor of Bank of England on Monday 3 September. The final decision was that of the Chancellor of the Exchequer, but his decision was taken on the basis of a joint recommendation of the Governor of the Bank of England and the Chairman of the FSA' (Treasury Committee Report, p. 54).

There was then much – quite confused – discussion whether such emergency lending to Northern Rock would have to be announced or could be kept covert. The problem was that this support clearly involved a material change in the financial condition of Northern Rock, and the question was whether European and/or British rules on market disclosure required that such a change be publicly announced. The Board of Northern Rock were legally advised that such an announcement was needed. So it was arranged for the early morning of Monday 17 September, but, as already described, this was overtaken by the leak to the BBC on the evening of Thursday 13 September.

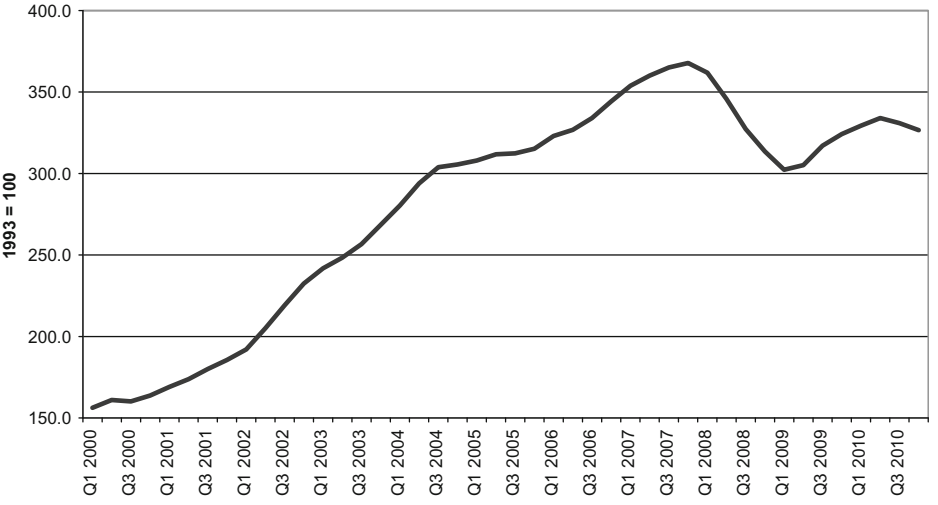
## The subsequent history

Throughout this whole episode, the management of Northern Rock, and all the regulatory authorities (the Bank of England, FSA and Treasury), were adamant in their claims that the Rock's mortgage assets were of 'high quality'. They no doubt

were, in the sense that all the proper procedures and paperwork had been gone through in the correct manner, with none of the fiddles and deficiencies that had besmirched the US sub-prime market. But Northern Rock had concentrated on UK mortgages and was expanding very rapidly (averaging 20% p.a.) at the top of a huge housing boom. Moreover, as noted in the FSA's internal report on 'Lessons Learned', p. 38,

> 'A key element of Northern Rock's product range was the "Together" product, which represented 31% of gross mortgage lending by the firm in 2006 and 26% in the first six months of 2007. The key features of Together were that it offered a secured loan of up to 95% loan to value (LTV) coupled with an unsecured loan of up to 30% LTV all at a single rate and serviced by one monthly payment. Together was essentially targeted at first-time buyers and the unsecured element was designed to finance associated house-purchase or home-making costs. The scope for a loan of up to 125% LTV meant the product was regarded by many as particularly high-risk.'

It was inevitable that in a sharp bust in the housing market (one such was clearly on its way by the latter half of 2007 – Figure 4) a sizeable proportion of the Rock's mortgage borrowers would move into negative equity. Although the legal arrangements (mortgages were recourse, not non-recourse loans) and UK social culture meant that arrears of payment, default and foreclosure would still be far less than in the USA, the forward-looking expectation by end-2007 must have been that not only was Northern Rock illiquid, but that it would very likely also become insolvent; even though in September 2007, on a backwards-looking basis, it was solvent (with relatively few arrears comparatively – not such a difficult achievement when house prices had been




*Source*: The Nationwide House Price Index (http://www.nationwide.co.uk/hpi/historical.htm).

**Figure 4** UK house prices, seasonally adjusted, 2000–2010. *Source*: The Nationwide House Price Index (http://www.nationwide.co.uk/hpi/historical.htm).

rising so rapidly. Also Northern Rock was, apparently unusually quick to foreclose, which thereby kept arrears low.)

Be that as it may, the announcement by the Chancellor, Alistair Darling, on Monday 17 September that all existing deposits at Northern Rock would be guaranteed (new deposits there also became fully guaranteed on 9 October) stopped the run; but this was not until some £4–5 billion more had been withdrawn, raising the necessary liquidity assistance from the Bank of England to over £25 bn. Meanwhile, the government guarantee covered an additional £30 bn. So, having rescued Northern Rock depositors at taxpayers' expense, the main concern now was to limit that bill.

Besides its mortgage assets, whose potential future value depended on an increasingly fragile housing market, Northern Rock had an existing infrastructure of branches, staff and systems that could give a new entrant access to the (oligopolistic) UK retail banking market. In pursuit of such an entry a wide group of institutions expressed some interest in acquiring Northern Rock, putative names apparently including Virgin Group, Oliphant, Cerberus, J.C. Flowers, Lloyds (again), Lehman Bros, Bradford & Bingley and Tesco. By the time the deadline for such bids occurred (4 February 2008), only two remained on the table, from Virgin and an in-house bid. After a study of the options, the Treasury, assisted by Goldman Sachs, decided that (temporary) public ownership would be cheaper for the taxpayer than either bid, so on 17 February 2008, the Chancellor announced the nationalization of Northern Rock.

In the meantime, not surprisingly, most of the prior top management and Board had resigned and been replaced. When Northern Rock was nationalized on 20 February, the question arose as to what compensation, if any, should be paid to the shareholders, an issue driven forward by a couple of hedge funds who had bet on a recovery. The authorities decreed that the appropriate amount for compensation should be the likely valuation of Northern Rock *absent any official support*, which was, of course, as an independent valuer determined, precisely zero, which was accepted in all the law courts who have so far tried the case.

Since nationalization, Northern Rock has split into two parts at the outset of 2010: the good part, Northern Rock plc, and the bad part, Northern Rock (Asset Management) plc, whose role is to manage the run-down of the worst assets (it was later merged with Bradford & Bingley plc into a single holding company, UK Asset Resolution UK). The good part, Northern Rock plc, continues, with conflicting pressures to expand its loans (at a time when other banks are not doing so) on the one hand and to reduce its book to allow the government to exit nationalized banking on the other hand. In February 2010 the government removed the 100% guarantee for Northern Rock depositors, leaving such depositors in the same position as depositors in other UK banks, with the first £50,000 fully guaranteed. Meanwhile some of the Bank of England's support loan, transferred to HMT in the summer of 2008, has been repaid, (the June 2010 half-yearly balance sheet shows some £22.5 billion of such debt to HMT still outstanding), and there are intermittent rumours of private sector interest in buying Northern Rock plc back from the government. On 18 January 2011 it was announced in the Press that the public sector owner UK Financial Investments (UKFI) was asking investment banks and other possible advisers to apply to assist with funding a private sector buyer.

## Conclusions

The story of the run on Northern Rock is an unhappy one. The management of Northern Rock adopted an excessively expansionary and risky business plan, and neither the Board nor their supervisor, the FSA, checked them. This was partly due to illusion and delusion, on the part of management, that they could always fund in wholesale markets and that the UK housing market would not collapse; and on the part of the FSA that full implementation of Basel II was a sufficient guarantee that all would be well. Its downfall damaged a lot of reputations, not least that of 'principles based', 'light touch' UK regulation.

CHARLES A. E. GOODHART

*See Also* **Bank of England; credit crunch chronology; eurozone crisis 2010; foreclosures.**

### *Bibliography*

There are three main public sector reports. The best is that by The Treasury Committee of the House of Commons: *The Run on the Rock*, Fifth Report of Session 2007–8, Vol. 1, HC56-1 (26/biblScope> January 2008).

The FSA Internal Audit Division then examined *The Supervision of Northern Rock: A Lessons Learned Review* in their Report of March 2008.

These reports only covered events leading up to and surrounding the run itself. For an account of the aftermath, the National Audit Office (NAO) report on *HM Treasury: The Nationalisation of Northern Rock*, 20 March 2009 (www.nao.org.uk/publications/0809/northern.rock.aspx), should be consulted.

For further details of events since the run, this article has relied on Wikipedia, notably *Nationalisation of Northern Rock* (http://en.wikipedia.org/wiki/ Nationalisation_of_Northern_Rock), as modified on 2 January 2011, and other Press sources, notably *Timeline: the Northern Rock Crisis* by the *Guardian* (http://www.guardian.co.uk/business/2008/mar/26/northernrock.

The most comprehensive academic account of the event is Milne, A. and Wood, G. 2009. Shattered on the Rock? British financial stability from 1866 to 2007. *Journal of Financial Regulation*, 10(2), 89–127. Geoffrey Wood both gave evidence to, and was the adviser of, the Treasury Committee for this Report.

For a more general study of the circumstances and implications of the event, see Bruni, F. and Llewellyn, D. (eds.) 2009. *The Failure of Northern Rock: A Multi-dimensional Case Study*. SUERF – The European Money and Finance Forum, Vienna.

# shadow banking: a review of the literature

## 1) What Is Shadow Credit Intermediation?

The shadow banking system is a web of specialised financial institutions that channel funding from savers to investors through a range of securitisation and secured funding techniques. Although shadow banks—the institutions that constitute the shadow banking system—conduct credit and maturity transformation similar to that of traditional banks, they do so without the direct and explicit public sources of liquidity and tail risk insurance available through the Federal Reserve's discount window and the Federal Deposit Insurance Corporation. Shadow banks are therefore inherently fragile, not unlike the commercial banking system prior to the creation of the public safety net. This definition closely follows that of Pozsar *et al.,* (2010).

*A) Definition*

In the traditional banking system, intermediation between savers and borrowers occurs in a single institution. Through the process of funding loans with deposits, banks engage in credit, maturity, and liquidity transformation. Credit transformation refers to the enhancement of the credit quality of debt issued by the intermediary through the use of priority of claims. For example, the credit quality of senior deposits is better than the credit quality of the underlying loan portfolio due to the presence of junior equity. Maturity transformation refers to the use of short-term deposits to fund long-term loans, which creates liquidity for the saver but exposes the intermediary to roll-over and duration risks. Liquidity transformation refers to the use of liquid instruments to fund illiquid assets. For example, a pool of illiquid whole loans might trade at a lower price than a liquid-rated security secured by the same loan pool, as certification by a credible rating agency would reduce information asymmetries between borrowers and savers.

Savers entrust their funds to banks in the form of deposits, which banks use to fund loans to borrowers. Savers furthermore own the equity and long-term debt issuance of the banks. Deposits are guaranteed by the FDIC, and a liquidity backstop is provided by the Federal Reserve's discount window. Relative to direct lending (that is, savers lending directly to borrowers), credit intermediation provides savers with information and risk economies of scale by reducing the costs involved in screening and monitoring borrowers and by facilitating investments in a more diverse loan portfolio.

Adrian, Ashcraft: Federal Reserve Bank of New York (e-mail: tobias.adrian@ny.frb.org, adam.ashcraft@ny.frb.org). This paper was prepared for the *New Palgrave Dictionary of Economics*. The authors thank Nicola Cetorelli and Andrei Shleifer for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

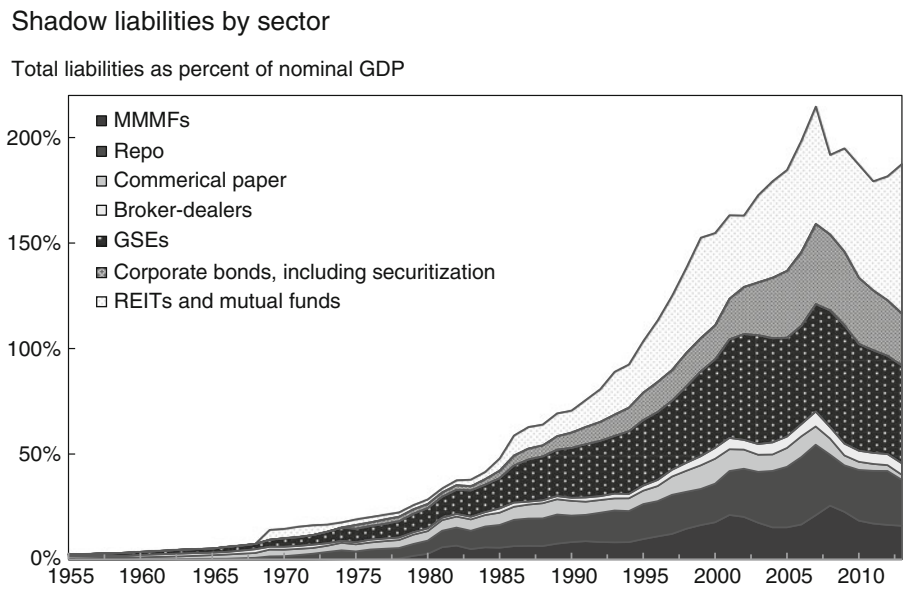
Shadow banking activity is removed from official public-sector enhancements, but typically receives indirect or implicit enhancements. Official enhancements to credit intermediation can be classified into four levels of strength:

1. A liability with direct official enhancement must reside on a financial institution's balance sheet, whereas off-balance-sheet liabilities of financial institutions are indirectly enhanced by the public sector.
2. Activities with direct and implicit official enhancement include debt issued or guaranteed by the government-sponsored enterprises (GSEs), which benefit from an implicit credit put to the taxpayer. The implicit nature of support implies that the intermediary receives the benefit of credit and put options to the public sector, but typically would not pay their full marginal social cost. It is not surprising that, with such a subsidy, these intermediaries would grow very large.
3. Activities with indirect official enhancement generally include the off-balance-sheet activities of depository institutions, such as unfunded credit card loan commitments and lines of credit to conduits. The focus here is on the accounting and capital arbitrage activities by financial institutions. Capital requirements have typically been tied to accounting rules, so transactions to remove assets from the balance sheet have historically reduced regulatory capital. While recent accounting reform has reduced the scope for this form of arbitrage going forward, it was an important part of the narrative of the recent credit cycle.
4. Finally, activities with indirect and implicit official enhancements include asset management activities such as bank-affiliated hedge funds and money market mutual funds (MMMFs), as well as the securities lending activities of custodian banks.

Credit intermediation activities that take place without official credit enhancements are said to be unenhanced. For example, the securities lending activities of insurance companies, pension funds, and certain asset managers do not benefit from access to official liquidity. We define shadow credit intermediation to include all credit intermediation activities that are implicitly enhanced, indirectly enhanced, or unenhanced by official guarantees established on an *ex ante* basis.

*B) Measurement*

To illustrate how shadow credit intermediation has evolved over the past few decades, Figure 1 presents the liabilities of financial businesses in the shadow sector, derived from U.S. Flow of Funds data. In particular, it documents the liabilities relative to GDP for each of money market mutual funds, repurchase agreements, commercial paper, broker-dealers, Government-Sponsored Enterprises (GSEs), corporate bonds (including securitization), as well as real estate investment trusts (REITs) and mutual funds. While the level of shadow of liabilities relative to GDP was negligible until the mid-1960s, it had increased to a peak of 215 percent of GDP in 2007 before collapsing to as low as 179 percent following the recent financial crisis. The main driver of the increased importance of shadow liabilities is the growth in REITs and mutual funds. As of this writing, these types have combined liabilities equal to 71 percent of GDP, which interestingly is up from a pre-crisis peak of 55 percent. The second driver is growth in the liabilities of GSEs, which remains in the historically high range at 46 percent. While




Note: MMMFs refers to money market mutual funds; GSEs refers to government sponsored enterprises, and REITs refers to real estate investment trusts.

Figure 1: **Shadow liabilities by sector. Source: Federal Reserve Financial Accounts of the United States.**

Money Market Mutual Funds, Repo and Commercial Paper combined increased to as high as 63 percent in 2007, they have declined significantly and remain at 40 percent.

Figure 2 breaks out the liabilities of the financial business into one of four major categories: 1) traditional maturity transformation, including bank deposits and interbank liabilities; 2) traditional credit transformation including term debt issued by banks and bank holding companies as well as reserves of pensions and life insurance companies, in addition to depository loans not elsewhere classified; 3) shadow maturity transformation, including MMMFs, repo, open market paper, and security broker- dealer credit and payables; and 4) shadow credit transformation, including GSEs, term debt issued by nonbanks, mutual fund shares, REIT mortgage debt, and loans categorised as "other." The figure suggests several striking patterns. First, the amount of traditional maturity transformation relative to GDP in the financial system has been fairly stable throughout the period, with a post-crisis increase related to extraordinary actions by the central bank. At the same time, the relative importance of traditional maturity transformation in funding credit has been significantly diminished over time. Banks owned half of all financial liabilities in 1960, but now have less than 15 percent.

Second, the reduced importance of bank maturity transformation is being offset partly by shadow maturity transformation. The consequence is that the fraction of the

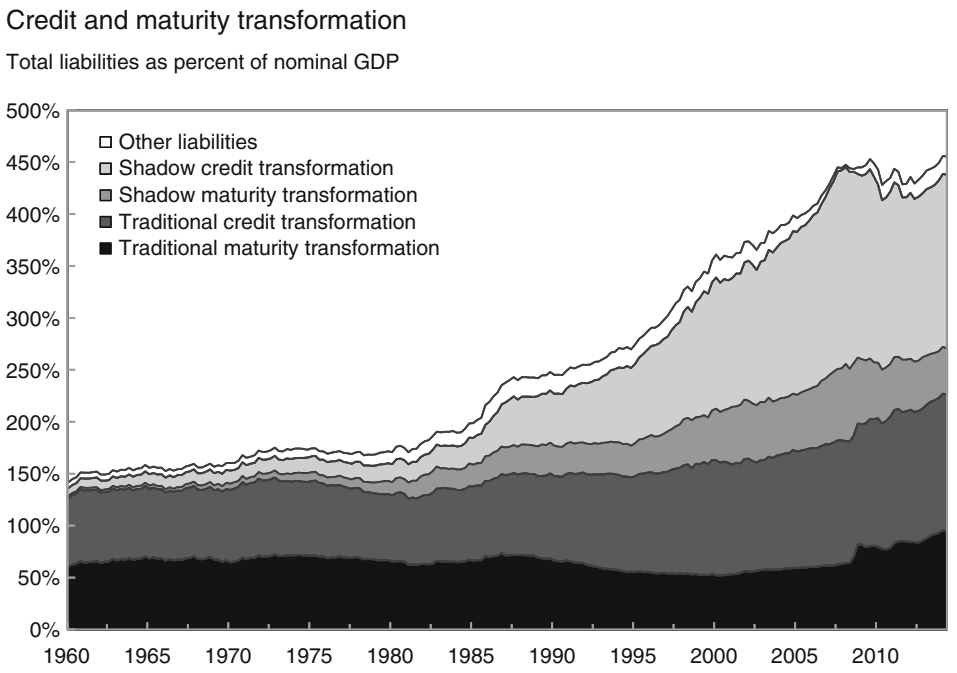



Figure 2: **Credit and maturity transformation. Source: Federal Reserve Financial Accounts of the United States.** Traditional maturity transformation includes net interbank liabilities (line 28) plus checkable (line 29) and savings (line 30) deposits of depository institutions. Traditional credit transformation includes reserves of life insurance companies (line 43) and pensions (line 44) plus corporate debt issued by banks and holding companies plus loans from depository institutions NEC (line 37). The latter is calculated by subtracting from total corporate debt (line 36) the amount issued by holding companies (line 10) and banks (line 5) from L212. Shadow maturity transformation includes from L107 MMMFs (line 31), repo (line 32), commercial paper (line 34), and security broker-dealer credit (line 41) and payables (line 42). Shadow credit transformation includes GSEs (line 35), REITs (line 39), mutual fund shares (line 40), and other loans (line 38).

aggregate money supply issued by shadow intermediaries has increased significantly, representing almost half of money-like liabilities in 2005. While the figure illustrates that the amount of credit funded through shadow maturity transformation even at the peak is modest (approximately 10 percent), the growing importance of shadow money in the aggregate supply of money was an important factor in amplifying the shocks to the economy more broadly.

Third, while maturity transformation by shadow intermediaries has increased over the period, the larger story is clearly the increased role of term debt markets in funding credit. In particular, the amount of shadow credit transformation increased from an insignificant amount in 1960 to as much as 164 percent of GDP in 2007. The increase in market funding for credit is driven not only through shadow credit transformation by the GSEs and securitisation, but also through traditional credit transformation by

the increased importance of mutual funds and REITs, which together are now also about 129 percent of GDP. Credit transformation broadly now represents about 293 percent of GDP, corresponding to two-thirds of all financial sector liabilities.

Overall, the Figure documents that shadow banking broadly is not displacing activity that existed in the traditional sector. If anything appears to be associated with a significant deepening of credit markets.

*C) Examples*

The types of activities, institutions, and vehicles that are part of the shadow banking system are constantly evolving. The examples provided in this subsection are by no means exhaustive, but they do represent parts of the shadow banking system that have been particularly important at some point in time and some that still are.

**ABCP Conduits**

Asset-backed commercial paper (ABCP) is commercial paper collateralised by a specific pool of financial assets. ABCP is usually issued by bankruptcy-remote special-purpose vehicles (SPVs), such as ABCP conduits or special investment vehicles (SIVs). Both ABCP conduits and SIVs obtain credit ratings on the issued paper. ABCP ratings are largely based on the credit profile of banks providing credit and liquidity support by commercial banks, while SIV ratings are based on the credit quality of the assets as well as the overall funding strategy of the SIV. Single-seller ABCP conduits are backstops to the working capital needs of large nonbank finance companies and receive such support from a single commercial bank, while multi-seller conduits fund the working capital needs of smaller nonbanks and receive the support of multiple institutions. Similarly, SIVs can either be affiliated with a single banking institution, or obtain support from multiple institutions.

The bankruptcy remoteness of all of these entities implies that the collateral backing the ABCP is exempt from the potential bankruptcy of the institution that provides the backup lines of credit and liquidity. The maturity of ABCP is between one and 180 days, exposing the ABCP to rollover risk, a source of fragility for ABCP issuers that will be discussed later. There were a few examples of ABCP issuers that did not receive unconditional enhancements from commercial banks. One is the Canadian ABCP market, where investors were forced to hold defaulted paper. In addition, extendible ABCP effectively transfers the rollover risk to investors, thus requiring a higher rate of return.

Structured investment vehicles (SIVs) are specialised financial institutions that conduct shadow maturity transformation. On the asset side of SIVs are securitised assets such as ABS, MBS, CDOs, CLOs, CMOs, or financial sector debt. These assets are funded through issuance of ABCP, medium-term notes (MTN), or long-term notes (LTN). In order to achieve a credit rating on their liabilities, SIVs obtain backup lines of credit from commercial banks. SIVs were first created in 1988, effectively moving the financing of ABS from the balance sheet of Citigroup to an off-balance-sheet SIV. While some SIVs are closely associated with particular financial institutions, others operate independently of any particular institution. Since the financial crisis of 2007-09, SIVs have stopped operating. SIVs resemble commercial banks in many ways, but both assets and liabilities are tradable, and liquidity and credit backstops are provided by private institutions.

ABCP has provided funding flexibility to borrowers and investment flexibility to investors since the 1980s, when ABCP was used as a way for commercial banks to fund customer trade receivables in a capital-efficient manner and at competitive rates. ABCP became a common source of warehousing for ABS collateral in the late 1990s. The permissible off-balance-sheet structure facilitated balance-sheet size management, with the associated benefits of reduced regulatory capital requirements and leverage. ABCP funding has also been a source of fee-based revenue. For corporate users, ABCP benefits include some funding anonymity, increased commercial paper (CP) funding sources, and reduced costs relative to strict bank funding. Over time, ABCP conduits expanded from the financing of short-term receivables used as collateral to a broad range of loans, including auto loans, credit cards, student loans, and commercial mortgage loans. At the same time, as the market developed, it came to embed much more maturity mismatch through funding longer-term assets, warehoused mortgage collateral, etc. Securities arbitrage vehicles are one particular example of a shadow banking institution that performed substantial amounts of maturity transformation. These vehicles used ABCP to fund various types of securities, including collateralised debt obligations (CDOs), asset-backed securities (ABS), and corporate debt.

ABCP experienced a run that began in the summer of 2007, when the sponsor of a single-seller mortgage conduit, American Home, declared bankruptcy, and three mortgage programmes extended the maturity of their paper. On August 7, BNP Paribas halted redemptions at two affiliated money market mutual funds when it was unable to value ABCP holdings. Covitz, Liang, and Suarez (2012) use data from the Depository Trust Clearing Corporation (DTCC) to document an investor run on more than 100 programmes, one-third of the overall market. While runs were more likely on programmes with greater perceived subprime mortgage exposure, weaker liquidity support, and lower credit ratings, there is also evidence of investor runs that were unrelated to specific programme characteristics.

#### ABS issuers

Asset-backed securities (ABS) are collateralised claims on pools of loans, mortgages, or receivables. The cash flow and income from ABS are structured into tranches, which receive credit ratings. For example, the super senior AAA tranche might represent 80 percent of the total value of the ABS, the mezzanine BBB tranche might represent 15 percent of the total value, and the remainder may be allocated to an equity tranche. Such pooling and tranching of the ABS are referred to as securitisation as the ABS' value is securitised by its collateral.

Securitisation activity is at the heart of shadow banking, as it allows credit originators to sell pools of credit to other institutions, thereby transferring the credit risk. Securitised products such as ABS are sold to banks, shadow banks, and real money investors. The underlying assets of ABS consist of receivables from credit cards, auto loans, mortgages, and aircraft leases, among others. Even royalty payments and movie revenues have been securitised. Securitisation techniques such as ABS represent a major form of financial innovation in recent decades and are tightly linked with both the credit cycle and the development of the shadow banking system.

Legally, the ABS is structured as a bankruptcy-remote SPV. ABS typically perform no maturity transformation, but do achieve credit and liquidity transformation. Credit transformation is achieved through diversification. For example, the ABS collateral might consist of subprime mortgage loans, while much of the ABS' liabilities consist of AAA assets. Liquidity transformation occurs because any individual mortgage or loan of the ABS collateral might be illiquid due to adverse selection problems, yet a pool of such assets might be liquid. However, the liquidity of the ABS depends crucially on the business cycle, as ABS become more illiquid during downturns, particularly during financial crises.

One special form of ABS is the collateralised debt obligation (CDO), which is secured by a smaller number of loans or by bonds. For other forms of ABS, collateral consists of a large number of individual loans, mortgages, or receivables. For CDOS, however, the collateral can be corporate bonds, structured credit products such as ABS, or pools of agency mortgage-backed securities (MBS). When the collateral of a CDO is ABS, it is sometimes called an ABS CDO. When collateral is MBS, the CDO is called a collateralised mortgage obligation (CMO). There are also collateralised loan obligations (CLOs), which are CDOs with syndicated loans as collateral. The underlying loans of CLOs are often leveraged loans, used to restructure the funding of corporations to allow for more leverage.

Historically, the first CMO was issued by Salomon Brothers and First Boston in 1983 for Freddie Mac, and the first CDO was issued by Drexel Burnham Lambert for Imperial Savings Association in 1987. The credit quality of ABS CDOs is often enhanced through CDOs on the underlying mezzanine tranches of the ABS that are re-securitised. This enhancement reduces the credit risk of the CDO's collateral and allows the issuance of AAA tranches from an underlying pool of mezzanine tranches, which can in turn be funded in shorter-term markets. CDO issuance peaked in 2007 and then totally collapsed in the aftermath of the 2007-09 financial crisis.

#### Tri-party Repo

A repurchase agreement (repo) is the sale of securities together with an agreement that the seller will buy back the securities at a later date. Most repo contracts are short term—between one and 90 days—although there are repos with much longer maturities. Repos are over-collateralised, and the difference between the value of the collateral and the sale price is called the repo haircut. In addition, the repurchase price is greater than the sale price, the difference constituting the repo rate, which is, in economic terms, an interest rate on a collateralised loan. In a repo transaction, the party buying the collateral acts as a lender.

The distinguishing feature of a tri-party repo is that a clearing bank acts as an intermediary between the two parties to the repo. The clearing bank is responsible for the administration of the transaction, including collateral allocation, marking to market, and substitution of collateral. The tri-party structure ensures that both the borrower and the lender are protected against the default of the other, as the collateral resides with a third party. The U.S. tri-party repo market represents a major source of funding for security broker-dealers. The market peaked at slightly above $2.8 trillion in 2008 and is currently slightly below $1.7 trillion.

Investors in tri-party repo are primarily money market mutual funds and other cash-rich investors such as corporate treasury functions, while the borrowers are large securities dealers with inventories of securities to finance. Clearing banks unwind these trades each afternoon and return the cash to the investors. But because the dealers retain a portfolio of securities that need financing on a 24-hour basis, they must extend credit to the other dealers against these securities for several hours between that afternoon unwind and the settlement of new repos in the early evening. That way, those dealers can repay their investors and avoid defaulting on the obligations.

Since the enactment of the Bankruptcy Amendments and Federal Judgeship Act of 1984, repos on Treasury, federal agency securities, bank certificates of deposits, and bankers' acceptances have been exempted from the automatic stay in bankruptcy. The bankruptcy exception ensured the liquidity of the repo market by assuring lenders that they would get speedy access to their collateral in the event of a dealer default. In 2005, the safe harbour provision was expanded to repos written on broader collateral classes, including certain mortgage-backed securities. This broadening of acceptable collateral for the exemption from the automatic stay for repos allowed the repo market to fund credit collateral—and thus directly fund the shadow banking system.

It should be noted that the tri-party repo market is only a subset of other repo and short-term, collateralised borrowing markets. While broker-dealers conduct their funding primarily in the tri-party repo market, their lending occurs mainly in DVP (delivery versus payment) repo or GCF (general collateral finance) repo. In contrast to a tri-party repo, DVP repos are bilateral transactions that are not settled on the books of the clearing banks. Instead, settlement typically occurs when the borrower delivers the securities to the lender. Adrian *et al.* (2013) discuss various forms of repo and securities lending, and Fleming and Garbade (2003) describe GCF repo, which is conducted among dealers.

Copeland, Martin, and Walker (2011) document the collateral composition in the tri-party market, as well as the repo market conventions, using data from July 2008 to early 2010. They show that, during this period, several hundred billion dollars of collateral in the tri-party repo market consisted of collateral such as equities, private-label ABS, and corporate credit securities without any eligibility for public sources of liquidity or credit backstops. Krishnamurthy, Nagel, and Orlov (2011) complement this finding by looking directly at the collateral of MMMFs. While they find that the majority of the $3.5 trillion MMMFs' collateral is of high quality, they do document several hundred billion dollars of private-label ABS securities funded by MMMFs. However, the overall amount of private-label ABS funded in the repo market by MMMFs is less than 3 percent of total outstanding.

Adrian and Shin (2009, 2010a) study the role of repo for security broker-dealers and document the growth of the sector since the 1980s. A distinguishing feature of the balance sheet management of security broker-dealers is the procyclicality of their leverage. Balance sheet expansions tend to coincide with expansions in broker-dealer leverage, while balance sheet contractions are achieved via deleveraging. Adrian and Shin show that repos play the crucial role in this leverage cycle of the broker-dealers: The majority of the adjustment in balance sheet size tends to be achieved through adjustments in the size of the repo book. While Adrian and Fleming (2005) point out that the net funding

of dealers in the repo market tends to be small, Adrian and Shin (2010a) argue that the overall balance sheet size of financial intermediaries can be viewed as an indicator of market liquidity. When gross balance sheets are reduced through deleveraging, financial market liquidity tends to dry up.

#### Money Market Funds

Money market mutual funds are open-ended mutual funds that invest in short-term securities such as Treasury bills, commercial paper (including ABCP), and repo. MMMFs were first created in 1971 in response to Regulation Q, which restricted the interest that commercial banks can pay on deposits. Since then, money market funds have represented an alternative to bank deposits from investors' point of view, with yields that are typically more attractive than bank deposits. The money market sector peaked at around $3.5 trillion in 2008. MMMFs are regulated by the SEC under the Investment Company Act of 1940.

Money market funds seek a stable net asset value (NAV), which is generally $1.00, meaning that they aim never to lose money. If a fund's NAV drops below $1.00, it is said to "break the buck." In September 2008, the day following the Lehman Brothers bankruptcy, the Reserve Primary Fund broke the buck and triggered a run on MMMFs. Other fund managers reacted by selling assets and investing at only the shortest of maturities or by reallocating to Treasury bills, thereby exacerbating the funding difficulties for other instruments such as commercial paper and repo.

Wermers (2011) investigates in more detail the role of investment flows into and out of money market mutual funds, focusing particularly on the period of the financial crisis. Wermers shows that institutional investors were more likely to run than retail investors, and institutional investors tended to spread such run behaviour across various MMMF families. Institutional MMMF investors can thus be viewed as a transmission channel for contagious runs. Kacperczyk and Schnabl (2011) analyse the impact of the organisational structure of MMMFs on their risk-taking behaviour. In particular, they ask how the risk-taking differs between stand-alone funds and the funds that are owned by larger holding companies, such a bank holding companies. Kacperczyk and Schnabl find significant differences in the risk-taking of stand-alone MMMFs relative to the funds that have implicit guarantees from financial conglomerates. During the financial crisis of 2008, when systemic risk increased and conglomerates became relatively more exposed to systemic risk, stand-alone mutual funds increased their risk-taking behaviour relatively more. Conversely, in the run-up to the crisis, when measured systemic risk was low, MMMFs that were part of conglomerates took on relatively more risk.

## 2) Why Does Shadow Credit Intermediation Exist?

The term "shadow banking" was coined by McCulley (2007) and was picked up by policymakers (see, for example, Tucker (2010)). The first articles on shadow banking appeared in 2008 (Pozsar (2008) and Adrian and Shin (2009)). A comprehensive overview of the shadow banking system can be found in Pozsar *et al.,* (2010). An update on regulatory reforms relating to shadow banking can be found in Adrian and Ashcraft (2012). Academic studies of shadow banking include Gorton and Metrick (2011, 2012),

Gennaioli, Shleifer, Vishny (2012b), Stein (2010), and Acharya, Schnabl, and Suarez (2010). In addition to the academic literature by financial economists, legal scholars have contributed to the shadow banking literature (e.g., Schwarcz (2012) and Ricks (2010)).

The Financial Stability Board (FSB) has initiated international working groups on shadow banking (see FSB (2011, 2012)). The objective of the FSB is to enhance the regulation and oversight of the shadow banking system. The FSB is leading this work because of the global reach of shadow banking, which acts as an international systemic risk transmitter in times of crisis. In response to the tightened financial regulation, international shadow bank regulatory arbitrage might very well be growing in the future, making an adequate regulatory framework and monitoring system imperative. FSB (2012) presents a classification of shadow banking working groups, with the aim to develop a framework for policy recommendations and monitoring. The classifications are 1) banks' interactions with shadow banks, 2) money market mutual funds, 3) other shadow banking entities, 4) securitisation activity, and 5) securities lending and repos. Finally, industry groups are also studying shadow banking. The Institute of International Finance (2012) has put forward a framework for policymaking in relation to shadow banking. In addition, the Securities Industry and Financial Markets Association (SIFMA) has multiple workstreams on the topic of shadow banking.

There are three broad explanations for the existence of shadow banks: A) innovation in the composition of aggregate money supply; B) capital, tax, and accounting arbitrage; and C) other agency problems in financial markets. We discuss each of these explanations. Empirically, they are intertwined, and it is difficult to attribute relative magnitudes to each of them.

*A) Innovation in the Composition of Aggregate Money Supply*

Drawing motivation from the narrative of Gorton and Metrick (2011), it is possible to view shadow credit intermediation as financial innovation in the composition of aggregate money supply. Money plays a crucial role in the economy, acting not only as a store of value, but also as a unit of account and means of exchange. The rapid loss of confidence in the value of money has been a root cause of financial panics across countries and over time.

The earliest forms of money, commodity money, were made of precious metals, having inherent value by being comprised of gold or silver. However, commodity money was eventually replaced with fiat money, which has little intrinsic value, but is instead backed only by the issuer's promise to convert the notes into a commodity. In particular, the banking system of the early 1800s was characterised by banks that issued notes backed by the promise of convertibility into gold or silver coin.

Banking charters were tightly restricted by state legislatures. In the Free Banking Era (1837 to 1862), there was free entry into the sector for any banker with adequate initial capital, but banks were required to deposit state or federal government bonds with face value equal to the value of notes issued with a state representative. While one might have thought that the presence of collateral would have made free banking stable, the period was characterised by a series of panics, and almost one-third of all free banks

ultimately failed. The root cause of these panics is a subject of debate in the academic literature; reductions in the value of state debt likely played a prominent role. Jaremski (2010) documents that failure rates of free banks is correlated with state bond prices, but does not find the same for charter banks. Rockoff (1991) suggests that the existence of market discounts on state bonds not only reduced confidence by note holders, limiting their liquidity and value, but also created scope for "wildcat banking," where implicit leverage between the face value of bank notes and the market value of state government bonds permitted banks to have extraordinary leverage. The scope for panics created by concern about the value of bank notes was eliminated by the passage of the National Banking Acts in 1863 and 1864. This legislation replaced bank notes with a national currency backed by the deposit of U.S. Treasury bonds, enacted a ceiling on the aggregate circulation of notes, and set required reserves on both notes and deposits in order to encourage banks to hold safer portfolios.

While the National Banking Acts created confidence in currency, financial innovation in the composition of money in the form of bank deposits had already occurred. While bank notes were secured, deposits were secured only by the general assets of the bank. When depositors lost confidence in the solvency of a bank, they would insist that the bank honour its obligation to convert deposits into specie. As banks had a limited supply of specie in reserve, they could not accommodate large runs by depositors, which increased the incentives of depositors to run at the first sign of trouble. State governments made numerous attempts to stabilise deposits through insurance schemes, but most of them failed. As a result, the industry created collectives known as clearinghouses, which carefully monitored the financial condition of their members and stood behind their liabilities in the event of a run by depositors.

The first clearinghouses were established by New York banks in 1853. Gorton (1985) documents that when one member faced a run, the clearinghouses suspended the production of bank-specific financial information and instead published financial information on all members together. In order to prevent the costly liquidation of illiquid assets like loans, the clearinghouses issued loan certificates to members, secured by members' assets. These certificates could be used in the clearing process in place of currency, which freed up currency to accommodate withdrawals by depositors. During the panics of 1893 and 1907, the clearinghouses issued loan certificates directly to the public, permitting depositors to replace their claims on a bank with a claim on the clearinghouse. While the creation of the Federal Reserve in 1913 was intended to bring stability to the banking system by replacing the system clearinghouses, the central bank did not begin to act as a lender of last resort until well after the Great Depression. Consequently, it was the enactment of federal deposit insurance in 1933 that first brought stability to demand deposits.

Over the past thirty years, significant innovations in the composition of the aggregate money supply have made the financial system more vulnerable to a loss of confidence by the holders of money. In particular, money market mutual funds were developed in the 1970s in response to limits on the ability of depository institutions to pay interest on checking accounts, as well as in response to a need for limits on deposit insurance, which left large depositors exposed to bank risk. One of the main investments of money market mutual funds is overnight repurchase agreements, the

equivalent of bank notes secured by collateral, most often U.S. Treasury obligations. Seeking stability, financial innovation transformed uninsured deposits into an instrument that looks like an insured deposit in the form of an overnight repurchase obligation.

Sunderam (2012) explores the extent to which shadow banking liabilities constitute substitutes for high-powered money. He shows in a simple model that shadow banking liabilities should constitute substitutes for money in the private sector's asset allocation. Empirically, Sunderam shows that shadow banking liabilities respond to money demand, extrapolating that heightened money demand can explain about half of the growth of ABCP in the mid-2000s. He also confirms that regulatory changes to ABCP played a significant role in the growth of the shadow banking system, a theme that we turn to in the next section.

*B) Capital, Tax, and Accounting Arbitrage*

We define shadow banking activities as banking intermediation without public liquidity and credit guarantees. The value of public guarantees was rigorously modelled by Merton (1977) using an options pricing approach. Merton and Bodie (1993) propose the functional approach to financial intermediation, which is an analysis of financial intermediaries in relation to the amount of risk-sharing that they achieve via guarantees. Pozsar *et al.,* (2010) provide a comprehensive overview of shadow banking institutions and activities that can be viewed as a functional analysis of market-based credit intermediation. Many of their insights are comprised in maps of the shadow banking system that provide a blueprint of the funding flows. Levitin and Wachter (2011) provide a quantitative assessment of the role of implicit guarantees for the supply of mortgages. There is also a large literature that studies the implicit guarantees of the GSEs (see Passmore, Sherlund, and Burgess (2005), Frame and White (2005), and Acharya *et al.,* (2011)).

Acharya, Schnabl, and Suarez (2011) document that the rapid expansion of ABCP since 2004 resulted from changes in regulatory capital rules. In particular, FASB issued a directive in January 2003 (FIN 46) and updated the directive in December 2003 (FIN 46A) suggesting that sponsoring banks should consolidate assets in ABCP conduits onto their balanced sheets. However, U.S. banking regulators clarified that assets consolidated onto balance sheets from conduits would not need to be included in the measurement of risk-based capital and instead used a 10 percent credit conversion factor for the amount covered by a liquidity guarantee. Acharya, Schnabl, and Suarez document that the majority of guarantees were structured as liquidity-enhancing guarantees aimed at minimising regulatory capital, instead of credit guarantees, and that the majority of conduits were supported by commercial banks subject to the most stringent capital requirements. Moreover, the authors document that conduits were sponsored by banks with low economic capital as measured by the ratio of the book value of equity to assets. Finally, the authors find that investors in conduits with liquidity guarantees were repaid in full, while investors in conduits with weaker guarantees suffered small losses, suggesting there was no risk transfer despite the capital relief.

The motivation for capital arbitrage is consistent with the mispricing of explicit credit and liquidity put options associated with deposit insurance and access to official

liquidity, as well as the presence of a perception that large banks are "too big to fail," which permits them to engage in excessive leverage maturity transformation. The presence of minimum capital and liquidity standards mitigates these incentives, and the ability of banks to evade binding standards permits them to maximise the value of these put options.

*C) Other Agency Problems in Financial Markets*

Ashcraft and Schuermann (2008) describe seven important informational frictions that existed in the securitisation of subprime mortgage credit prior to the financial crisis, although these frictions can be generalised to all securitisation transactions. They include asymmetric information problems between the lender and originator (predatory lending and borrowing), between the lender and investors, between the servicer and investors, between the servicer and borrower, between the beneficiary of invested funds and asset managers, and between the beneficiary of invested funds and credit rating agencies. In addition, asymmetric information between investors and issuers results in risk-insensitive cost of funding. For example, Keys *et al.* (2010) document that mortgage borrowers with FICO scores just above a threshold of 620 perform significantly worse than borrowers with FICO scores just below 620. As it is more difficult to securitise loans below that threshold, the authors argue that this result is consistent with issuers exploiting asymmetric information, disrupting the otherwise monotone relationship between borrower credit scores and performance.

Although securitisation has a relatively short history, it is a troubled one. The first known securitisation transactions in the United States occurred in the 1920s, when commercial real estate (CRE) bond houses sold loans to finance CRE to retail investors through a vehicle known as CRE bonds. Wiggers and Ashcraft (2012) document the performance of these bonds, which defaulted in large numbers following the onset of the Great Depression. Although the sharp deterioration in economic conditions played an important part in explaining their poor performance, so did aggressive underwriting and sales of the bonds in small denominations to unsophisticated retail investors.

During the 1990s no fewer than five different sectors of ABS ran into trouble, including but not limited to home equity, home improvement lending, manufactured housing, equipment leasing, and franchise loans. In each of these cases, there was generally meaningful risk retention by a sponsor using securitisation as a source of funding. However, one common theme appears to have been the aggressive pursuit of gain-on-sale-related earnings from securitisation in advance of an initial public offering, and this was often achieved through competition on underwriting standards. In contrast, the challenges of securitisation in the 2000s were concentrated in multisector CDOs in 2002 as well as RMBS and CMBS in 2005-07. These credit cycles were more likely to involve firms using securitisation for arbitrage and were used as a source of fee income with minimal intended risk retention, although many of them were left holding warehouses of loans as the financial crisis unfolded.

Over-reliance on credit ratings can create problems when the rating agencies face their own agency problems. For example, Mathis, McAndrews, Rochet (2009) analyse a dynamic model of ratings where reputation is endogenous and the market environment may vary over time. The authors' model predicts that a rating agency is likely to issue

less accurate ratings in boom times than it would during recessionary periods. Moreover, the authors demonstrate that competition among rating agencies yields similar qualitative results. Xia and Strobl (2012) document that the conflict of interest caused by the issuer-pays rating model leads to inflated corporate credit ratings. In particular, the authors compare the ratings issued by an issuer-paid rating agency with an investor-paid agency and demonstrate that the difference between the two is more pronounced when issuer-paid agency's conflict of interest is particularly severe. For example, the issuer-paid agency has more favourable ratings for firms with more short-term debt, a newly appointed CEO or CFO, and a lower percentage of past bond issues rated by the agency. However, the authors find no evidence that these variables are related to corporate bond yield spreads, which suggests that investors may be unaware of incentive problems at the issuer-paid agency. Cohen (2010) documents significant relationships between variables that should not affect a CRA's view of the credit risk of conduit/fusion CMBS transactions issued during 2001-07, but that would affect issuers' and CRAs' incentives in an environment where rating shopping was present.

## 3) How Does Shadow Credit Intermediation Work?

Pozsar *et al.,* (2010) make a distinction between the "internal," "external," "independent," and "government sponsored" shadow banking system. The internal system consists of shadow banking activities conducted under the auspices of bank holding companies. The external system comprises shadow banking activities that are conducted by major nonbank financial institutions such as nonbank-affiliated broker-dealers or insurance companies. Independent shadow banking institutions are specialised shadow banking vehicles that are independent of any major financial institutions. Finally, the government-sponsored shadow banking system consists of institutions that provide credit intermediation services with implicit government guarantees. Before discussing the various shadow banking institutions in detail, we review the "seven steps" of shadow credit intermediation (Figure 3).

### *A) The Seven Steps of Shadow Credit Intermediation*

The shadow banking system is organised around securitisation and wholesale funding. Loans, leases, and mortgages are securitised and thus become tradable instruments. Funding is also in the form of tradable instruments, such as commercial paper and repo. Savers hold money market balances, instead of deposits with banks.

The shadow banking system decomposes the credit intermediation into a chain of wholesale-funded, securitisation-based lending. Through the shadow intermediation process, the shadow banking system transforms risky, long-term loans (subprime mortgages, for example) into seemingly credit-risk-free, short-term, money-like instruments. Shadow credit intermediation is performed through chains of nonbank financial intermediaries in a multistep process that can be interpreted as a "vertical slicing" of the traditional bank's credit intermediation process into seven steps. Pozsar *et al.,* (2010) explain the seven steps of shadow bank credit intermediation:

1. Loan origination (auto loans and leases, nonconforming mortgages, etc.) is performed by finance companies.

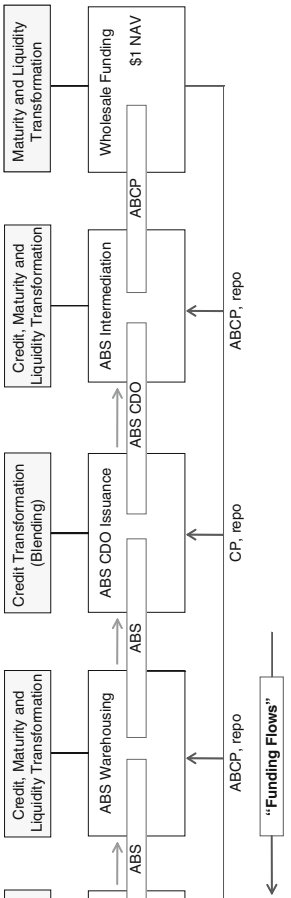



Figure 3: **The shadow credit intermediation process. Source: Pozsar *et al.*, (2010).**

2. Loan warehousing is conducted by single- and multi-seller conduits and is funded through asset-backed commercial paper (ABCP).
3. The pooling and structuring of loans into term asset-backed securities (ABS) is conducted by broker-dealers' ABS syndicate desks.
4. ABS warehousing is facilitated through trading books and is funded through repos, total return swaps, or hybrid and repo conduits.
5. The pooling and structuring of ABS into CDOs is also conducted by broker-dealers' ABS.
6. ABS intermediation is performed by limited-purpose finance companies (LPFCs), structured investment vehicles (SIVs), securities arbitrage conduits, and credit hedge funds, which are funded in a variety of ways including, for example, repo, ABCP, MTNs, bonds, and capital notes.
7. The funding of all the above activities and entities is conducted in wholesale funding markets by funding providers such as regulated and unregulated money market intermediaries (for example, 2(a)-7 MMMFs and enhanced cash funds, respectively) and direct money market investors (such as securities lenders).[1] In addition to these cash investors, which fund shadow banks through short-term repo, CP, and ABCP instruments, fixed-income mutual funds, pension funds, and insurance companies also fund shadow banks by investing in their longer-term MTNs and bonds.

Not all intermediation chains involve all seven steps, and some might involve even more steps. For example, an intermediation chain might stop at the second step if a pool of prime auto loans is sold by a captive finance company to a bank-sponsored multi-seller conduit for term warehousing purposes. In another example, ABS CDOs could be further repackaged into a CDO^2, which would elongate the intermediation chain to include eight steps. Typically, the poorer an underlying loan pool's quality at the beginning of the chain (for example, a pool of subprime mortgages), the longer the credit intermediation chain that would be required to "polish" the quality of the underlying loans to the standards of money market mutual funds and similar funds. As a rule of thumb, the intermediation of low-quality long-term loans (nonconforming mortgages) involved all seven or more steps, whereas the intermediation of high-quality short- to medium-term loans (credit card and auto loans) involved usually three steps (and rarely more). The intermediation chain always starts with origination and ends with wholesale funding, and each shadow bank appears only once in the process.

*B) Commercial Banks and Shadow Banking*

Per definition, credit intermediation activity on the balance sheets of commercial banks does not constitute shadow banking, as it has access to official liquidity and credit guarantees by the Federal Reserve and the FDIC. However, commercial banks can be involved in shadow banking activities in several ways. For example, they can provide credit and liquidity lines to shadow bank entities such as conduits, ABS issuers, or SIVs. Commercial banks in turn are owned by bank holding companies (BHCs).

[1]Money market mutual funds in the United States are regulated under Rule 2a-7 of the Securities and Exchange Commission's (SEC) Investment Company Act of 1940.

Mandel, Morgan, and Wei (2012) provide a detailed analysis of commercial banks' sponsorships of shadow banking activities.

Many shadow banking activities are conducted under the auspices of BHCs. For example, a BHC might own a wealth management unit with a money market mutual fund, which we would consider a shadow bank internal to the BHC. Another example is tri-party repo funding by the broker-dealer subsidiaries of BHCs. A third example is the use of ABCP conduits, which are off balance sheet to the BHC, but are sponsored by the commercial bank subsidiary of the BHC via credit and liquidity lines.

One gauge of the extent of shadow banking activity by BHCs is their organisational complexity. While traditional banking is done in a single entity, BHCs tend to have hundreds or thousands of subsidiaries, most of which do not have direct, explicit access to public credit and liquidity puts. For example, Avraham, Selvaggi, and Vickery (2012) document that each of the five largest BHCs in the U.S. had over 1,500 subsidiaries in 2012, with the largest one owning more than 3,000. While some of these subsidiaries are foreign banks, most of them are nonbank subsidiaries in the United States. The majority of the subsidiaries are funds, trusts, and financial vehicles that are typically engaging in shadow banking activities. In fact, Copeland (2012) shows that these shadow banking activities of bank holding companies have been increasing over time and represent a quantitatively important share of the holding companies' total earnings.

Cetorelli and Peristiani (2012) investigate the role of BHCs in asset securitisations. They assess quantitatively the degree to which commercial banks are involved in ABS, CMBS, and CDO issuance, and in servicing and underwriting securitisations. For nonagency ABS, the BHC market share is between 35 and 75 percent for underwriting, issuance, and servicing and close to 100 percent for trust services. In contrast, for private-label mortgages, the market share of BHCs has increased dramatically over the past twenty years to over 60 percent for underwriting, servicing, and issuance. This suggests that the presence of BHCs in shadow banking activities relating directly to securitisations is substantial.

Bord and Santos (2012) study the role of banks in the originate-to-distribute model of credit intermediation. They base their study on data from the Shared National Credit Program (SNC). The data are collected by the national supervisory banking agencies (Federal Reserve, the FDIC, or the OCC), which track credit held by federally supervised institutions. Unlike any form of publicly available data, the SNC allows Bord and Santos to track the ownership of loans by various institutions over time. The authors document that more than 75 percent of syndicated credit lines are bought by syndicate participant banks and that they stay with those banks after three years. The share of term loans owned by syndicate banks has fallen from around 75 percent in the mid-1990s to around 30 percent in the mid-2010s. For term loans, shadow banking organisations have thus emerged as more and more important investors over the past twenty years. Buyers of term loans that are particularly important are investment managers and collateralised loan obligations (CLOs). Bord and Santos conclude that the share of term loans sold to the shadow banking system amounted to less than 10 percent in 1993 and rose to over 30 percent by 2007. While loan originations are conducted almost exclusively by commercial banks, the ultimate owners of term loans are thus split among banks and shadow banks.

Besides the subsidiaries associated with BHC involvement in securitisation activities, the largest nonbank BHC subsidiaries consist of finance companies, broker-dealers, and wealth management units including mutual, hedge, and money market mutual funds. In many respects, the financial crisis of 2007-09 has led to a financial system where the BHCs own a larger share of nonbank subsidiaries that conduct shadow banking activities. For example, the five largest independent broker-dealers prior to the crisis all were absorbed by, or transformed into, BHCs. Similarly, some of the largest independent issuers, originators, and servicers of private-label mortgages were absorbed by BHCs. So while the two decades in the run-up to the financial crisis saw the emergence of a shadow banking system that was partially independent from BHCs, the financial crisis led, perhaps paradoxically, to a migration of independent shadow banking activity into BHCs. Cetorelli (2012) shows that, as of 2011, BHCs controlled about 38 percent of the assets of the largest insurance companies, 41 percent of total money market mutual fund assets, and 93 percent of the assets of the largest brokers and dealers. Moreover, very little securities lending and related cash collateral reinvestments take place without the services provided by the main custodian banks.

The trend toward consolidation of shadow banking in BHCs since the crisis is, however, counteracted by a powerful force: the enhanced prudential standards of BHCs. Tighter capital and liquidity requirements will arguably lead to an increased incentive for some forms of credit intermediation to migrate out of BHCs and into the shadow banking system. This trend has been observed in proprietary trading, which has largely migrated from the BHCs to independently run hedge funds in anticipation of the Volcker Rule (see Duffie (2012) for an academic assessment of the rule). Furthermore, the CLO market continues to thrive and is at least partially independent from BHCs. A broader movement of securitisation-related activity from BHCs to shadow banking institutions can be expected over time.

*C) External and Independent Shadow Banking*

External shadow banking entities are regulated institutions that are independent of any entities with direct, explicit government backstops and that conduct shadow banking activities. Examples of such institutions are stand-alone broker-dealers; independent wealth management firms that run money market funds; credit hedge funds; and finance companies that are affiliated with industrial companies such as the auto loan subsidiaries of car manufacturers. Independent shadow banks are institutions independent of the government safety net; they are set up to engage exclusively in certain shadow banking activities. In contrast to the institutions of the external shadow banking system, independent shadow banks do not have non-shadow banking activities as their primary business, but rather specialise only in shadow banking. Examples of independently run shadow banks are nonbank affiliated SIVs, stand-alone money market mutual funds, independent CDOs and CLOs, and the majority of ABS and private-label RMBS and CMBS.

The fifth and sixth steps of the credit intermediation chain rely heavily on private credit risk repositories to perform originate-to-distribute securitisations. Private risk repositories specialise in providing credit transformation services in the shadow banking system and include mortgage insurers, monoline insurers, certain subsidiaries of large, diversified insurance companies, credit hedge funds, and credit derivative product

companies. These entities, as investors in the junior equity and mezzanine tranches of loan pools, all provide risk capital to the shadow banking system, thereby supporting credit extension to the real economy. Different credit risk repositories correspond to specific stages of the shadow credit intermediation process. As such, mortgage insurers specialise in insuring or wrapping whole mortgage loans; monoline insurers specialise in wrapping ABS tranches (or the loans backing a specific ABS's tranches); and large, diversified insurance companies, credit hedge funds, and credit derivative product companies specialise in taking on the risks of ABS CDO tranches through CDS. There are also overlaps, with some monolines wrapping both ABS and ABS CDOs, for example. Effectively, the various forms of credit put options provided by private risk repositories absorb tail risk from loan pools, turning the enhanced securities into securities that are free from credit risk (at least from investors' perceptions prior to the crisis). This in turn means that any liability that issued against these assets is perceived to be free of credit risk as well, just as if it is FDIC-insured.

The perceived credit-risk-free nature of traditional banks' and shadow banks' liabilities stems from two very different sources. In the case of traditional banks' insured liabilities (deposits), the credit quality is driven by the counterparty—the U.S. taxpayer. As a result, insured depositors invest less effort into examining a bank's creditworthiness before depositing money than they would if they were uninsured. In the case of shadow banks' liabilities (repo or ABCP, for example), perceived credit quality is driven by the "credit-risk free" nature of collateral that backs shadow bank liabilities, as it is often enhanced by private credit risk repositories. The credit puts provided by private credit risk repositories are alternatives to the credit transformation performed by 1) the credit-risk-based calibration of advance rates and attachment points on loan pools backing top-rated ABCP and ABS tranches, respectively; 2) the credit-risk-based calibration of haircuts on collateral backing repo transactions; 3) the capital notes supporting LPFCs' and SIVs' portfolios of assets; and 4) the pooling and repackaging of non-AAA-rated term ABS into ABS CDOs. The credit puts of private credit risk repositories are also similar in function to the wraps provided by Fannie Mae and Freddie Mac on conforming mortgage pools. Just as these government-sponsored, public credit risk repositories "borrowed" the AAA-rating of the federal government and extended it to pools of mortgage loans (turning them into credit-risk-free rate products), the private credit risk repositories were effectively "borrowing" the AAA rating of their parent.

*D) Government-Sponsored Shadow Banking*

In many ways, the modern shadow banking system originated in the government sector. Securitisation was first conducted by government-sponsored enterprises (GSE), which are comprised of the FHLB system (1932), Fannie Mae (1938), and Freddie Mac (1970). The GSEs have dramatically impacted the way in which banks are funded and the way in which they conduct credit transformation: The FHLBs were the first providers of term warehousing of loans, and Fannie Mae and Freddie Mac pioneered the originate-to-distribute model of securitised credit intermediation.

Like banks, the GSEs fund their loan and securities portfolios with a maturity mismatch. Unlike banks, however, the GSEs are funded not through deposits, but through capital markets, where they issue short- and long-term agency debt securities. These

agency debt securities are bought by money market investors and real money investors such as fixed-income mutual funds. The funding functions performed by the GSEs on behalf of banks and the way in which GSEs are funded are the models for wholesale funding markets. The GSEs use several securitisation techniques. They use term loan warehousing services provided by the FHLBs. They also use credit risk transfer and transformation through credit insurance provided by Fannie Mae and Freddie Mac. Securitisation functions are provided by Fannie Mae and Freddie Mac. Maturity transformation is conducted on the GSEs' balance sheets through retained portfolios. These securitisation techniques first used by the GSEs were adopted and imitated by banks and nonbanks to generate the nongovernmental shadow banking system. The adaptation of these techniques gave rise to the securitisation-based, originate-to-distribute credit intermediation process.

## 4) Why Does Shadow Credit Intermediation Need to Be Regulated?

During the financial crisis of 2007-09, the shadow banking system collapsed. ABCP conduits experienced a series of runs. Of the five major investment banks, one failed, two were acquired by banks, and two were transformed into bank holding companies. Securitisation activity totally stopped, and many shadow banking institutions such as SIVs and CDOs all but disappeared. The collapse of shadow banking institutions and shadow banking activities occurred both on the asset and liability sides. On the asset side, the main issues were the underwriting standards. On the liability side, the main issues were related to the fragility of wholesale funding. As a result of these fragilities, the government sector set up a variety of backstops for the shadow banking system. These backstops consisted of both liquidity facilities and solvency guarantees. They were created because of the potential for shadow bank distress to spill over to other institutions and damage the real economy.

### *A) Asset Quality*

Because they are tailored to take advantage of mispriced tail risk, shadow banking institutions accumulate assets that are particularly sensitive to tail events. At a deep level, the question becomes, how can the mispricing of tail risk exist in a world with fully rational actors? Shouldn't financial market participants be able to calculate tail risk probabilities, implicit guarantees, and various tail risk enhancements? And shouldn't these calculations lead to the proper assessment of tail risk? The literature has provided two distinct, complementary answers. The first relies on the behavioural explanation of "neglected risk." The second relies on information opacity in a rational world. We will discuss each of these explanations in turn.

Evidence from psychology and behavioural finance argues that market participants are fundamentally biased against the rational assessment of tail risk. Gennaioli, Shleifer, and Vishny (2012a) develop a theory of individual decision making based on the behavioural evidence, positing that actors neglect risk. In a later paper, Gennaioli, Shleifer, and Vishny (2012b) apply this theory to the economics of the shadow banking system. They model a world where investors systematically ignore the worst state of the world, generating overinvestment and overpricing during the boom and excessive collapse of real activity and the financial sector during the bust.

Their theory is possibly the most parsimonious narrative of the boom and bust of the shadow banking system. In fact, much empirical evidence is consistent with such a theory. Credit rating agencies modelled only small or no declines in aggregate housing prices, and investors in securitised products often did not understand the amount of risk exposure that was embedded in the products. Meanwhile, the prices of tail risk far into the future, far out of the money options relating to mortgage credit, were surprisingly cheap. An early paper warning of the financial system's exposure to such tail risk was presented by Rajan (2005), who pointed to precisely this phenomenon by asking whether financial innovation had made the world riskier.

Neglected risks are one way to interpret the widely perceived risk-free nature of highly rated structured credit products, such as the AAA tranches of ABS. Coval, Jurek, and Stafford (2009) point out that these AAA tranches behave like catastrophe bonds that load on a systemic risk state. In such a systemic risk state, assets become much more correlated than in normal times. The underestimation of correlation enabled financial institutions to hold insufficient amounts of liquidity and capital against the puts that underpinned the stability of the shadow banking system, which made these puts unduly cheap to sell. As investors tend to overestimate the value of private credit and liquidity enhancement purchased through these puts, the result is an excess supply of cheap credit. Adrian, Moench, and Shin (2009) document the close correspondence between the pricing of risk and the fluctuations of shadow bank and broker-dealer balance sheets. Times of low-risk premia tend to be associated with expanding balance sheets—in fact, intermediary balance-sheet developments predict the pricing of risk across many asset classes. Neglected risk can manifest itself through over-reliance on credit ratings by investors. For example, Ashcraft *et al.* (2011) document that subprime MBS prices are more sensitive to ratings than *ex post* performance, suggesting that funding is excessively sensitive to credit ratings relative to informational content.

Dang, Gorton, and Holmström (2009) present an alternative theory where, in a world with fully rational market participants, assets are highly exposed to tail risk. Theirs is a theory of information opacity that can serve as a rationalisation of credit problems for the shadow banking system. According to this theory, debt contracts are optimal because they generate opacity. Opacity, in turn, minimises adverse selection and provides the least possible incentives to collect information. This insight justifies the growth of relatively opaque securitised products in the run-up to the crisis. Mortgages and loans were packaged into MBS and ABS and funded by CDOs, SIVs, and MMMFs that had relatively little information about the underlying credit quality. However, Dang, Gorton, and Holmström show that systemic risk is exacerbated once a bad shock hits informationally opaque, debt-funded economies. The intuition is that a bad shock leads to an increase in private information collection, which exacerbates the incorporation of adverse information in market prices. As a result, adverse selection starts to accumulate as systemic crises deepen.

The above theory complements the explanation by Gennaioli, Shleifer, and Vishny (GSV) discussed earlier. While Dang, Gorton, and Holmström (DGH) emphasise adverse selection as an amplification mechanism, GSV emphasise awareness of risk. In GSV, the riskiness of the worst state of the world is simply neglected, and this neglect is based on behavioural arguments. In contrast, in DGH's model, the opacity of

financial contracts in good times is an equilibrium outcome that maximises the liquidity of financial contracts. The commonality between the two theories is that the severity of financial crisis is neglected, either rationally or behaviourally. As a result, the tail risk embedded in debt securities is underpriced from an *ex post* point of view. In both DGH and GSV, the assets that are accumulated during the boom experience large asset price declines during times of crisis. Such theories of neglected risk thus provide a rationalisation for the accumulation of risk exposure to the housing market that was the major aggregate risk of the shadow banking system. In the theories of DGH and GSV, securities such as ABS and CDOs that obscure the underlying credit risks arise naturally. Such securities, in turn, generate large losses in times of crisis.

Such explanations of the boom and bust cycle of securitised credit products need to be complemented with the additional insights from theories discussed in Section 2. Those alternative theories provide additional reasons that give rise to the systematic mispricing of tail risk. In particular, the implicit or indirect access to government backstops via credit and liquidity puts from bank holding companies and insurance companies leads to the underpricing of tail risk and the excessive buildup of systemic risk. The agency conflicts arising at various stages of the shadow banking system, and particularly the misaligned incentives of credit rating agencies, will lead to excessive risk taking in good times, with associated excessive credit losses in times of crisis. Finally, the failure to provide adequate financial disclosure based on accounting rules generates an additional market failure that facilitates excessive risk taking with the associated large losses in downturns.

*B) Funding Fragility*

The financial frictions that lead to excessive risk taking and exacerbated credit losses during downturns also interact with the fragility of funding. Per definition, funding sources for shadow banking activities are uninsured and thus runnable. In many ways, the fragility of shadow banks due to the run-ability of liabilities resembles the banking system of the 19$^{th}$ century, prior to the creation of the Federal Reserve and the FDIC. During that time, bank runs were common, and they often had severe consequences for the real economy.

The shadow banking system's vulnerability to runs bears resemblance to bank runs as modelled by Diamond and Dybvig (1983). Shadow banks are subject to runs because assets have longer maturities than liabilities and tend to be less liquid as well. While the fundamental reason for commercial bank runs is the sequential servicing constraint, for shadow banks the effective constraint is the presence of fire sale externalities. In a run, shadow banking entities have to sell assets at a discount, which depresses market pricing. This provides incentives to withdraw funding—before other shadow banking depositors arrive.

However, the analogy between bank runs and shadow bank runs goes only so far. The reason is that shadow banking entities do not offer demand deposits, but instead obtain funding in wholesale money markets such as commercial paper or repo. Martin, Skeie, and von Thadden (2011) provide a model for a run in repo markets that takes the empirical facts of the Bear Stearns and Lehman crises as a starting point. In their model, repo borrowers face constraints due to the scarcity of collateral and the liquidity

of collateral. Under sufficiently adverse conditions, self-fulfilling runs can occur. The model focuses in particular on the differences between the tri-party repo market and the bilateral repo market (see Adrian *et al.,* (2013) for an overview of both markets). Arguably, runs occurred in both markets, but they were of very different natures. While the run in the bilateral market was characterised by a sharp increase in haircuts (as documented by Gorton and Metrick (2012)), the run in the tri-party repo market materialised as a simple withdrawal of funding with a rather limited impact on the level of haircuts (see Copeland, Martin, and Walker (2011)). Runs in the ABCP market were equally characterised by a withdrawal of funding (see Covitz, Liang, and Suarez (2012)).

Funding fragility of shadow banking institutions can also be interpreted as the result of the leverage cycles of market-based financial institutions. Such leverage cycles refer to equilibrium outcomes, where asset values and balance sheet capacity of intermediaries are determined endogenously. The friction in models of leverage cycles is due to the funding constraints of intermediaries, which reflect the incentive problems discussed earlier. Theories of intermediary leverage cycles have been proposed by Fostel and Geanakoplos (2008), Brunnermeier and Pedersen (2009), Brunnermeier and Sannikov (2011), Garleanu and Pedersen (2011), and Adrian and Boyarchenko (2012). Such theories of leverage cycles have the commonality that intermediaries are subject to collateral constraints, as is the case for repo and ABCP funding. The tightness of the collateral constraints depends on the underlying risk of assets, the liquidity of assets, and the collateral values. As economic conditions deteriorate, the leverage cycle acts as an amplification mechanism to underlying shocks.

Adrian and Boyarchenko (2012) show that their theory of intermediary leverage cycles has strong empirical support. Intermediary balance sheets exhibit strongly procyclical leverage, meaning that leverage expands in booms. This procyclical behaviour of leverage is a hallmark of shadow banking, as documented by Adrian and Shin (2009). Shadow bank leverage tends to be high when balance sheets are large and credit intermediation is expanding. Furthermore, equity is countercyclical, both in the theory and in the data, as intermediaries tend to hold as little equity as possible during booms, but are forced to raise equity during downturns when the market risk increases. Adrian and Boyarchenko (2012) also document the close link between intermediary balance sheets and asset prices. Over time, expanding leverage tends to coincide with compressed risk premia and inflated asset prices. In busts, risk premia widen, generating asset price busts. In addition, market volatility is countercyclical. As a result, the funding of intermediaries tends to collapse during times of crisis.

The advantage of the general equilibrium theories of leverage cycles is that they allow welfare analysis, thus tying funding fragility in financial crises to possible policy interventions. In general, the market equilibrium is not welfare optimising, and policies that mitigate the cycle can enhance welfare. In the context of shadow banking, this implies that the run-up of shadow banking just before the crisis and its collapse during the crisis are exacerbating the financial cycle in a way that does not enhance welfare.

*C) Liquidity Backstops*

While the next section discusses structural changes to the regulatory environment that aim at reducing the fragility of the shadow banking system, and ultimately at reducing

the amplitude of the leverage cycle, this subsection focuses on *ex post* policies that were deployed during the financial crisis of 2007-09 to stabilise the collapse of shadow banking. This collapse reflected problems on the asset and liability side—in other words, both credit problems and liquidity problems. While the Federal Reserve initiated programmes that primarily aimed at liquidity injections, the U.S. Treasury's programmes sought to resolve credit problems.

The financial crisis started in 2007 with the collapse of ABCP conduits and SIVs. The majority of those conduits were single-seller conduits that were under the umbrella of particular BHCs. As the asset quality of the ABCP conduits and SIVs deteriorated, money market investors withdrew funding, forcing the sponsoring BHCs to seek other sources of funding. As a result, assets from the conduits and SIVs were re-intermediated and funded in unsecured markets, such as the Libor market. As a result, funding in the Libor market became disrupted, and the Federal Reserve initiated two programmes to address the funding liquidity shortage. The Term Auction Facility (TAF) provided term funding for commercial banks, effectively replacing the term funding that evaporated in the ABCP market (see Armantier, Krieger, and McAndrews (2008) for a description of the TAF). In addition, foreign exchange swaps provided term funding for foreign banks that did not have access to the TAF via the Fed's discount window. Institutions that held U.S. dollar assets that could no longer be funded in the ABCP market were thus able to obtain funding in foreign currencies, and swap into dollar funding at the foreign central bank by using the Fed's foreign exchange swaps.

The next dysfunction during the financial crisis occurred in the repo market. In early 2008, haircuts in the DVP repo market started to increase substantially, leading to forced deleveraging of many fixed-income and credit hedge funds. For example, Carlyle's CCC fund, which invested in agency mortgages and was funded in the DVP repo market, had to declare bankruptcy in February 2008. Subsequently, the repo funding shortage also impacted the tri-party repo market. In the week of March 11, 2008, Bear Stearns was no longer able to obtain tri-party repo funding. In order to prevent these funding difficulties from spreading to other institutions, the Federal Reserve introduced the Primary Dealer Credit Facility (PDCF) on March 15, 2008 (see Adrian, Burke, and McAndrews (2009) for a detailed exposition of the facility). The PDCF allowed primary dealers to obtain funding from the Fed and thus effectively allowed the dealer sector to join depository institutions in having access to last-resort lending. In addition, the Term Security Lending Facility (TSLF) allowed dealers to exchange agency mortgage collateral against Treasury collateral (see Fleming, Hrung, and Keane (2010) for detail on the TSLF).

Following Lehman's bankruptcy on September 15, 2008, money markets experienced a run, leading to funding shortages of ABCP, CP, and repo issuers. The Federal Reserve introduced two facilities to address these money market dislocations: the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF) and the Commercial Paper Funding Facility (CPFF). The CPFF offered a funding source to commercial paper issuers that replaced money market funding in the aftermath of Lehman. The CPFF was constructed to be a self-liquidating facility due to terms that made it attractive during the financial crisis when spreads were unusually large but that

were uneconomical in a more normal spread environment. (Adrian, Kimbrough, and Marchioni (2011) describe the CPFF in greater detail.)

The Term Asset-Backed Securities Loan Facility (TALF) was created to help market participants meet the credit needs of households and small businesses by supporting the issuance of asset-backed securities (ABS) collateralised by auto loans, student loans, credit card loans, equipment loans, floorplan loans, insurance premium finance loans, loans guaranteed by the Small Business Administration, residential mortgage servicing advances, or commercial mortgage loans. Ashcraft, Malz, and Pozsar (2012) describe the facility in detail.

These liquidity facilities have the commonality of expanding the central bank's lending of last resort to institutions of the shadow banking system that do not have direct, explicit access to public liquidity backstops. The fragility of shadow banks due to vulnerable assets and liabilities makes them vulnerable to excessive collapse in times of adverse financial cycles. Lending of last resort aims to insulate real economic activity from such disorderly collapses. While lending of last resort is a necessary action *ex post*, once crises materialise, the anticipation of such action can result in distorted risk-taking incentives. As a result, structural reforms, reviewed in the next section, aim at mitigating incentives for excessive risk taking in the shadow banking system *ex ante*.

## 5) How Should Shadow Credit Intermediation Be Regulated?

Adrian and Ashcraft (2012) review recent regulatory changes to the shadow banking system in detail. We provide a short overview of those reform efforts, focusing on three areas: 1) reforms relating to money markets, 2) implications of banking regulation for the shadow banking system, and 3) reforms of securitisation and credit ratings. It should be noted that all these reform efforts are under way at this time.

### *A) Money Market Reforms: ABCP, Repo, and Money Market Mutual Funds*

#### Reforms Relating to ABCP Conduits

In June 2009, the Financial Accounting Standards Board (FASB) announced the Statements of Financial Accounting Standards (FAS) 166 and 167, amending existing accounting rules for consolidation of securitisation transactions. Sponsors of securitisation transactions have generally interpreted this new guidance as requiring accounting consolidation in the event that a first-loss position and loan servicing are retained by the sponsor for securitisation transactions.

Following revisions to the accounting rules, the U.S. banking agencies clarified in September 2009 that depository institutions would have to hold regulatory capital against consolidated securitisation transactions and ABCP conduits. The movement of assets onto the balance sheet will result in an increase in capital requirements under the minimum leverage ratio, an increase in risk-weighted assets and capital requirements given the inability of banks to use a 10 percent credit conversion factor for liquidity guarantees, and a requirement that banks provision for losses on loans held in consolidated conduits and securitisation trusts. The close link between regulatory capital and accounting treatment has eliminated the scope for using securitisation of loans serviced by the sponsor to reduce capital requirements. Furthermore, Section 331 of the Dodd-Frank Act requires FDIC assessments on consolidated assets minus tangible

equity of large banks rather than the historical practice of counting only deposit liabilities. The consolidation of conduits onto bank balance sheets means that banks will pay assessments on these liabilities, making conduit sponsorship more expensive.

#### Reforms Relating to Tri-party Repo

Reforms in the tri-party repo market are ongoing. An important friction in the tri-party repo market is the dependence of market participants on intraday credit of the custodian banks. In 2009, an industry task force sponsored by the Federal Reserve Bank of New York was created with the aim of reducing market participants' dependence on intraday credit. The task force has shortened the window of the daily unwind, moving it from 8:30 in the morning to 3:30 in the afternoon. However, between 3:30 and the settlement of all repos, the dealers are still dependent on the credit of the clearing banks.

Another major source of systemic risk in the tri-party repo market is vulnerability to the default of a major dealer. Such an event exposes that clearing bank to counterparty credit risk. Moreover, it leads to a potentially destabilising transfer of risk across market participants and directly impacts the dealers' clients who are no longer able to obtain leverage through the dealer in question. The vulnerability of short-term funding markets with respect to single institutions is a major concern for the stability of these funding markets. The tri-party repo task force has not been successful in identifying a solution to the problem of how money market fund investors would be able to liquidate collateral in the event a large broker-dealer became insolvent. As long as the tri-party repo market accepts a significant amount of collateral other than U.S. Treasury and agency securities (such as private-label ABS and corporate bonds), the tri-party market will remain prone to runs and constitute a source of systemic risk.

The major broker-dealers that are the most important borrowers in the tri-party repo market have become subject to tighter regulation. In particular, one of the consequences of the financial crisis has been that two of the formerly five major investment banks have been transformed into bank holding companies and two have merged with bank holding companies. The fifth dealer, Lehman Brothers, declared bankruptcy, and its dealer subsidiary was acquired by foreign banks. As a result, all of the formerly major independent investment banks are now regulated on a consolidated basis by the Federal Reserve and will be subject to the reformed Basel capital and liquidity standards. In addition, the Dodd-Frank Act instituted enhanced prudential standards for large bank holding companies and the designation of "systemically important nonbank financial institutions." Furthermore, the Orderly Liquidation Authority provides the FDIC with the authority to act as receiver for the resolution of nonbank financial institutions (including bank holding companies) for which a systemic risk determination has been made. A currently open question concerns the regulation of the major U.S. broker-dealers owned by foreign banking organisations.

#### Reforms Relating to MMMFs

MMMFs have undergone some reform since the financial crisis of 2007-09. In particular, the SEC has put new restrictions on 2a-7 funds to limit risk and maturity transformation and reliance on ratings. However, these restrictions do not address the key friction that exists in the market, which is implicit support for a stable net asset value

(NAV) by plan sponsors and the official sector through historical experience. The MMMF rules as amended in 2010 also increase the funds' incentives to lend for short tenors and decrease their incentives to look through to the collateral. The SEC rules incent MMMFs to act as unsecured rather than secured investors—which is a problem from a financial stability point of view. However, these reforms continue to leave MMMFs as a source of systemic risk.

The susceptibility of MMMFs to runs is illustrated in the aftermath of Lehman's failure in September 2008 when a stand-alone MMMF—the Reserve Fund—broke the buck. In the following weeks, institutional investors broadly withdrew from MMMFs, thus forcing massive liquidations of MMMF assets, which led to downward pressure on prices, and instances of funds breaking the buck. When investors run on money market funds, they are forced to liquidate assets, putting downward pressure on the mark-to-market values of the money market mutual fund assets and potentially leading to more funds breaking the buck. The market friction that makes MMMFs unstable is thus the stable NAV that gives rise to fire sale dynamics that occur when investors withdraw investments.

As a result of the stable NAV rule, investors of MMMFs effectively treat the funds like demand deposit accounts. In fact, many MMMFs market the funds as alternatives to demand deposits. However, MMMFs have no explicit backstop that would protect them against declines in asset values. MMMFs do rely on implicit discretionary support by sponsors. However, in a financial crisis, investors cannot necessarily count on the sponsors to provide support. The second friction we would highlight in the context of MMMFs is thus the implicit guarantee provided by the funds' sponsors.

One of the proposals for further reforms of MMMFs is to abandon the stable NAV rule and operate MMMFs with floating NAV. Money funds would then mark their asset values to market at all times. This would remove some of the incentives for investors to run. However, changing money funds from stable to floating NAVs would not remove all incentives of investors to run. In the presence of some illiquidity in the asset market, early withdrawal of funds can lead to temporary under-valuations of assets and provide incentives for early withdrawal. There are, indeed, some instances of runs in certain European countries that have money market funds with floating NAVs.

he SEC is currently considering a range of reform options, and the Financial Stability Oversight Council (FSOC) has highlighted the need for MMMF reforms in its annual reports of 2011 and 2012. In general, these were intended to address the fact that MMMFs have a number of characteristics—including a stable NAV, redemption upon demand, and extremely risk-averse investors—that interact to make these entities vulnerable to runs. Several of these proposals entail the creation of liquidity and capital buffers. The former provide additional near-cash assets to deal with redemptions, while the latter enhances the loss absorption capacity available to deal with a credit event. Broadly speaking, two kinds of buffers can be set up: *ex ante* and *ex post*.

One type of *ex ante* buffer is to create a private emergency liquidity facility, capital reserve, or insurance. Regulated fixed NAV funds would benefit from an *ex ante* buffer, but be forced to pay the cost. Another approach to an *ex ante* buffer is for individual funds to set aside resources in advance to absorb losses should they occur, serving the same purpose as capital reserves in traditional banks. As an alternative, the Investment

Company Institute has proposed a private-sector "liquidity bank" that would provide a backstop but might itself benefit from access to official liquidity.

An *ex post* buffer does not require any resources to be set aside in advance, but is created by taking steps to ensure that investors absorb losses when they occur and that they cannot flee and leave the losses behind. In particular, such measures are designed to forestall investors redeeming shares at a NAV of one dollar once a credit event or liquidity event has begun. A variable NAV may be helpful in this regard, because it could, if properly computed, adjust rapidly in response to losses or liquidity shocks. However, this would be a fundamental change in the nature of MMMFs.

The Squam Lake Group (2011) put forward a proposal for MMMFs to have two share classes. The senior tranche would be a stable net asset value fund that would be backed by a liquidity buffer amounting to $x$ percent of the current NAV. The liquidity buffer could be implemented in at least four different ways. Most recently, McCabe *et al.,* (2012) has investigated the feasibility of requiring a two-share system for MMMFs (labelled A and B shares). McCabe calibrates the returns to the tranches under realistic scenarios about asset returns and run risk. The advantage of this two-share proposal is that it would preserve the stable NAV feature of money market funds under much more severe circumstances than is currently the case, while lowering the returns to investors into the stable NAV shares only slightly. However, it should be noted that the two-share proposal does not fully protect funds against runs in all states of the world. For any realistic capital requirement, there are some tail events that will induce the fund to have to unwind. McCabe *et al.,* (2012) have proposed that MMMFs be made subject to a "minimum balance at risk" (MBR). The MBR would be a small fraction of each MMMF investor's balance demarcated to absorb losses if the fund is liquidated. This feature accounts for the credit risk that MMMFs hold. Furthermore, redemptions of the MBR would be delayed by thirty days, thus accounting for the illiquidity of MMMF assets in times of crisis and reducing fire sale incentives. Large redemptions would subordinate part of the MBR, creating a disincentive to redeem if the fund is likely to have losses and thus reducing incentives for investors to run.

### *B) Banking Regulation Reforms: Capital and Liquidity*

#### Capital Requirements for Securitisation Exposures

In February 2011, regulators announced planned changes to the treatment of securitisation exposures held by banks in the trading book. In general, assets held in the trading book face lower capital charges than those in the banking book given the stated intent of the institution to actively trade, and the presumption of regulators was that the institution will be able to exit the position before incurring credit losses. However, the behaviour of banks during the recent financial crisis suggested that these institutions were unwilling to trade out of positions, given the large decline in prices relative to projected losses. The proposed revisions to the Market Risk Amendment of Basel II recognise this behaviour and require banks to hold capital against securitisation exposures in the trading book as if they were in the banking book, eliminating banks' ability to hold less capital against these exposures.

#### FDIC Safe Harbour

In September 2010, the FDIC approved revisions to its safe harbour from repudiation powers in receivership. In particular, as receiver of a failed bank, the FDIC has the authority to repudiate contracts, which could possibly include the sale of assets to a bankruptcy-remote trust as part of a bank-sponsored securitisation transaction. Historically, the FDIC created a safe harbour from use of this authority tied to the accounting treatment of the transaction. However, the aforementioned changes to FAS 166/167 implied that many securitisation transactions would now be consolidated on a bank's balance sheet, implying that investors would no longer benefit from the existing safe harbour.

In the new safe harbour, the FDIC requires bank-sponsored securitisations to meet minimal standards for capital structure; disclosure requirements to be aligned with the SEC's proposed revisions to Regulation AB; and documentation, compensation, and risk retention to be aligned with the inter-agency implementation of Dodd-Frank 941. The rule has more stringent requirements for bank-sponsored RMBS transactions, including the need for a 5 percent cash reserve for twelve months to fund representations and warranties and a requirement that compensation to rating agencies be based in part on the performance of the underlying transactions. The stated motivation for using the safe harbour in this fashion is to protect the FDIC as guarantor of bank deposits from the bank's investment in securitisation transactions. As the scope of the rule applies only to banks sponsoring securitisation transactions, it is possible that, when binding, it will shift securitisation activity to the nonbank sector.

#### Bank Liquidity Regulation

In December 2010, the Basel Committee proposed new liquidity requirements for banks. In particular, in addition to capital requirements, banks would have to meet two liquidity standards: a liquidity coverage ratio (LCR) and a net stable funding ratio (NSFR). The LCR is intended to promote short-term resilience of a bank's liquidity risk profile by ensuring that it has sufficient high-quality liquid assets to survive a significant stress scenario lasting for one month. In particular, the bank is required to hold unencumbered high-quality liquid assets in an amount no less than 100 percent of total net cash outflows over the next thirty days in a stress scenario. The NSFR is intended to promote resilience over a longer time horizon by creating additional incentives for banks to fund their activities with more stable sources on an ongoing basis. In the NSFR requirement, stable funding is defined as "the portion of those types and amounts of equity and liability financing expected to be reliable sources of funds over a one-year time horizon under conditions of extended stress." The amount of required stable funding is a function of the liquidity characteristics of the institution's financial exposures. Collectively, these liquidity rules are expected to have an impact on the costs of providing liquidity guarantees to ABCP conduits, as banks will now be required to hold an adequate level of unencumbered high-quality liquid assets for draws on lines underlying the exposures in the conduits, as well as any ABCP with a maturity of thirty days or less. Moreover, new proposed liquidity requirements for banks could make backup lines more expensive by requiring an adequate level of liquid assets to meet stress liquidity needs for a thirty-day time horizon.

*C) Credit Market Reforms: Securitisation and Credit Ratings*

**FDIC Safe Harbour**

In April 2010, the SEC proposed revisions to Regulation AB that provide guidance on required disclosure by sponsors of securitisation transactions. The proposal by the SEC was largely confirmed in Section 942 of the Dodd-Frank Act. These rules were re-proposed in April 2011, partly in response to the Section 932A requirement to remove references to credit ratings and partly in response to comments on the original proposal. The motivation for revisions to the rule is the conclusion that investors did not have adequate information or time to conduct due diligence on new issue securitisation transactions, that market participants over-relied on credit ratings, and that incentive misalignment exists between sponsors and investors.

**Risk Retention**

In April 2011, regulators jointly proposed rules implementing Section 941 of Dodd-Frank, requiring that sponsors retain meaningful risk of securitisation transactions. In the proposal, the sponsor of a securitisation transaction is required to hold at least 5 percent of an eligible form of risk retention measured using par value. Eligible forms of risk retention generally include vertical retention, where the sponsor retains a fraction of every tranche; horizontal retention, where the sponsor retains a first-loss position; and a specific combination of the two forms of equal size, referred to as L-shaped. The sponsor is not permitted to sell or hedge the retained interest for the life of the transaction and is not permitted to pledge for nonrecourse financing. The proposed rule provides for exemptions from risk retention for securitisations sponsored by U.S. government agencies, for government-sponsored enterprises as long as they are in receivership, and for qualified loan pools that meet strict underwriting requirements. The proposed rule permits the sponsor of a CMBS transaction to sell a horizontal tranche to a B-piece investor that re-underwrites every loan in the transaction and permits the sponsor of an ABCP transaction to recognise risk retention by the originator in the underlying receivables being financed in the conduit.

**Credit Rating Agencies**

The 2010 Dodd-Frank Act includes a range of provisions intended to improve rating agency incentives and performance. Under Sections 7 and 11 of the Securities Act of 1933, when an issuer includes statements in a prospectus from experts like lawyers or accountants, the prospectus must also include consent to liability from the expert. While Rule 436(g) historically exempted credit rating agencies from this requirement, this exemption was removed by Dodd-Frank. However, the exemption has never gone into effect because the SEC issued a no-action letter, acknowledging refusal by the credit rating agencies to consent to expert liability, threatening to bring new issuance to a halt. As of this writing, the repeal of 436(g) is still not in force.

Dodd-Frank made amendments to Rule 17g-5 in order to provide investors with more views on the creditworthiness of structured finance products and to improve the quality of ratings by limiting rating shopping. In particular, these amendments require a rating agency hired by an issuer to disclose the rating assignment and obtain representation from the arranger that they will provide information to both hired

and certified non-hired nationally recognised statistical rating organisations (NRSROs). In practice, this provision has not been used to produce many shadow ratings, given concerns by NRSROs about legal liability they would face from disclosing this information without explicit permission of the issuer.

The Credit Rating Agency Reform Act of 2006 (Rating Agency Act) mandated that the SEC establish a registration and oversight programme for NRSROs. While the SEC was given formal oversight authority for the rating agencies for the first time, the law expressly prohibited regulation of the substance of credit ratings or the procedures and methodologies. In June 2007, the SEC adopted new rules establishing a regulatory programme for NRSROs. These rules require NRSROs to have written policies and procedures to prevent the misuse of material nonpublic information and to manage certain conflicts of interest. They also disallow certain other conflicts of interest outright and prohibit NRSROs from engaging in certain unfair, coercive, or abusive practices.

## 6) Conclusion

In this paper, we define shadow credit intermediation to be the intermediation of credit without the direct or explicit support of the U.S. taxpayer. Over the past sixty years, the fraction of credit funded using commercial banks' maturity transformation has declined significantly, as market-based credit intermediation has increased in importance. The decline of credit intermediation financed through traditional banks has been offset only in part by the rise of shadow banks.

We explore three motivations for the growing importance of shadow banking. First, we highlight financial innovation in the composition of the aggregate money supply. Second, we discuss the incentives of financial institutions to avoid taxes, accounting rules, or capital requirements. Third, we review the presence of agency problems in financial markets, which create perverse incentives.

We provide a topology for understanding different parts of the shadow intermediation process as well as different sectors that engage in shadow banking. We furthermore articulate the market failures that ultimately justify the need for regulation of this sector. In particular, we discuss how complexity can result in neglected risk by investors, permitting the buildup of systemic risks. Moreover, we note the well-documented externalities associated with runs on institutions involved in maturity transformation. We also document how the public sector has provided lending-of-last-resort facilities during the financial crisis in order to shield real economic activity from a run on the shadow banking system.

Finally, we provide an overview of new rules targeting shadow banking and highlight the uneven impact they will have on the likely size of the shadow banking sector in the future. While changes to accounting and capital requirements will reduce incentives by banks to engage in types of arbitrage activities at the core of the financial crisis—that is, ABCP and securitisation activity—significant increases in the overall level and risk sensitivity of capital will provide strong incentives for credit intermediation to be funded outside of the banking system.

### *Bibliography*

Acharya, V., Schnabl, P. and Suarez, G. 2010. Securitisation Without Risk Transfer. NBER Working Paper 15730.

Acharya, V., Richardson, M., Nieuwerburgh, S. V. and White, L. J. 2011. Guaranteed to Fail: Fannie Mae, Freddie Mac, and the Debacle of Mortgage Finance. Princeton University Press.

Adrian, T. and Ashcraft, A. B. 2012. Shadow Bank Regulation. *Annual Review of Financial Economics*. Forthcoming.

Adrian, T. and Boyarchenko N. 2012. Intermediary Leverage Cycles and Financial Stability. *Federal Reserve Bank of New York Staff Reports*, 567.

Adrian, T., Begalle, B., Copeland, A. and Martin, A. 2013. Repo and Securities Lending. In: *Quantifying Systemic Risk Measurement: NBER Research Conference Report Series*, (eds. J. G. Haubrich and A. W. Lo). University of Chicago Press. Forthcoming.

Adrian, T., Burke, C. and McAndrews, J. 2009. The Federal Reserve's Primary Dealer Credit Facility. Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 15(4), 1–12.

Adrian, T. and Fleming, M. 2005. What Financing Data Reveal About Dealer Leverage. Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 11(3), 1–7.

Adrian, T., Moench, E. and Shin, H. S. 2009. Financial Intermediation, Asset Prices, and Macroeconomic Dynamics. Federal Reserve Bank of New York *Staff Report*, 422.

Adrian, T., Kimbrough, K. and Marchioni, D. 2011. The Federal Reserve's Commercial Paper Funding Facility. Federal Reserve Bank of New York *Economic Policy Review*, 17(1), 25–39.

Adrian, T., Shin and H. S. 2009. The Shadow Banking System: Implications for Financial Regulation. *Banque de France Financial Stability Review*, 13, 1–10.

Adrian, T., Shin and H. S. 2010a. Liquidity and Leverage. *Journal of Financial Intermediation*, 19(3), 418–37.

Armantier, O., Krieger, S. and McAndrews, J. 2008. The Federal Reserve's Term Auction Facility. Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 14(5).

Ashcraft, A., Goldsmith-Pinkham, P., Hull, P. and Vickery, J. 2011. Credit Ratings and Security Prices in the Subprime MBS Market. *American Economic Review*, 101(3), 115–19.

Ashcraft, A. B. and Schuermann, T. 2008. Understanding the Securitisation of Subprime Mortgage Credit. *Foundations and Trends in Finance*, 2(3), 191–309.

Ashcraft, A. B., Malz, A. and Pozsar, Z. 2012. The Federal Reserve's Term Asset-Backed Securities Loan Facility. Federal Reserve Bank of New York *Economic Policy Review*. Forthcoming.

Avraham, D., Selvaggi, P. and Vickery, J. 2012. A Structural View of Bank Holding Companies. Federal Reserve Bank of New York *Economic Policy Review*, 18(2), 65–82.

Bord, V. and Santos, J. C. 2012. The Rise of the Originate-to-Distribute Model and the Role of Banks in Financial Intermediation. Federal Reserve Bank of New York *Economic Policy Review*, 18(2), 21–34.

Brunnermeier, M. K. and Pedersen, L. H. 2009. Market Liquidity and Funding Liquidity. *Review of Financial Studies*, 22(6), 2201–2238.

Brunnermeier, M. K. and Sannikov, Y. 2012. A Macroeconomic Model with a Financial Sector. Working Paper. Princeton University.

Cetorelli, N. 2012. A Principle for Forward-Looking Monitoring of Financial Intermediation: Follow the Banks! Federal Reserve Bank of New York *Liberty Street Economics Blog*, July 23, 2012. Available at http://libertystreeteconomics.newyorkfed.org/2012/07/a-principle-for-forward-looking-monitoring-of-financial-intermediation-follow-the-banks.html.

Cetorelli, N. and Peristiani, S. 2012. The Role of Banks in Asset Securitisation. Federal Reserve Bank of New York *Economic Policy Review*, 18(2), 47–64.

Cohen, A. 2011. Rating shopping in the CMBS market. Presented at Regulation of Systemic Risk, Washington, DC, Sep. 15–16.

Copeland, A., Martin, A. and Walker, M. 2011. Repo Runs: Evidence from the Tri-party Repo Market. Federal Reserve Bank of New York *Staff Report*, 506.

Copeland, A. 2012. Evolution and Heterogeneity among Larger Bank Holding Companies: 1994 to 2010. Federal Reserve Bank of New York *Economic Policy Review*, 18(2), 83–93.
Coval, J., Jurek, J. and Stafford, E. 2009. The Economics of Structured Finance. *Journal of Economic Perspectives*, 23(1), 3–25.
Covitz, D., Liang, N. and Suarez, G. 2012. The Evolution of a Financial Crisis: Panic in the Asset-backed Commercial Paper Market. *Journal of Finance*. Forthcoming.
Dang, T. V., Gorton, G. and Holmström, B. 2009. Opacity and the Optimality of Debt for Liquidity Provision.
Yale/MIT Working Paper.
Diamond, D. and Dybvig, P. 1983. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy* 91:401–19.
Duffie, D. 2012. Market Making Under the Proposed Volcker Rule. Working Paper. Stanford University.
Financial Accounting Standards Board. 2012. Briefing Document: FASB Statement 166 and 167. Available at http://www.fasb.org/cs/ContentServer?c=FASBContent_ C&pagename=FASB% 2FFASBContent_C%2FNewsPage&cid=1176155633483.
Financial Stability Board. 2011. Shadow Banking: Strengthening Oversight and Regulation. Available at http://www.financialstabilityboard.org/publications/r_111027a.pdf.
Financial Stability Board. 2012. Progress Report to the G20 on Strengthening the Oversight and Regulation of Shadow Banking. Available at http://www.financialstabilityboard.org/publications/ r_120420c.pdf.
Fleming, M. and Garbade, K. 2003. The Repurchase Agreement Refined: GCF Repo. Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 9(6), 1–7.
Fleming, M., Hrung, W. B. and Keane, F. M. 2010. Repo Market Effects of the Term Securities Lending Facility. *American Economic Review: Papers and Proceedings*, 100, 591–6.
Fostel, A. and Geanakoplos, J. 2008. Leverage Cycles and the Anxious Economy. *American Economic Review*, 98(4), 1211–1244.
Frame, W. S. and White, L. J. 2005. Fussing and Fuming over Fannie and Freddie: How Much Smoke, How Much Fire? *Journal of Economic Perspectives*, 19(2), 159–184.
Garleanu, N. and Pedersen, L. H. 2011. Margin-Based Asset Pricing and Deviations from the Law of One Price, *Review of Financial Studies*, 24(6), 1980–2022.
Gennaioli, N., Shleifer, A. and Vishny, R. 2012a. Neglected Risks, Financial Innovation, and Financial Fragility. *Journal of Financial Economics*. Forthcoming.
Gennaioli, N., Shleifer, A. and Vishny, R. 2012b. A Model of Shadow Banking. *Journal of Finance*. Forthcoming.
Gorton, G. 1985. Clearinghouses and the Origin of Central Banking in the United States. *Journal of Economic History*, 4(2), 277–283.
Gorton, G. and Metrick, A. 2011. Regulating the Shadow Banking System. Brookings Paper on Economic Activity, 261–312.
Gorton, G. and Metrick, A. 2012. Securitised Banking and the Run on Repo. *Journal of Financial Economics*, 104, 425–51.
Holmström, B. and Tirole, J. 1998. Private and Public Supply of Liquidity. *Journal of Political Economy*, 106(1), 1–40.
International Institute of Finance. 2012. Shadow Banking: A Forward-Looking Framework for Effective Policy.
Jaremski, M. 2010. Free Bank Failures: Risky Bonds Versus Undiversified Portfolios. *Journal of Money, Credit, and Banking*, 42(8), 1565–1587.
Kacperczyk, M. and Schnabl, P. 2011. Does Organizational Form Affect Risk Taking? Evidence from Money Market Mutual Funds. New York University Working Paper.
Keys, B., Mukherjee, T., Seru, A. and Vig, V. 2010. Did Securitisation Lead to Lax Screening? Evidence from Subprime Loans. *Quarterly Journal of Economics*, 125(1), 307–62.

Krishnamurthy, A., Nagel, S. and Orlov, D. 2011. Sizing Up Repo. NBER/CEPR/Stanford/Northwest Working Paper.
Levitin, A. and Wachter, S. 2011. Explaining the Housing Bubble. Research Paper 10-15, Institute for Law and Economics, University of Pennsylvania.
Mandel, B., Morgan, D. and Wei, C. 2012. The Role of Bank Credit Enhancements in Securitisation. Federal Reserve Bank of New York *Economic Policy Review*, 18(2), 35–46.
Martin, A., Skeie, D. and von Thadden, E. 2011. Repo Runs. Federal Reserve Bank of New York *Staff Report* 444.
Mathis, J., McAndrews, J. and Rochet, J. C. 2009. Rating the raters: Are reputation concerns powerful enough to discipline rating agencies? *Journal of Monetary Economics*, 57(5), 657–74.
McCabe, P. 2011. An A-/B-share capital buffer proposal for money market funds. Working Paper.
McCabe, P. E., Cipriani, M., Holscher, M. and Martin, A. 2012. The Minimum Balance at Risk: A Proposal to Mitigate the Systemic Risks Posed by Money Market Funds. Federal Reserve Bank of New York *Staff Report*, 564.
McCulley, P. 2007. Teton Reflections. *PIMCO Global Central Bank Focus*.
Merton, R. C. 1977. An Analytical Derivation of the Cost of Deposit Insurance and Loan Guarantees. *Journal of Banking and Finance*, 1, 3–11.
Merton, R. C. and Bodie, Z. 1993. Deposit Insurance Reform: A Functional Approach. Carnegie-Rochester Conference Series Public Policy 38, 1–34.
Passmore, W., Sherlund, S. M. and Burgess, G. 2005. The Effect of Housing Government-Sponsored Enterprises on Mortgage Rates. *Real Estate Economics*, 33(3), 427–463.
Pozsar, Zoltan. 2008. The Rise and Fall of the Shadow Banking System. Available at http://www.economy.com/sbs.
Pozsar, Z., Adrian, T., Ashcraft, A. B. and Boesky, H. 2010. Shadow Banking. Federal Reserve Bank of New York *Staff Report*, 458. Available at http://www.newyorkfed.org/research/staff_reports/sr458.pdf.
Rajan, R. 2005. Has Financial Development Made the World Riskier? Proceedings of the Federal Reserve Bank of Kansas City Economics Symposium, 313–69.
Ricks, M. 2010. Shadow Banking and Financial Regulation. Columbia Law and Economics Working Paper 370.
Rockoff, H. 1991. "Lessons from the American Experience with Free Banking" in Unregulated Banking: Chaos or Order?" edited by Forrest Capie and Geofrey Wood. London: Macmillan Academic and Professional.
Schwarcz, S. 2012. Regulating Shadow Banking. Boston University *Review of Banking and Financial Law*.
Squam Lake Group. 2011. Reforming Money Market Funds. Squam Lake Group Staff Report.
Stein, J. 2010. Securitisation, Shadow Banking, and Financial Fragility. *Daedalus*, 139(4), 41–51.
Sunderam, A. 2012. Money Creation and the Shadow Banking System. Harvard Business School Working Paper.
Tucker, P. 2010. Shadow banking, financing markets and financial stability. Remarks by Mr Paul Tucker, Deputy Governor for Financial Stability at the Bank of England, at a Bernie Gerald Cantor (BGC) Partners Seminar, London, 21 January 2010.
Wermers, R. 2011. Runs on Money Market Mutual Funds. University of Maryland Working Paper.
Wiggers, T. and Ashcraft, A. B. 2012. Defaults and Losses on Commercial Real Estate Bonds during the Great Depression Era. Federal Reserve Bank of New York *Staff Report*, 544.
Xia, H. and Strobl, G. 2012. The Issuer-Pays Rating Model and Ratings Inflation: Evidence from Corporate Credit Ratings. SSRN Working Paper.

## speculative bubbles

We maintain that a speculative bubble exists if the market price of an asset differs from its fundamental value – the expected present value of the stream of future dividends attached to the asset. In an economy with a finite sequence of trading dates, the fundamental theorem of asset pricing (see Dybvig and Ross, 1987) guarantees that the equilibrium market price of any asset equals its fundamental value. But in some economies with an infinite sequence of trading dates, this result does not hold, and speculative bubbles may arise. An investor might buy an asset at a price higher than its fundamental value if she expects to sell it later on at a higher price – Harrison and Kreps (1978) call this process 'speculative behaviour'. In general equilibrium models, however, agents take prices as given and trade assets to transfer income across time and states. These models do not contemplate 'speculative behaviour' as it is usually understood. Therefore, the term 'speculative bubble' may seem inappropriate in some theoretical frameworks. Santos and Woodford (1997) talk broadly about 'asset pricing bubbles'.

There have been famous historical examples of sudden asset price increases followed by an abrupt fall as the Dutch 'tulipmania' (1634–7), the 'Mississippi bubble' (1719–20) and the 'South Sea bubble' (1720). Kindleberger (1978) argues that these are examples of bubbles, whereas Garber (2000) provides market-fundamental explanations for these episodes. More recently, we have seen sharp changes in stock and housing markets. The Japanese stock and land prices experienced a sharp rise in the late 1980s and a dramatic fall in the early 1990s. During the 'technology bubble', the Nasdaq Composite Index rose by more than 300 per cent between August 1996 and March 2000, and then fell sharply, reaching the August 1996 level in October 2002. This pattern has been especially intense for the Internet-related sector (Ofek and Richardson, 2003).

There is a vast literature following the variance-bound tests proposed by LeRoy and Porter (1981) and Shiller (1981) that finds significant excess volatility of stock prices (see Gilles and LeRoy, 1991, for a survey). The violation of these variance bounds suggests that asset prices are not determined by fundamental values (see Flood and Hodrick, 1990, and Cochrane, 1992 for a discussion). Various tests have been proposed to detect the presence of rational bubbles in asset prices (see Camerer, 1989, and Cuthbertson, 1996, for a survey). But these tests have important shortcomings. Estimating the fundamental values of an asset is usually a complex task. Hence, rejections of the null hypothesis could be due to an incorrect specification of the fundamental value and not necessarily to the existence of a bubble (Flood and Hodrick, 1990). Even in the most famous apparent bubble episodes, some authors have provided a fundamentalist explanation (see, for example, Donaldson and Kamstra, 1996; Pástor and Veronesi, 2006). To avoid the uncertainty associated with the specification of the fundamental value, Diba and Grossman (1988a) develop a test to detect bubbles based on the investigation of the stationary properties of asset prices and dividends. The main drawback of this test, as Evans (1991) shows, is its limited power to detect periodically collapsing bubbles. Given the severe problems in establishing empirically the existence

of bubbles, it is of great importance to understand the theoretical conditions under which bubbles may exist.

If all traders are rational, a backward induction argument precludes the existence of bubbles for assets traded at a finite sequence of dates. More specifically, assume that the economy ends at time $T$, and there is an asset that provides a dividend of $d_T$ at time $T$. Then the price of the asset at $T–1$ must be equal to the present value of $d_T$. By backward induction a bubble cannot exist at any point in time $t$ less that $T$. Hence, a rational bubble begins on the first date of trading. Moreover, in present value terms the size of the bubble must be constant. (This is usually called the martingale property of bubbles.) Diba and Grossman (1988b) argue that negative rational bubbles cannot exist because it would imply that investors expect that the price of the asset will become negative at a finite future date. Tirole (1982) concludes that, in an economy with a finite number of infinitely lived traders, any asset must be valued according to its market fundamental. However, Tirole (1985) shows that under certain circumstances a deterministic overlapping generations economy allows for the existence of bubbles. In infinite-horizon optimization economies, bubbles are not compatible with the transversality condition: the present value of optimal asset holdings must converge to zero. But by definition the discounted price of the asset will converge to the size of the bubble. Hence, either the asset is in zero net supply or the size of the bubble is equal to zero.

Santos and Woodford (1997) explore the existence of asset pricing bubbles in an infinite-horizon competitive framework, allowing for potentially incomplete markets, arbitrary borrowing limits and incomplete participation of agents (this framework considers jointly economies with a finite number of infinitely lived households and overlapping generations economies). They show that the price of any asset in positive net supply must be equal to its fundamental value, provided that the present value of aggregate wealth is finite. This latter condition is satisfied empirically (see Abel et al., 1989) since in industrialized economies the aggregate share of income that goes to capital is greater than the investment rate. Loewenstein and Willard (2000) extend these results to a finite horizon economy where assets are negotiated continuously. Some key conditions underlying the negative results of Santos and Woodford (1997) are rational expectations, symmetric information and competitive behaviour.

This analysis has important implications for monetary theory because it precludes the existence of valued fiat money as a store of wealth in a broad class of economies. Santos (2006) extends these results to an economy with liquidity constraints and proves that these constraints must be binding infinitely often for all agents in the economy. Hence, in his simple model the aggregate value of the money supply must be equal to the value of aggregate output infinitely often. This is in the spirit of the quantity theory of money. On a related matter, the absence of rational bubbles guarantees that the initial real value of public debt is equal to the present value of future net public revenues. This is a necessary condition to establish the validity of the 'fiscal theory of the price level' (Sims, 1994; Woodford, 1995).

The presence of bubbles has also been explored in theoretical frameworks with asymmetric information or boundedly rational agents. Allen, Morris and Postlewaite (1993) find necessary conditions for the existence of bubbles in a model with

asymmetric information and a finite sequence of trading dates, and provide examples satisfying these conditions. The existence of a bubble is possible because there is private information which is not common knowledge (all agents know that all agents know, and so on, ad infinitum) that the stock price will fall. Everybody realizes that the stock is overpriced but each agent expects to sell at a higher price before the true value becomes publicly known.

Bubbles may appear in the presence of agency problems associated with short-run optimization behaviour. Allen and Gorton (1993) show that for some compensation schemes a manager may purchase a stock with some prospect of capital gains although with certainty the price will fall below its current level at some point in the future. Allen and Gale (2000) develop a model in which intermediation by the banking sector leads to an agency problem that results in asset bubbles. Investors borrow from banks to buy a risky asset, and they can default in the case of low payoffs. Hence risky assets are more attractive, and therefore investors bid up asset prices.

The behavioural finance literature (see Barberis and Thaler, 2003; Shleifer, 2000 for a survey) often assumes that some agents – called noise traders – are not fully rational. In models in which noise traders and rational agents coexist, the price of an asset can deviate from its fundamental value if rational agents are limited in their capacity to eliminate the mispricing. Shleifer (2000) describes bubbles as the interaction between a significant number of positive feedback investors (who buy securities when prices rise and sell when prices fall), and rational arbitrageurs who anticipate the bursting of bubbles. In this framework, rational arbitragers buy initially after a good-news event to increase the price of the asset and to stimulate the demand of the positive feedback traders; later, they undo their position before the bubble explodes. Abreu and Brunnermeier (2003) develop a model in which noise traders coexist with rational arbitrageurs who become aware of the existence of a bubble sequentially. These rational arbitrageurs would like to exit the market just before the bubble bursts, because before bursting the asset displays high capital gains. The bubble can explode for exogenous reasons, or endogenously when a sufficient number of arbitrageurs decide to abandon the market. In this setting some news could facilitate synchronization and, as a consequence, the bursting of the bubble. Scheinkman and Xiong (2003) develop a model in which overconfidence generates disagreements among agents regarding asset fundamentals. They show that the price of an asset can be above its fundamental value.

In summary, asset prices seem rather volatile – more than their fundamental values. By definition this implies the existence of speculative bubbles. Most empirical exercises to detect the presence of bubbles seem inconclusive. The conditions under which general equilibrium models generate bubbles seem rather fragile, since optimizing agents are unwilling to accumulate arbitrary amounts of wealth. Most recent work has explored the existence of bubbles in economies with limited rationality, asymmetric information and strategic behaviour. The main challenge for these approaches is to explain the mechanisms that lead agents to hold overpriced assets. Specifically, if agents accumulate those assets for arbitrary reasons, then these exercises will not be very enlightening.

MIGUEL A. IRAOLA AND MANUEL S. SANTOS

*See Also* **arbitrage pricing theory; excess volatility tests; noise traders; present value.**

### *Bibliography*

Abel, A.B., Mankiw, N.G., Summers, L.H. and Zeckhauser, R.J. 1989. Assessing dynamic efficiency: theory and evidence. *Review of economic studies* 56, 1–20.

Abreu, D. and Brunnermeier, M.K. 2003. Bubbles and crashes. *Econometrica* 71, 173–204.

Allen, F. and Gale, D. 2000. Bubbles and crises. *Economic Journal* 110, 236–55.

Allen, F. and Gorton, G. 1993. Churning bubbles. *Review of Economic Studies* 60, 813–36.

Allen, F. Morris, S. and Postlewaite, A. 1993. Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory* 61, 206–29.

Barberis, N. and Thaler, R. 2003. A survey of behavioral finance. In *Handbook of the Economics of Finance*, vol. 1, ed. G.M. Constantinides, M. Harris and R.M. Stulz. Amsterdam: North-Holland.

Camerer, C. 1989. Bubbles and fads in asset prices. *Journal of Economic Surveys* 3, 3–41.

Cuthbertson, K. 1996. *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*. Chichester: Wiley & Sons Ltd.

Cochrane, J. 1992. Explaining the variance of price dividend ratios. *Review of Financial Studies* 5, 243–80.

Diba, B.T. and Grossman, H.I. 1988a. Explosive rational bubbles in stock prices. *American Economic Review* 78, 520–30.

Diba, B.T. and Grossman, H.I. 1988b. The theory of rational bubbles in stock prices. *Economic Journal* 98, 746–54.

Donaldson, R.G. and Kamstra, M. 1996. A new dividend forecasting procedure that rejects bubbles in asset prices: the case of 1929's stock crash. *Review of Financial Studies* 9, 333–83.

Dybvig, P.H. and Ross, S.A. 1987. Arbitrage. In *The New Palgrave: A Dictionary of Economics*, vol. 1, ed. J. Eatwell, M. Milgate and P. Newman. London: Macmillan.

Evans, G.W. 1991. Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review* 81, 922–30.

Flood, R.P. and Hodrick, R.J. 1990. On testing for speculative bubbles. *Journal of Economic Perspectives* 4(2), 85–101.

Garber, P.M. 2000. *Famous First Bubbles*. Cambridge, MA: MIT Press.

Gilles, C. and LeRoy, S. 1991. Econometric aspects of the variance-bounds tests: a survey. *Review of Financial Studies* 4, 753–91.

Harrison, J.M. and Kreps, D.M. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92, 323–36.

Kindleberger, C.P. 1978. *Manias, Panics and Crashes: A History of Financial Crises*. New York: Basic Books.

LeRoy, S.F. and Porter, R.D. 1981. The present value relation: test based on implied variance bounds. *Econometrica* 49, 555–74.

Loewenstein, M. and Willard, G.A. 2000. Rational equilibrium asset-pricing bubbles in continuous trading markets. *Journal of Economic Theory* 91, 17–58.

Ofek, E. and Richardson, M. 2003. DotCom mania: the rise and fall of internet stock prices. *Journal of Finance* 58, 1113–37.

Pástor, L. and Veronesi, P. 2006. Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81, 61–100.

Santos, M.S. 2006. The value of money in a dynamic equilibrium model. *Economic Theory* 27, 39–58.

Santos, M.S. and Woodford, M. 1997. Rational asset pricing bubbles. *Econometrica* 65, 16–57.

Scheinkman, J.A. and Xiong, W. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 11, 1183–219.

Shiller, R.J. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–36.

Shleifer, A. 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford: Oxford University Press.

Sims, C.A. 1994. A simple model of the determination of the price level and the interaction of monetary and fiscal policy. *Economic Theory* 4, 381–99.

Tirole, J. 1982. On the possibility of speculation under rational expectations. *Econometrica* 50, 1163–81.

Tirole, J. 1985. Asset bubbles and overlapping generations. *Econometrica* 53, 1499–528.

Woodford, M. 1995. Price level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy* 43, 1–46.

## South Sea bubble

The South Sea Company was founded in 1711, in the expectation that peace between Spain and England after the end of the War of the Spanish Succession would produce profitable trading opportunities with the 'South Seas' (that is, Spanish America). The company's trading activity remained intermittent and unprofitable throughout the 1710s. In 1719, a new scheme was launched – the conversion of government debt into equity of the South Sea Company. Debt-holders of the 1710 lottery loan were offered the option to convert their holdings into company shares. The government agreed to make interest payments to the company instead of to debt-holders. As old (and illiquid) loans were swapped for liquid company shares, debt-holders gained. The government negotiated a lower rate of interest, and the South Sea Company made a modest profit. The 1719 equity-for-debt swap is generally seen as Pareto-improving.

The 1720 conversion scheme differed in important ways. Key elements included (*a*) the absence of a fixed conversion ratio – higher prices of South Sea stock meant that more debt could be bought with each share, (*b*) issuance of new stock on instalment, with only a small down payment required, (*c*) massive lending against shares, and (*d*) a high degree of corruption in the awarding of the contract. The South Sea conversion also shared important characteristics with John Law's Mississippi scheme in France, which produced a similar run-up (and crash) of prices half a year earlier.

Both the Bank of England and the South Sea Company competed for the contract to convert government bonds into equity. After bribes to MPs, ministers, and members of the court (of about £1.3 million), the South Sea Company won the right to perform the conversion in March 1720. By this time, the price of its shares had increased to 255, from 128 at the beginning of the year. The share prices of other companies moved up and down in parallel with South Sea stock, but less sharply (see Figure 1). The company proceeded to issue fresh shares in four subscriptions, and offered to convert debt into shares on (modestly) generous terms. By late June, prices had risen to 765, and forward prices during the summer rose as high as 950. When regular trading resumed, prices began to weaken, but the fourth subscription was still strongly oversubscribed. In September, prices fell quickly. By the year end they had almost declined to their January level.

### Interpretations

Since Mackay's classic *Extraordinary Popular Delusions and the Madness of Crowds*, the South Sea bubble has often been cited as a prime example of irrational investor behaviour. In contrast, Peter Garber (2000) argued that the share prices increased in line with 'changing view(s) of market fundamentals'. If the scheme had succeeded in improving economic conditions in England as a whole (as John Law's logic would predict; Verde, 2004), the firm's large capital base might have allowed it to pursue profitable ventures. Yet most of these remained vague, and the company had no track record of successfully making money from anything other than financial transactions.

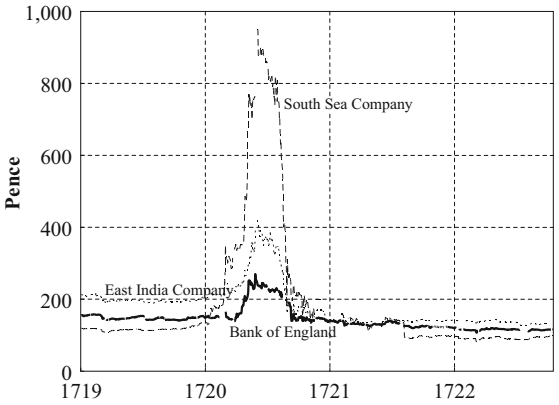



**Figure 1** Share price of major English companies, 1718–22. *Source*: Neal (1990); data from ICPSR, Study No. 1008.

It is doubtful whether future profits could ever have been high enough to justify the company's market capitalization in the summer of 1720. Even Garber accepts that prices above 400 are hard to square with reasonable expectations of future profits. Easy credit, investor preferences for lottery-like payoffs (as a result of shares being sold with only a small down payment), and restricted free float (caused by company lending against its own shares) may have contributed to the start of the bubble.

Recent work has focused on the reasons why the bubble, once under way, could have expanded greatly. Dale (2004) argues that apparent mispricing of subscription receipts proves investor irrationality, while others have argued that the gap can be explained by the option-like nature of receipts. Temin and Voth (2004) examined the trades of a goldsmith bank, Hoare's, which made large profits buying and selling South Sea stock in 1720. They argue that the bank was aware of the overpricing, but invested in South Sea stock regardless. Predictability of investor sentiment made it rational to 'ride' the bubble, and to sell out with a profit as soon as it began to deflate. This strategy is similar to hedge fund behaviour on Nasdaq in the late 1990s (Brunnermeier and Nagel, 2004). If other large investors faced similar incentives, the lack of a coordinated early attack becomes easier to understand. The role of market microstructure imperfections was probably limited, as opportunities to sell short were abundant. However, the nature of the settlement process and the artificial reductions of free float engineered by the company may have contributed to the bubble.

## Consequences

The rise and crash of share prices in 1720 had few direct economic consequences. As prices declined, former debt-holders demanded compensation. Parliament investigated the scheme in which it had played such an important role. Directors had most of their assets expropriated. In contrast to the resolution of the Mississippi bubble in France, those who had tendered government bonds for company shares received partial compensation in the form of fresh government debt. The political consequences were possibly more formidable than the immediate economic repercussions. Leading politicians

who had taken bribes, such as the Chancellor of the Exchequer, John Aislabie, were forced out of office and incarcerated. Robert Walpole, sometimes referred to as England's first prime minister, distinguished himself both through his opposition to the scheme and competent handling of its fallout. He succeeded Aislabie at the Exchequer and remained in power until 1742.

The collapse of the South Sea bubble is sometimes seen as a factor behind the Bubble Act. This appears to be erroneous, as the Act was passed before the bubble deflated (Carswell, 1993). Its passage and rigorous enforcement after the summer of 1720 probably owed more to the company's efforts to support its own sagging share price. Because of the Act, new equity issues became very rare for almost a century. The Act was repealed only in 1825.

HANS-JOACHIM VOTH

*See Also* **bubbles; Law, John.**

### *Bibliography*

Brunnermeier, M. and Nagel, S. 2004. Hedge funds and the technology bubble. *Journal of Finance* 59, 2013–40.

Carswell, J. 1993. *The South Sea Bubble*. Stroud: Sutton.

Dale, R. 2004. *The First Crash*. Princeton: Princeton University Press.

Garber, P. 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. Cambridge, MA: MIT Press.

Mackay, C. 1995. *Extraordinary Popular Delusions and the Madness of Crowds*. San Francisco: Essential Library.

Neal, L. 1990. *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason*. New York: Cambridge University Press.

Temin, P. and Voth, H.J. 2004. Riding the South Sea bubble. *American Economic Review* 94, 1654–68.

Verde, F. 2004. Government equity and money: John Law's system in 1720 France. Working paper, Federal Reserve Bank of Chicago.

# subprime mortgage crisis, the

## The rise and fall of subprime lending

A subprime mortgage loan is a residential mortgage loan that is particularly risky for some reason. The elevated risk may stem from the credit history of the borrower, the lack of a large down payment, or a monthly payment that is large relative to the borrower's income (see Chapter 2 of Muolo and Padilla (2010) for a history of subprime residential lending). Subprime loans were unlikely to meet the credit-quality standards of the two government-sponsored entities (GSEs), Fannie Mae and Freddie Mac, which package so-called prime mortgage loans into mortgage-backed securities for sale to outside investors, and which eliminate credit risk by guaranteeing principal and interest payments to investors if the borrower defaults. While the GSEs did not have hard-and-fast cutoffs regarding borrower credit quality, they were historically less likely to securitize loans made to borrowers with poor credit histories.

The subprime lending industry boomed in the 2000s as financial markets found ways to package subprime mortgages into marketable securities without the credit guarantees of the GSEs. The typical subprime deal offered investors a menu of securities with different income streams. Investors desiring the least risk would purchase the top "tranches" of the security, which generated highly predictable payments, while those willing to accept greater risk would purchase lower tranches. Any credit losses from the mortgages underlying the security would be allocated from the bottom up, so the top tranches would experience no credit losses unless the tranches below them were completely wiped out. Of course, investors holding the lower tranches were compensated for this risk with higher returns. Since most subprime loans ended up in securities, the typical subprime borrower was connected to his or her ultimate lender – a purchaser of the subprime security – through a long chain of financial intermediaries. A stylized description of the process might proceed as follows: a mortgage broker or loan officer would meet with the borrower and find a suitable mortgage product. A mortgage banker would originate this mortgage by supplying the money, perhaps after taking out a warehouse loan from a bank. The originating lender would then sell this mortgage to a Wall Street investment bank, using part of the proceeds to pay off the warehouse loan. The investment bank (such as Bear Stearns or Lehman Brothers) would assemble a large number of subprime mortgage loans into a tranched 'private-label' (that is, non-GSE) security for sale to investors around the world. This description is stylized because in practice different parties might play different roles at different times; a unit of an investment bank might originate and sell off one group of mortgages, while at the same time another part of the bank was packaging a different group of mortgages into a security. But the main organizing principle of the process was that intermediaries collected fees or other types of income, then passed the credit risk of their mortgages on to the next entity in the chain.

Loans made through this so-called 'originate-to-distribute' model rose and fell with the health of the US housing market. Using data from the Home Mortgage Disclosure Act (HMDA), Gramlich (2007, p. 6) notes that subprime originations totalled only $35 billion in 1994. By 2005, these originations reached $625 billion, representing 20% of all originations in that year and a 26% annual growth rate over the proceeding decade.

Even during the best of times, it was no secret that subprime mortgages performed much worse than prime mortgages. According to data collected by the Mortgage Bankers Association, the quarterly 90-day delinquency rate for subprime mortgages ranged between 2.3% to 3.6% in the boom years of 2003–2005. The comparable range for prime mortgages during the same period was 0.28% to 0.41%, nearly an order of magnitude smaller. Subprime lending was profitable despite the high delinquency rate, however, due in large part to the high interest rates that subprime mortgages carried.

Problems in the subprime market began to arise in late 2005 and 2006. By the end of 2007, the 90-day delinquency rate in the subprime market hit 5.4%. (This rate reached a maximum of 15.0% in the first quarter of 2010.) Sources of funds for subprime lending began to dry up: between 2005 and February 2007, 15 subprime lenders either ceased operations or offered themselves for sale (Barr, 2007). Troubles in the subprime market became a full-blown crisis in the summer of 2007. On 22 June, Bear Stearns pledged up to $3.2 billion to bail out one of its hedge funds because of subprime-related losses. A few weeks later, on 1 August, this and another fund owned by Bear Stearns declared bankruptcy. Eight days later, a French bank, BNP Paribas, suspended three of its funds because it could not value their mortgage investments. Continued losses on subprime lending contributed directly to the government-assisted sale of Bear Stearns to JPMorgan Chase on 16 March 2008, as well as the bankruptcy of Lehman Brothers six months later.

One of the key questions to arise from the subprime crisis was how the collapse of this market could have had such a big impact on the world's financial system. An answer comes by noting where these losses occurred. Adrian and Shin (2010), citing data in Greenlaw *et al.* (2008), write that the total exposure to subprime losses totalled about $1.4 trillion. Of this amount, fully two-thirds was borne by a 'leveraged sector' consisting of firms like commercial or investment banks, which borrow money to fund their investments. 'Thus, although securitization was meant to transfer credit risk to those who were better able to bear it, the balance-sheet management of financial intermediaries appears to have achieved the opposite outcome – of concentrating risks in the financial intermediary sector itself' (Adrian and Shin, 2010, p. 611).

During the past two decades, many economists have studied what happens when financial intermediaries like banks run into trouble. A deterioration of assets in these leveraged financial institutions reduces their capital cushions, which in turn can force them to shrink their balance sheets further. At banks, balance-sheet reductions are typically accomplished by a fall in lending, which adversely impacts the ability of creditworthy firms to fund their operations. Economic activity suffers as a result.

A second reason why the subprime crisis caused such trouble concerned the type of losses it generated. The wave of subprime defaults caused losses on subprime securities to extend into their intermediate ('mezzanine') tranches. Many of these tranches had been further securitized into collateralized debt obligations (CDOs), along with other

forms of debt (for example, student loans and credit card receivables). Like a mortgage-backed security, the motivating idea behind a CDO is diversification. It was thought to be extremely unlikely that all the constituent forms of debt in a CDO would go bad at the same time, just as it was assumed that there was no large source of aggregate default risk for the individual mortgages in a subprime security. Unfortunately, both assumptions turned out to be wrong. However, because CDOs often included a disproportionately large share of (eventually worthless) mezzanine tranches from subprime securities, CDOs often suffered losses that were much more severe than those on subprime securities.

In hindsight, participation in the subprime market turned out to be disastrous for many subprime borrowers and investors. In the next section, we discuss alternative explanations for why these people wound up in such unfortunate positions.

*The insider/outsider theory*

What might be called the insider/outsider theory of the crisis contends that insiders in the securitization process (mortgage brokers, mortgage bankers and investment bankers) took unfair advantage of the outsiders at the opposite ends of the chain (borrowers and investors). At the start of the chain, borrowers were disadvantaged by unscrupulous mortgage brokers and loan officers, who hid the true costs of subprime loans. One oft-cited example is a broker or loan officer failing to mention that a loan with low initial interest rate would 'reset' to a much higher interest rate after two or three years. The typical subprime loans, called '2/28s' or '3/27s', were fully amortizing, 30-year mortgages that had fixed interest rates for two or three years. The interest rates then began to float by a set margin over a short-term benchmark interest rate. While the initial fixed rates on these hybrid loans were sometimes called 'teaser' rates, the rates were actually quite high relative to rates on comparable prime loans (Foote *et al.*, 2008a). Subprime loans were also likely to carry prepayment penalties that made it costly for borrowers to refinance before the reset occurred.

At the opposite end of the chain, investors were said to be exploited by the investment banks who packaged the loans into securities. These banks allegedly obfuscated the quality of the loans by creating unnecessarily complex instruments, or by simply misrepresenting what the securities included. Stiglitz (2010, p. 77) sums up the insider/outsider theory succinctly: 'The wheelings and dealings of the mortgage industry in the United States will be remembered as the great scam of the early twenty-first century'.

The insider/outsider theory suggests a simple answer to the fundamental question of why so many people made bad investments: the big losers in the subprime crisis were outsiders, deceived by better-informed insiders into thinking they were making reasonable decisions when in fact they were being misled. The structure of the originate-to-distribute model meant that no one in the securitization chain had the incentive to correct the misconceptions of either borrowers or lenders. The result was a dramatic expansion in credit supply that sent housing prices higher. When borrowers were finally hit with unaffordable payments, they defaulted, leading to widespread losses on securities held by investors. The defaults also sparked a flood of foreclosures that sent housing prices lower during the ensuing housing bust.

*The bubble theory*

An alternative theory of the crisis holds that borrowers and lenders were deceived not by devious insiders but instead by a devious idea: house prices could rise rapidly more or less indefinitely or, at worst, fall only gently. In this narrative, the financial system did what it is designed to do: match borrowers to lenders. These borrowers and lenders did not have to be coaxed into their decisions by financial market insiders. On the contrary, both subprime borrowers and lenders were eager to cash in on the biggest real-estate boom in American history.

A fundamental component of this alternative theory is that a bubble developed in the American housing market in the early 2000s. In economic parlance, a bubble occurs whenever the market price of a long-lived asset rises for reasons that have nothing to do with its fundamental value. Because people expect the price to rise tomorrow, there is plenty of demand for the asset today, and sure enough, the price goes up. This type of speculative cycle does not continue indefinitely and, at some point, both prices and expectations of further appreciation collapse.

Examples of past episodes generally thought to be speculative bubbles include the tulip craze of 1630s Holland, the Florida land craze of the 1920s, and the Internet stock boom of the late 1990s. In none of these episodes did the securitization of credit play a role, and the bubble theory contends that securitization did not spark the subprime crisis either. The only fuel needed for the bubble was the self-fulfilling expectation that house price appreciation (HPA) would continue to be positive. Buyers eager for capital gains stretched themselves to buy the largest homes they could. Subprime borrowers were not troubled if their monthly payments were large relative to their incomes, because their expected capital gains would offset (and perhaps even surpass) the monthly payments their mortgages required. If the borrowers ever got into financial trouble, they could sell their houses for more than enough money to pay off their debts. At the other end of the chain, investors were eager to lend to subprime homeowners, because these mortgages paid high interest rates, and because rapid HPA made the characteristics of subprime borrowers irrelevant to the success of their investments. Even if borrowers made no down payments and thus had no equity at origination, strongly positive HPA meant that borrowers soon would have positive equity, and thus could get themselves out of any trouble. In short, the decisions of both subprime borrowers and investors can be explained without resorting to insider/outsider conflict, as long as both borrowers and investors believed that house prices would keep going up.

Of course, the sensitivity of subprime outcomes to HPA caused problems in 2006, when house prices slammed into reverse. Falling prices increase the probability that a borrower's equity becomes negative. Some of these 'underwater' owners will default because they no longer believe that keeping their mortgage current makes financial sense (so-called 'ruthless' or 'strategic' defaults). Other underwater owners will suffer adverse life events, including job loss, divorce or illness. These owners may want to stay in their homes, but they will be unable to keep their mortgages current (because of the adverse life events) and at the same time they are unable to sell their houses to fully discharge their debts (because of negative equity). The combination of adverse life events and negative equity therefore results in a 'double-trigger' default (see Foote *et al.* (2008b) for a simple two-period model outlining the default decision of an underwater

borrower). Because subprime borrowers generally began their homeownerships with low initial equity and high payments relative to incomes, and because their damaged credit histories suggest binding liquidity constraints, it is not surprising that subprime homeownerships turned out to be highly sensitive to falling house prices. Proponents of the bubble explanation believe that 'exploding' subprime mortgages are not required to explain the surge of subprime defaults after 2005. A combination of falling house prices and vulnerable subprime homeowners will suffice.

## Testing theories of the subprime crisis

*Evaluating the insider/outsider theory*

Many economists are now investigating the role that securitization played in the recent housing cycle. One empirical example is Mian and Sufi (2009), which studies patterns in ZIP-code-level data on income, mortgage originations, credit scores and defaults. The authors find that ZIP codes housing a disproportionate share of borrowers with low credit scores as of 1996 enjoyed a larger relative expansion of securitized mortgage credit from 2002 to 2005. Predominantly subprime ZIP codes also experienced higher default rates after 2005, which are correlated most strongly with growth in mortgage credit raised through the securitization channel. 'This result hints at moral hazard on behalf of originators as a factor contributing to the expansion in credit supply, although we believe more research is needed on this precise mechanism' (Mian and Sufi, 2009, p. 1454).

Other research uses loan-level data to test the insider/outsider theory. The marketers of subprime securities collected a great deal of loan-level data on the mortgages they handled, and they made this data available to potential investors and researchers. In the remainder of this subsection, we discuss some research that uses these data, along with some theoretical considerations that are relevant to the insider/outsider narrative.

**Were interest-rate resets responsible for the subprime crisis?** The view that interest-rate resets sparked the wave of subprime defaults informed the government's anti-foreclosure policy early in the crisis. But this view finds little support in loan-level data: most subprime borrowers who defaulted did so without an increase in their payments. Table 1 presents statistics from a large sample of troubled subprime loans from the crisis, defined as loans on whom lenders initiated foreclosure proceedings between 2007 and 2010. The top part of the table shows that most delinquency spells for troubled borrowers started when their payments equalled, or fell short of, their initial payment. The share fluctuates, but at no point in the crisis could more than 14% of troubled borrowers tie their problems to a payment higher than the payment due when they originated the loan. The lower panel of the table illustrates why. Most troubled loans were either fixed-rate mortgages or adjustable-rate mortgages prior to the reset of the loan. Even among loans that did reset, a significant share saw their payments remain the same. Because post-reset interest rates were set to be a fixed margin above a benchmark short-term interest rate (such as the six-month LIBOR), after 2008 interest-rate resets generally caused payments to fall.

In an expansive study, Mayer *et al.* (2009) use loan-level data to investigate the relationship between subprime defaults and the novel features of subprime mortgages,

**Table 1** *Did payment increases cause the crisis? Sample includes subprime first-lien mortgages in McDash/LPS dataset on which lenders initiated foreclosure proceedings from 2007 to 2010. All loans originated after 2005. Source: McDash Analytics, a subsidiary of Lender Processing Services (LPS), LLC.*

| | | 2007 | 2008 | 2009 | 2010 | Total |
|---|---|---|---|---|---|---|
| Prior to delinquency spell that lead to foreclosure... | | | | | | |
| % of loans with... | Payment increase | 5% | 14% | 14% | 9% | 11% |
| | Payment reduction | 0% | 0% | 4% | 9% | 2% |
| | No change since orig | 95% | 86% | 81% | 82% | 87% |
| FRM share | | 29% | 36% | 42% | 58% | 38% |
| ARMs 30 days delinquent prior to reset | | 65% | 47% | 21% | 6% | 41% |
| ARMs that reset but payment stayed the same or fell | | 1% | 3% | 23% | 27% | 10% |
| # obs in dataset (10% national sample) | | 12,318 | 13,457 | 9,480 | 5,475 | 40,730 |
| # obs in thousands | | 123 | 135 | 95 | 55 | 407 |

including prepayment penalties and resets. Consistent with other research, they find that these features fail to explain the sharp increase in subprime defaults (Mayer *et al.*, 2009, p. 48):

> Our conclusions run counter to the popular perception that unorthodox mortgage features are responsible for the surge in defaults. At first glance, the fact that the most common subprime mortgage was a confusing and complicated product – a short-term hybrid with a prepayment penalty – and that delinquency rates were highest on these products suggest that the mortgage type itself must be to blame. We suggest instead that default rates were highest on these products because they were originated to the borrowers with the lowest credit scores and highest loan-to-value ratios.

**Were prime borrowers steered into subprime mortgages?** Loan-level datasets also revealed that as the housing boom progressed, the average credit scores of subprime borrowers were stable or even rose (Foote *et al.*, 2008a; Mayer *et al.*, 2009). Some have interpreted this as evidence that potentially prime borrowers were steered into subprime mortgages by unscrupulous brokers, who were often paid more for delivering high-interest loans to originators (Brooks and Simon, 2007). But a deeper look at the data reveals a more complex picture. A low credit score is not the only reason that a loan could be labelled subprime. Prime borrowers also needed to have non-trivial down payments and reasonable payment-to-income ratios. They also needed to fully document their incomes. Foote *et al.* (2008a) and Gerardi *et al.* (2009) show that these additional risk characteristics generally worsened over time as credit scores improved. One potential explanation is that mortgage brokers misled borrowers into taking loans with worse risk characteristics, even when less-risky loans would have better served borrowers. An alternative explanation, consistent with the bubble theory, is that borrowers wanted to stretch themselves into the most expensive houses they could, because they expected house prices to keep rising.

**Did worsening credit standards cause the subprime crisis?** The worsening of these subprime risk characteristics led naturally to the hypothesis that declining credit standards explain the poor performance of subprime loans originated at the height of the

housing boom. Gerardi *et al.* (2009) investigate this issue with an econometric model that links defaults of individual subprime loans to their risk characteristics. The model implies that worsening risk characteristics explain relatively little of the bad performance of loans made in 2006, when housing prices peaked. Specifically, the problem with the 2006 loans was not that this vintage included more loans with adjustable rates, zero down payments and reduced documentation (though it did). The problem with the 2006 vintage was that loans with adjustable rates, zero down payments and reduced documentation performed drastically worse in 2006 than otherwise identical loans earlier. What was the factor that caused so many defaults among the later loans? One possibility is that the loans were worse on dimensions that are unobservable to the econometrician. Another possibility is that the earlier loans were made when house prices were rising, while the 2006 loans were made just as prices were beginning to fall.

**Did investors in subprime securities know what they were buying?** The question of what investors knew about subprime securities is addressed as much by the *existence* of large, loan-level datasets on subprime loans as by any specific fact that these datasets reveal. As noted earlier, these datasets were constructed from loan-level files supplied by the issuers of subprime securities. The data included in these files were extensive, including not only the loan and borrower information mentioned earlier in this section but also geographic information such as ZIP code. Similarly, issuers of CDOs that included tranches of subprime securities provided a list of these securities, so that CDO investors would know about all the hundreds of thousands of loans that the CDO contained. To be sure, gathering and evaluating this information took resources. But investors in structured securities were institutions and not individuals – widows and orphans did not buy CDOs of subprime securities – and the costs of such analysis were small relative to the sums of money involved. In fact, investors and researchers devoted vast resources to careful and sophisticated analysis of the data, much of which was widely disseminated through analyst reports and even, in some cases, books about mortgages and mortgage-backed securities. To think of an obvious contrast, banks view detailed loan-level information for their portfolio loans as proprietary information not to be shared with anyone.

A related question is whether investors knew how to use this data to predict mortgage performance. Housing economists generally model mortgage performance as a function of house prices (which determine the borrower's equity position), interest rates (which are relevant even for fixed-rate loans, via the prepayment channel), and loan- and borrower-level risk characteristics (including loan-to-value ratios, documentation status and credit scores). Perhaps surprisingly, economists using these models in academia, government and private firms are actually quite good at forecasting the performance of residential mortgages *if* they know how house prices and interest rates will evolve. As Gerardi *et al.* (2009) show, the relationship between house prices and mortgage defaults was well understood before the crisis and mortgage analysts, despite their close ties to the business, drew attention to the risks. One conference-call presentation from UBS in September 2005 was titled 'Subprime Home Equities: It's (Almost) All About Home Prices'. Another study, from Lehman Brothers Fixed Income Research (2005), predicted that collateral losses on subprime securities would top 17% if home prices fell by 5% for three consecutive years. It stressed that losses of this size would have rendered even the mezzanine tranches of many subprime securities worthless.

**Is the insider/outsider theory consistent with information economics?** The insider/outsider narrative can also be evaluated theoretically using the classic lemons model of Akerlof (1970). Akerlof argued that markets can break down whenever sellers have more information about the goods being traded than buyers do. His insight was that rational sellers exploit their private information but that in general equilibrium, buyers anticipate seller behavior and either offer less for the good or, in extreme cases, simply refuse to trade. Crucially, when they do trade, they receive exactly what they expect. Applied to subprime lending, investors in subprime securities are the buyers while those inside the securitization chain are the sellers. If subprime securities did in fact suffer from large informational asymmetries, Akerlof's theory might have explained a failure of this market to get off the ground. But Akerlof's model cannot explain the explosive growth of subprime lending during the housing boom, nor can it explain the unanticipated losses in the housing bust. In short, Akerlof's model does not imply that informed agents will gain at the expense of uninformed agents, so it is not consistent with the basic insider/outsider explanation of the crisis.

To illustrate the disconnect between analysis of the crisis and information economics, it is instructive to look at a paper, Keys et al. (2009), which is widely cited as evidence that asymmetric information played a major role. Keys et al. claim to show that lenders screened loans more intensively when they planned to hold them in their own portfolio as compared to when they planned to sell them to others. Even if we take their empirical claim, disputed by Bubb and Kaufman (2009), as correct, standard information economics says it tells us nothing about why investors lost money. According to Akerlof, investors will anticipate that lenders screen loans originated for sale less intensively and offer correspondingly lower prices. The evidence from Gerardi *et al.* (2009), discussed above, shows that analysts using the data provided by issuers did not systematically over-predict the performance of mortgages but rather based their optimistic forecasts of deal performance on the assumption of ever-rising house prices.

One possibility is that investors didn't understand the institutional details of the mortgage origination process and thus did not realize that there was an asymmetric information problem, but it is difficult to square this with the facts. The intermediaries were, as we discuss below, major investors in the securities and suffered massive losses, so it is hard to argue that they did not understand the origination process. Further, insiders were candid about their role in the process. A top executive at one of the largest originators of subprime loans, Countrywide Financial, quoted on the front page of the Money and Investing section of the *Wall Street Journal*, announced that, "We're looking to hold only pristine product on the balance sheet." (see Simon and Hagerty, 2005).

If asymmetric information was a potential problem, then how could the market grow so fast? One possibility is that insiders were aware of the market-killing potential of asymmetric information, so they tried to level the playing field by sharing as much information with outsiders as they could. Insider concerns about information would explain the existence of the extensive, micro-level datasets on subprime loans. Indeed, the detailed loan-level data used by Keys et al. to uncover the alleged asymmetric information problem was provided in real time by the issuers of securities and used by virtually every major investor in the mortgage securities market to measure and predict the performance of loans.

**Did insiders really profit at the expense of outsiders?** Additional evidence against the insider/outsider story comes by noting that insiders also suffered severe and sometimes fatal losses during the subprime crisis. Of the top ten issuers of private-label mortgage securities in 2006, three failed (Washington Mutual, Indymac and Lehman Brothers), two were forced to merge to avoid bankruptcy (Bear Stearns and Countrywide) and the remaining five had, as of November 2008, reported a total of $120 billion in losses, largely because of failed investments in mortgage-backed securities and CDOs composed of mortgage-backed securities (for losses, see Onaran and Pierson (2008)). Managers of these firms had large investments in the firms and, in some cases, investment vehicles created by the firms to buy mortgage-related securities. As a result, these managers suffered huge personal losses when their firms failed. Lehman staff owned a quarter of the company's stock and typically could not sell for five years after the stock was granted (Sorkin, 2010, p. 294). And top Bear Stearns executives were major investors in the two mortgage-related hedge funds whose massive losses played such a pivotal role in the crisis (Muolo and Padilla, 2010, p. 244).

*Evaluating the bubble theory*

A convincing empirical case for the bubble theory would have two components. First, it would show that investors and borrowers held highly optimistic expectations for house prices during the housing boom. Second, it would show that these expectations were unjustified given the fundamental drivers of housing prices, specifically future rents and interest rates. The fall in housing prices after 2006 could then be attributed to a correction of prices toward fundamentals, rather than an exogenous decline in mortgage credit, or the defaults and foreclosures suffered by subprime borrowers.

It is not difficult to find evidence supporting the first part of this story – widespread market optimism over HPA. To be sure, interested parties like the National Association of Realtors pushed a narrative of ever-rising house prices, but so did well-respected mortgage analysts. Gerardi *et al.* (2009) shows that analyst reports placed little probability on a steep decline in house prices, because nominal prices had not declined for decades, and because regional house-price declines had only been experienced in steep recessions. The possibility that house prices would fall by 30% from mid-2006 to late 2009 (the actual price decline, according to the national Case–Shiller repeat-sales index) would have struck most of these analysts as exceedingly remote. The worst-case scenarios of mortgage analysts generally featured consecutive yearly house-price declines of 5% or less. For example, the meltdown scenario for housing prices assumed by Lehman Brothers Fixed Income Research (2005) was three consecutive years of 5% declines in house prices, followed by positive annual HPA of 5% thereafter. Lehman analysts gave the meltdown scenario a 5% chance of occurring. The worst-case scenario in a September 2005 conference call presentation by Thomas Zimmerman, the executive director of the US Securitized Products Strategy Group at UBS, was for prices to fall by 2–3% for three consecutive years (Zimmerman, 2005). And more than a year later, in early 2007, Fannie Mae performed an internal stress test on its subprime portfolio that assumed two consecutive price declines of 5%, along with a two-percentage-point increase in interest rates (Hilzenrath, 2008).

It is much harder to prove the second part of the bubble case – that high house prices were not justified by fundamentals and thus were destined to fall. The main problem here is that the fundamental value of any asset (financial or otherwise) depends both on future income streams and the discount factors needed to turn these streams into present values. Because expectations of future rents and interest rates are essentially unobservable, they are hard to compare to current and expected future house prices. In fact, during the housing boom, the inability to know for sure whether house prices were close to fundamentals prevented professional economists from achieving consensus on the bubble question (Gerardi *et al.*, 2010). Many economists, including Baker (2002), were pessimistic about the future trajectory of housing prices, in the light of the massive increase in prices, relative to both rents and household incomes, that had already occurred. But other economists found that careful estimates of housing user costs had remained within their historical ranges for most US cities, thanks in part to declining interest rates (Himmelberg *et al.*, 2005).

Gerardi et al. (2010) contend that most economists were agnostic about the bubble question, which is not surprising given their training. The "Fundamental Theorem of Asset Pricing," the heart of modern financial economics, implies that asset prices are, to a first approximation, unpredictable. If it is likely that asset prices will fall tomorrow, then rational actors will rush to sell the asset. This rush to sell causes the price to fall today, not tomorrow. While economists have found some evidence of predictability in returns, Lo and MacKinlay (2001) argue that this predictability operates on a relatively small scale, so research has not "uncovered tremendous untapped profit opportunities" (2001, p. xxii).

An optimistic investor could always fend off a pessimistic economist by citing one of the massive forecasting errors in economic history, ranging from Irving Fisher's claim that stocks had reached a 'permanently high plateau' in 1929, to Robert Shiller's assertion of 'irrational exuberance' in stocks in 1996.

A related theoretical problem with the bubble story is that although bubbles have been intensively studied, consensus has yet to emerge on how bubbles form or why they pop (see Brunnermeier, 2008, for a survey). Bubbles appear frequently in laboratory settings, but economists are not sure why, as bubbles often arise when experimental subjects are given a great deal of information about the asset's fundamentals (see Smith, Suchanek, and Williams, 1988). Some analysts have suggested a link between the insider/outsider and bubble explanations by arguing that an out-of-control securitization process caused a housing bubble. For example, while much of the housing discussion in Stiglitz (2010) is a strident criticism of mortgage-industry insiders, the book also cites a bursting bubble as the main cause of the foreclosure wave (Stiglitz, 2010, pp. 1–2):

> The basic outlines of the story are well known and often told. The United States had a housing bubble. When that bubble broke and housing prices fell from their stratospheric levels, more and more homeowners found themselves 'underwater'. They owed more on their mortgages than what their homes were valued. As they lost their homes, many also lost their life savings and their dreams for a future – a college education for their children, a retirement in comfort. Americans had, in a sense, been living in a dream.

A challenge for these economists is to show how a bubble could have been started in this way. Nothing in Brunnermeier (2008), for example, explains how an insider/outsider conflict, or a relaxation of credit standards, leads to unjustified estimates of future asset-price appreciation. Models like the ones in Chapter 9 of Allen and Gale (2007) and Favilukis, Ludvigson and Van Nieuwerburgh (2009) show how policy changes in financial markets can generate increases in asset prices but the models require a policy reversal to generate a collapse. Mian and Sufi (2009) cite the classic credit cycles model of Kiyotaki and Moore (1997) to note that higher collateral values may increase the availability of credit to previously constrained borrowers. The additional borrowing, in turn, can push up collateral values further by increasing the effective demand for the collateral asset. But Kiyotaki and Moore (1997) is a perfect-foresight model. After an initial (unforeseen) shock that changes the value of the collateral asset, there are no more price surprises. The same is true of the textbook $q$-theory of investment, once the price of installed capital jumps in response to an unforeseen innovation in productivity. These predictions are impossible to square with the 30% decline of US house prices that had such devastating and unforeseen effects on the financial system.

Given the incomplete state of the theoretical literature on what causes bubbles, the bubble explanation is at best a proximate explanation of the subprime crisis. But while the bubble explanation is incomplete, it is also parsimonious. The only deviation from rationality it requires is the belief that house prices will continue to rise rapidly for the foreseeable future. If this belief was widespread, then millions of decisions by borrowers and investors make sense. By contrast, the insider/outsider theory requires that borrowers and sophisticated institutional investors independently, repeatedly, and systematically misunderstood the choices before them, even though they had been dealing with similar choices for decades without any problems.

## Policy implications

The question of whether the subprime crisis resulted from insider/outsider conflicts or a classic asset bubble has important implications for public policy. If insider/outsider conflicts are responsible, then regulators should try to align the incentives of borrowers and lenders and/or protect borrowers from unfair practices. Along these lines, the Dodd–Frank Wall Street Reform and Consumer Protection Act, passed by Congress in 2010 as a wide-ranging response to the financial crisis, established the Consumer Financial Protection Bureau to promulgate regulations for mortgages, credit cards and other financial products. Elizabeth Warren, the special adviser to the Treasury Secretary who was charged with setting up this bureau, drew a clear line between poor consumer regulation and the financial crisis in the months before the bureau began operations: "The crash of 2008 made it clear that the consumer agency should have been law years ago. It should have been in place before an out-of-control lending industry developed and marketed mortgages that many knew would explode" (Warren, 2010). Alternatively, if the subprime debacle was simply the latest manifestation of a classic asset bubble, then consumer regulation is less likely to prevent a future crisis, because all the actors involved will oppose restrictions on their activities. After all, it is hard to stop consenting adults. The alternative policy prescription is to make financial

institutions as robust as possible to unforeseen declines in asset prices, given our incomplete knowledge of what determines these prices. One way to do this is with enhanced capital regulations for financial institutions. Additionally, the bubble explanation would imply that because future bubbles are both possible and unpredictable, policymakers should possess the necessary tools to deal with the crises that these bubbles may spawn. One example of such a tool is the ability to wind down insolvent financial institutions in an orderly way, another component of the Dodd–Frank Act.

**Acknowledgements**

The views herein are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Boston or the Federal Reserve System.

CHRISTOPHER L. FOOTE AND PAUL S. WILLEN

*See Also* **Fannie Mae, Freddie Mac and the crisis in US mortgage finance; foreclosures.**

### *Bibliography*

Adrian, T. and Shin, H. S. 2010. The changing nature of financial intermediation and the financial crisis of 2007–2009. *Annual Review of Economics*, 2, 603–618.

Akerlof, G. A. 1970. The market for 'lemons:' quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488–500.

Allen, F. and Gale, D. 2007. *Understanding financial crises.*Oxford University Press, New York.

Baker, D. 2002. The run-up in home prices: a bubble. *Challenge*, 45(6), 93–119.

Barr, A. 2007. Subprime mortgage derviatives index plunges. *MarketWatch*, 23 February. Available at http://www.marketwatch.com/story/index-of-subprime-mortgage-derivatives-plunges-on-sector-woes.

Brooks, R. and Simon, R. 2007. Subprime debacle traps even very credit worthy. *Wall Street Journal*, 3 December.

Brunnermeier, M. K. 2008. Bubbles. In *The New Palgrave Dictionary of Economics* 2nd edn (eds. S. N. Durlauf and L. E. Blume). Palgrave Macmillan, Basingstoke. Available at http://www.dictionaryofeconomics.com/article?id=pde2008_S000278.

Bubb, R. and Kaufman, A. 2009. Securitization and moral hazard: evidence from a lender cutoff rule. *Federal Reserve Bank of Boston Public Policy Discussion Paper No. 09-5*. Available at http://www.bos.frb.org/economic/ppdp/2009/ppdp0905.pdf. Accessed 6 June 2011.

Coval, J., Jurek, J. and Stafford, E. 2009. The economics of structured finance. *Journal of Economic Perspectives*, 23(1), 3–25.

Favilukis, J., Ludvigson, S.C. and Van Nieuwerburgh, S. 2009. The macroeconomic effects of housing wealth, housing finance, and limited risk-sharing in general equilibrium Working paper, New York University.

Foote, C. L., Gerardi, K., Goette, L. and Willen, P. S. 2008a. Just the facts: an initial analysis of subprime's role in the housing crisis. *Journal of Housing Economics*, 17(4), 291–305.

Foote, C. L., Gerardi, K. and Willen, P. S. 2008b. Negative equity and foreclosure: theory and evidence. *Journal of Urban Economics*, 64(2), 234–245.

Gerardi, K., Willen, P., Sherlund, S. M. and Lehnert, A. 2009. Making sense of the subprime crisis. *Brookings Papers on Economic Activity*, 2008(2), 69–159.

Gerardi, K. S., Foote, C. L. and Willen, P. S. 2010. Reasonable people did disagree: optimism and pessimism about the U.S. housing market before the crash. *Federal Reserve Bank of Boston Public Policy Discussion Paper No. 10-5*. Available at http://www.bos.frb.org/economic/ppdp/2010/ppdp1005.htm. Accessed 6 June 2011.

Greenlaw, D., Hatzius, J., Kashyap, A. K. and Shin, H. S. 2008. Leveraged losses: lessons from the mortgage market meltdown. *Proceedings of the U.S. Monetary Policy Forum*. Available at http://research.chicagobooth.edu/igm/events/docs/USMPF-final.pdf. Accessed 6 June 2011.

Hilzenrath, D. S. 2008. Fannie's perilous pursuit of subprime loans. *Washington Post*, 19August. Available at http://www.washingtonpost.com/wp-dyn/content/article/2008/08/18/AR2008081802111_pf.html. Accessed 6 June 2011.

*Himmelberg, C., Mayer, C. and Sinai, T. 2005. Assessing high house prices: bubbles, fundamentals, and misperceptions. *Journal of Economic Perspectives*, 19(4), 67–92.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and VikrantVig. 2010. "Did securitization lead to lax screening? Evidence from subprime loans." *Quarterly Journal of Economics* 125(1): 307–362.

Kiyotaki, N. and Moore, J. 1997. Credit cycles. *Journal of Political Economy*, 105(2), 211–248.

Lehman Brothers Fixed Income Research. 2005. *U.S. ABS Weekly Outlook: HEL Bond Profile Across HPA Scenarios*. 15 August.

Lo, A.W. and MacKinlay A. C. 2001. *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, N.J. Available at http://press.princeton.edu/books/lo/.

Mayer, C., Pence, K. and Sherlund, S. M. 2009. The rise in mortgage defaults. *Journal of Economic Perspectives*, 23(1), 27–50.

Mian, A. and Sufi, A. 2009. The consequences of mortgage credit expansion: evidence from the U.S. mortgage default crisis. *Quarterly Journal of Economics*, 124(4), 1449–1496.

Muolo, P. and Padilla, M. 2010. *Chain of Blame: How Wall Street Caused the Mortgage and Credit Crisis*. Wiley, New York.

Onaran, Y. and Pierson, D. 2008. Banks' subprime-related losses surge to $591 billion: table. Bloomberg, 29 September. Available at http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aSlW.imTKzY8. Accessed 6 June 2011.

Simon, R. and Hagerty, J. R. 2005. How American lenders shelter themselves. *Wall Street Journal*, 22 September.

Smith, V. L., Suchanek, G. L. and Williams, A. W. 1988. Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica* 56(5): 1119–1151.

Sorkin, A. R. 2010. *Too Big to Fail: The Inside Story of How Wall Street and Washington Fought to Save the Financial System – and Themselves*. Penguin, London.

Stiglitz, J. E. 2010. *Freefall: America, Free Markets, and the Sinking of the World Economy*. Norton, New York.

Warren, E. 2010. *Priorities for the New Consumer Financial Protection Bureau*. Remarks at the Consumer Federation of America Financial Services Conference, 2 December. Available at http://www.treasury.gov/press-center/press-releases/Pages/tg987.aspx. Accessed 6 June 2011.

Zimmerman, T. 2005. Subprime home equities: it's (almost) all about house prices. *UBS Conference Call Presentation*, 26 September

# tulipmania

The Netherlands of 1634–7 was the scene of a curious speculation in tulip bulbs that has come to be known as the Dutch tulipmania. Single bulbs of rare and prized varieties such as Semper Augustus or Viceroy became worth a middle-sized fortune. In its most extreme final phase in January-February 1637, prices of even common varieties such as Switsers or Witte Kroone soared twentyfold within a month and then crashed back to their original values. That these were prices of easily reproducible horticultural products has added to the bemusement of generations of historians and economists.

In the succeeding 370 years, the *historical* tulipmania became, in itself, an obscure footnote to the *conceptual* tulipmania of economics and finance, a word warning of the obvious, delusional speculative excess that human behaviour in financial markets can create (see, for example, Kindleberger, 1996). It is interchangeable with words like 'bubble' or 'mania', which also arose from historically distant events such as the Mississippi or South Sea Bubbles or the more recent 'irrational exuberance'. These words have been used by economic theorists to emphasize an historical basis for the salience of unstable multiple equilibria in forward-looking financial and macroeconomic theories. They have also been used to justify ignoring financial market outcomes that contradict favoured asset pricing theories by means of an arbitrary invocation of the existence of a bubble.

## The traditional image of tulipmania

Modern references to the tulipmania usually depend on the brief description in Charles Mackay's *Extraordinary Popular Delusions and the Madness of Crowds* (1852). The tulip originated in Turkey and spread into western Europe in the mid-16th century. The tulip was immediately accepted by the wealthy as a beautiful and rare flower, appropriate for the most stylish gardens. The market was for durable bulbs, not flowers. The Dutch dominated the market for tulips, initiating the development of methods to create new flower varieties. The bulbs that commanded high prices produced unique, beautifully patterned flowers; common tulips were sold at much lower prices.

Beginning in 1634, non-professionals entered the tulip trade in large numbers. According to Mackay, individual bulb prices reached astronomic levels. For example, a single Semper Augustus bulb was sold at the height of the speculation for 5,500 guilders, a weight of gold equal to $66,000 evaluated at $600/oz. Mackay provided neither the sources of these bulb prices nor the dates on which they were observed, however.

Finally, and unexplained by Mackay, the frenzy suddenly terminated. According to Mackay, even rare bulbs could find no buyers at ten per cent of their previous prices, creating long-term economic distress. Mackay presented no evidence of immediate post-collapse transaction prices of the rare bulbs. Instead, he cited prices from bulb sales of 60 years, 130 years, or 200 years later as indicators of the magnitude of the collapse and of the obvious misalignment of prices at the peak of the speculation.

Moreover, Mackay provided no evidence of the general economic context from which the speculation emerged.

## The fundamentals of the tulipmania

Unfortunately, the fundamentals of markets in rare bulbs present a much more prosaic picture. The bulk of the speculation concerned highly prized tulips that were infected with mosaic virus. Mosaic virus had the effect of producing unique feathery patterns in the flower that could be reproduced only through propagation by budding, not by seeds. Hence, the rate of reproduction was much more limited that one might expect. Such bulbs were traded primarily among professionals. Their prices were supported by a strong demand by flower fanciers, not only in the rapidly growing Netherlands of the golden age but also by the wealthy nobility and merchants of surrounding countries. During the period of the tulipmania, 1634–7, the already high prices of such bulbs doubled or tripled. Over the course of decades or centuries, prices for these varieties converged to the low cost of reproduction, and this has been taken as evidence of folly.

However, an examination of the pricing of prized flower varieties throughout history reveals a similar pattern: prices of the prototype are very high, perhaps even representing a medium fortune. Then, as they are reproduced through succeeding generations, they become common. Just as for the value of a prized racehorse put out to stud, the high initial price represents the present discounted value of the valuation above cost of successive, expanding generations, wherein the value of any individual exemplar is bound to fall.

The more frenzied phase of the tulipmania described by Mackay took place from mid-1636 to February 1637, but especially in January, 1637. At this time trading, especially among the non-professionals, took place in newly organized 'colleges', which were located in taverns. The trading was not for actual bulbs but for contracts for forward delivery. Since bulbs had to remain in the ground through the winter, none were actually delivered on these contracts before the speculation ended. Contracts were not marked to market, and margin was not posted. A small, fixed amount of 'wine money' had to be delivered by the buyer, which provides the flavour of, if not the fuel for, what was happening during the frenzied trading in these taverns.

When this part of the speculation collapsed in February 1637, some city governments proposed winding up outstanding contracts with a ten per cent payment on contracted amounts if the buyer refused to accept delivery. Perhaps this is where Mackay got the notion that bulbs could not be sold at ten per cent of their previous value, even though a buyer might refuse the deal if prices had fallen only to 90 per cent of the contracted amount. There were very few takers even on this offer, but short sales were in any case unenforceable contracts under Dutch law.

When one looks at notarized contracts for actual bulbs, however, the picture is quite different. Some rare bulbs that were auctioned for high prices at the very peak of the speculation in February 1637 still sold for high, albeit much lower, prices in 1642. For example, an Admirael Liefkens bulb was sold for 1,345 guilders at the peak and for 220 guilders in 1642, an annual percentage decline in value of 36 per cent. This rate of

decline is comparable to the typical pattern of price behaviour for valued varieties in successive historical periods and does not indicate anything unusual in the mania.

It is the rare bulb price behaviour during the tulipmania that has been emphasized historically. But at the very end common bulbs sold in bulk shot up twentyfold and soon collapsed back to one-twentieth of the peak. It is this usually ignored bit of the episode that remains a puzzle.

## An historical background

The tulip market was introduced into the Netherlands during the Eighty Years' War of independence between the Dutch and the Spanish, and the tulipmania occurred in the middle of the Thirty Years' War as the two conflicts merged. The Spanish were thwarted in their attempts to subjugate the Netherlands, which consolidated its territory and eventually seized control of most of international shipping. The Thirty Years' War of 1618–48 was particularly destructive of the populations and economies of central Europe, with many principalities in the Holy Roman Empire losing one-third of their populations.

In every year of the war, the Dutch fielded large armies and supported large fleets, though the population of the Netherlands was no more than 1.5 million. The Dutch provided much of the strategic planning and finance for the Protestant effort, along with France, negotiating and financing the successive interventions of Denmark and Sweden on the Protestant side in the 1620s and 1630s.

From 1620 to 1645, the Dutch established near-monopolies on European trade with the East Indies and Japan, conquered most of Brazil, took possession of the Dutch Caribbean islands, and founded New York. In 1635 the Dutch formed a military alliance with Richelieu's France, which eventually placed the Spanish Netherlands in a precarious position. In 1639 the Dutch completely destroyed a second Spanish Armada of a size comparable to that of 1588. As a result of the war, Spain ceased to be the dominant power in Europe, and the Netherlands, though small in population and resources, temporarily became a major power centre because of its complete control over international trade and international finance. The Dutch were to 17th-century trade and finance as the British were to 19th-century trade and finance.

Sophisticated finance mechanisms evolved with the establishment of its trade and finance dominance. Amsterdam became the leading market for short- and long-term credit; and markets in stocks, commodity futures, and options materialized early in the 17th century. Trading of national loans of many countries centred on Amsterdam, as did a market in the shares of joint stock companies. The East India Company, founded in 1602, gradually gained control over east Asian trade and consistently paid out large dividends. Interest rates on Dutch markets were remarkably low for the times; for example, the East India Company paid no more than five per cent on advances during the 17th century.

There were some dark periods during this golden age, and it should be carefully noted that these occurred during the years of the tulipmania. From 1635 to 1637, bubonic plague ravaged the Netherlands. In July 1634 the Holy Roman Empire completely defeated Swedish forces in the Battle of Nordlingen, forcing a treaty on the

German Protestant principalities in the May 1635 Peace of Prague and releasing Spanish resources for the war against the Dutch. Along with the growing war weariness in the Netherlands, these events forced France to enter the Thirty Years' War militarily with the Dutch alliance in 1635. Initially unprepared, the French suffered major setbacks, culminating in an imperial invasion of northern France in August 1636.

## How should we interpret the tulipmania?

The tulipmania is an obscure event from distant history that provides a cornucopia of concepts even now. One can take one's pick from the following views, depending on personal taste:

- It was an outburst of speculative fever that serves to the present day as a warning of the dangers of market speculation.
- It was a curious event limited to the Dutch winter of 1636–7 in the middle of an outbreak of bubonic plague and at the time of the greatest success of the Catholic armies of the Empire in the Thirty Years War against the Protestants.
- It was a drinking game in which people without wealth made the equivalent of million euro bets with each other, with no intention or possibility of paying.
- It was a swing of fashion in the most wealthy society of the era, which caused the most exquisite of tulips to have a higher price than Rembrandt's *Night Watch*.
- It was a reasonable and well-calculated investment that still causes the most wonderful outburst of colour every Dutch springtime.

PETER GARBER

*See Also* **speculative bubbles.**

### *Bibliography*

Cooper, P. 1970. *New Cambridge Modern History, Volume 4: The Decline of Spain and the Thirty Years' War*. Cambridge: Cambridge University Press, 1970.

Garber, P. 2000. *First Bubbles: The Fundamentals of Early Manias*. Cambridge, MA: MIT Press.

Kindleberger, C. 1996. *Manias, Panics and Crashes: A History of Financial Crises*, 3rd edn. New York: Wiley.

Mackay, C. 1852. *Extraordinary Popular Delusions and the Madness of Crowds*, vol. 2, 2nd edn. London: Office of the National Illustrated Library.

Posthumus, N. 1929. The tulip mania in Holland in the years 1636 and 1637. *Journal of Economic and Business History* 1, 434–66.

Schama, S. 1987. *The Embarrassment of Riches*. New York: Alfred Knopf.

# Index