Central Banks and Short-Term Interest Rates: Bank of England Operations in the Sterling Money Market

by

Norbert Schnadt

London School of Economics and Political Science

A thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy at the University of London.

October 1994

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The policy instrument of central banks everywhere has usually been a short-term nominal interest rate. This means that central banks have adopted operating procedures whose goal has been to produce some desired level of money market interest rates. Although the Bank of England was in many respects the pioneer of these operating procedures, theoretical and empirical attention has focused almost exclusively on the Federal Reserve. This thesis aims to redress this imbalance by examining – in detail – the sterling money market and the operations of the Bank of England.

This task is carried out in two parts. *Part I* reviews central banks' use of the interest instrument more generally, beginning with an historical sketch of the evolution of central bank money market operations. This sketch is complemented by a critical discussion of two important concepts relating to such operations, namely interest rate smoothing and money base control. A simple analytical model is then developed to illustrate the determination of money market interest rates by the central bank.

Part II specifically concerns the money market operations of the Bank of England, and their implications for the behaviour of sterling money market interest rates. First, a model of the term structure of money market interest rates is derived. Its predicted behaviour in reaction to a change in the Bank's official rate is then empirically verified. Next, the yield on eligible bills – the Bank's intervention asset – is examined. It is argued that these assets carry an excess liquidity premium, arising from the Bank's constraints on their issue. Finally, an empirical model of the overnight interest rate – the UK equivalent of the federal funds rate – is developed, and the reasons for its volatility are investigated.

Acknowledgements

I am extremely fortunate, as a student interested in central banks, to have had Charles Goodhart as my supervisor. I thank him for getting me started in, and then guiding me so expertly through, the largely uncharted research terrain represented by the sterling money market. I am also indebted to John Whittaker for stirring my interest in central banks and for encouraging me to study what they actually do, rather than what many economics texts say they do.

Countless individuals working in the London markets – especially those at two venerable money market institutions, namely the Bank of England and Gerrard and National – patiently explained to me how things are done in the so-called "real" world. I thank them. Away from the "real" world, David Webb kindly made resources available to me at the Financial Markets Group. Without them I could not have completed this thesis. The success of the Group is a sure sign that those who are in the "real" world will continue to rely on those outside it for their insights. Thanks are also due to the following people: Dick Brealey and Jenny Ireland for their support of my work on the City Research Project; Richard Pattinson for invaluable data; Richard Payne for his primer on RATS and ARCH; and Dirk Schoenmaker and Spencer Dale for constructive comments on much of the work in this thesis.

Finally, I am hugely grateful to my wife Sarah - for supporting me (and my work), for putting up with me (and my work), but mostly for being my constant reminder that, whilst everything may be economics, economics isn't everything.

Contents

Figures

Tables

Preface

The contents of this thesis represent research done on a number of different – but closely related – fronts over a period of two years. Some of this material has already been published in a number of separate papers. Instead of reproducing this research here in its original form, I have substantially revised it, and have added yet further material, so as to create a more unified and continuous structure. The result is, hopefully, more substantial than the sum of the initial parts, and its organisation as a book reflects my intention to publish it in this form.

The central strand of my initial research was done for the City Research Project, a three year study into the international competitive position of London in financial services. This Project was funded by the Corporation of London and managed by the London Business School, and consisted of a collaborative effort on the part of numerous academics and consultants. My work was supervised by Charles Goodhart who, with characteristic insight, directed my efforts into areas which are of great interest to practitioners and central bankers, but into which few academics have yet ventured.

My contribution to this Project eventually took the form of two Subject Reports on domestic money markets and central bank operations, which were

the result of extensive fieldwork and interviews with money market participants and central bank officials in London, Paris, Frankfurt and New York. The first Subject Report, *The Domestic Money Markets of the UK, France, Germany and the US*, was a detailed comparative study of the domestic money markets and central bank operations in four countries. The second Subject Report, *European Monetary Union and the Sterling Money Market*, examined the implications of European Monetary Union for money market arrangements and central bank operations in Europe and in the UK, and made several recommendations concerning the sterling money market and the operating practices of the Bank of England. Both Reports were published by the London Business School (in conjunction with the Corporation of London) in January 1994, and generated considerable media interest[1]. A shortened version of the main recommendations, and of their motivation, was also submitted to, and published in *Banking World* ("Reforming the Sterling Money Market", April 1994).

The work done for the City Project forms the basis of two chapters – 3 and 4 – at the centre of this thesis. So as not to dilute the main focus, which is intended to concentrate on the UK, a considerable amount of material on other domestic money markets and on the implications of European Monetary Union has been omitted. What is included, instead, are the results of two other strands of research, also on the theme of central banking and money markets, which were initiated while I was completing my work on the City Project.

The first of these strands was partly analytical and partly historical, and comprised a joint effort with Forrest Capie (City University Business School) and Charles Goodhart to produce a monograph on the development of central banking. This research was funded by the Bank of England, and was prepared for a symposium marking the Tercentenary of the Bank. My own contribution to the monograph lay in the preparation of two appendices. One appendix describes the evolution of the open market operations of several of the major

[1]Articles reporting the Subject Reports appeared in the *Evening Standard*, the *Financial Times*, the *Daily Telegraph*, the *Independent*, *International Securities Lending* and *Euromoney*.

central banks. The other, more substantial, appendix provides a concise, standardised outline of the history of 32 central banks. Both appendices were compiled on the basis of information obtained directly from these central banks and provide an accurate, and hopefully authoritative, introduction to the development of central banking in these countries.

The monograph was presented at a symposium, arranged by the Bank of England, in June 1994. This symposium was attended by over 130 central bank Governors, and was possibly the largest such gathering on record. The monograph is forthcoming shortly in a volume to be published by the Cambridge University Press.

A number of points emerged from this historical examination which usefully complemented my research on contemporary money markets and central bank operations. These are incorporated in chapters 1 and 2. First, there was a remarkable continuity in the Bank of England's money market operations, whose basic structure was already in place by 1825. Second, many other central banks operated in more or less the same way: that is, they provided reserves to the banking system, on demand, by standing ready to discount eligible short-term securities. These stylised facts pointed to several – not entirely unknown – shortcomings in the economic literature on central bank operating procedures. For example, the common description of central bank operations as interest-rate "smoothing" was clearly overly simplistic and seemed to ignore the fact that central banks had already begun to use their discount rates as an instrument to pursue their objectives. Similarly, many of the proponents of money base control – and there are always proponents – somehow seemed to overlook the fact that early central banks had essentially tried, and had failed, to set strict quantitative limits on the supply of bank reserves. Indeed, part of the reason why they became central banks in the first place was that the banking system was potentially unstable in the face of limits on the supply of reserves.

The second strand of research was more analytical and focused on the issue of *how* the money market operations of the Bank of England actually affected sterling money market interest rates. Initial efforts took the form of joint research with John Whittaker (initially at Nottingham, and now

at Staffordshire University). The work examined the implications, for the money market yield curve, of the Bank of England's practice of lending over a range of maturities. This research resulted in a joint paper, "Optimal money market behaviour and sterling interest rates", which is forthcoming in *The Manchester School.* Chapter 5 contains a revised version of this paper.

Two further analytical papers form the basis of chapters 6 and 7. These inquire in greater detail into two features of the sterling money market which repeatedly came to my attention during my work on the City Project. The first concerns the liquidity premium on eligible bills (the assets purchased by the Bank in its daily operations), while the second relates to the volatility of the overnight interest rate. As it turns out, these features are closely related and suggest – even more strongly than the investigation for the City Project – that it may indeed be appropriate for the Bank of England to reconsider its current money market operations.

My studies of domestic money markets and central bank operations in these markets over the past two years have convinced me that these topics demand further research. The determination of short-term interest rates by central banks – rather than the determination of some monetary aggregate – was one of the earliest policy actions in a monetary economy. It has remained one of the most fundamental policy actions and affects countless economic decisions in an immediate and direct fashion. Yet research on this issue remains small compared with the volumes written about the central bank's influence on the money supply, and the role of money in economic decisions.

Similarly, money markets are less well understood than other financial markets. This is surprising since money markets are closely related to the importance which banks have assumed as borrowers, lenders and providers of payment services. Indeed, academics are almost always surprised to learn that the volume of trading in domestic money markets dwarfs that of (say) domestic stock markets. Yet academic research conducted on stock markets – particularly in recent years on their so-called microstructure – remains considerable as against the research efforts invested in money markets. It is hoped, therefore, that this thesis goes some way towards encouraging others that the sterling money market is an interesting and important research topic.

Much of the research described above has been presented at seminars – several of them at the Bank of England – and at conferences. These have been invaluable for eliminating errors of fact and (I hope) of reasoning from the contents of the thesis. In particular, I thank attendants at the Money, Macro and Finance Research Group term meetings and conferences in 1992, 1993 and 1994, and seminar participants at the Bank of England, the London Business School, Oxford University and the London School of Economics. The author, however, claims responsibility for any views expressed in this thesis, as well as any errors.

Introduction

The policy instrument of central banks everywhere has, almost without exception, been a short-term nominal interest rate. Sometimes, during periods of economic disruption such as wartime or depression, short-term (and even long-term) interest rates have been fixed directly, by statutory control. More typically, however, central banks have controlled short-term interest rates indirectly, by undertaking money market and other transactions whose effect has been to achieve a desired market interest rate.

The use of an interest rate instrument by central banks has a long history. There is ample evidence that, since their establishment, most central banks typically stood ready to discount certain financial instruments, thereby accommodating increased demands for currency and bank reserves and constraining the increases in short-term market interest rates which would otherwise have emerged (see Capie, Goodhart and Schnadt 1994). In other words, central bank discount rates have usually represented a ceiling, which could be raised or lowered at the discretion of the central bank, above which market rates seldom rose. Conversely, central banks, notably the Bank of England, would sometimes quell undesirably low market interest rates by selling financial assets to the private financial sector. In this way central banks lowered

the reserves of the banking system, forced banks to discount paper with them, and made their discount rates "effective" in the market.

Throughout their history, therefore, central banks have adopted operating procedures which have been variously described in the academic literature as interest rate "smoothing", interest rate "targeting", or interest rate "control". Essentially, these terms all point towards the fact that monetary policy is implemented via operating procedures which produce a desired level of short-term nominal interest rates. With the passage of time and the development of financial markets, particularly money markets, these operating procedures have naturally evolved and have become increasingly streamlined and sophisticated. And whilst there has been considerable convergence in operating procedures across central banks, historical factors and regulatory preferences still account for some important differences in the operating procedures of central banks[2].

Theoretical and empirical interest on central bank operating procedures and short-term interest rate determination has focused almost exclusively on the US. Thus the relevance of factors such as reserve requirements, bank reserve management, open market operations and discount window borrowing, have all usually been examined within the institutional context of the Federal Reserve System and the federal funds market. One of the many insights of this literature, however, is the importance of the institutional environment in conditioning the behaviour of economic agents and hence the behaviour of economic variables such as short-term interest rates. This places limits on the ability of this literature to explain behaviour within other contexts, such as the UK, where the operating procedures of the Bank of England, as well as arrangements in the sterling money market, differ considerably from those in the US. Consequently, these remain less widely understood, and are less frequently the subject of formal examination[3].

It is the aim of this thesis to redress this imbalance by examining more closely the operating procedures of the Bank of England (hereafter also the

[2]See Kneeshaw and Van den Bergh (1989) or Batten *et al* (1990)

[3]Notable exceptions are Goodhart (1984, 1987) and Llewellyn (1990).

Bank), and their impact on money market interest rates. This task is carried out in two parts, with the emphasis moving from the general (*Part I*) to the specific (*Part II*).

Part I comprises three chapters and begins, in chapter 1, with a brief historical sketch of the development of central bank money market operations and their use of an interest rate instrument. This is done, first, to validate the claim that was made at the outset of this paper, namely that central banks have virtually always used an interest rate instrument. Secondly, it illustrates the remarkable continuity in this aspect of central banking, which was developed largely by the Bank of England during the nineteenth century, and highlights that it has been common to many central banks.

Drawing on this historical sketch, chapter 2 questions two ideas which have arisen in the literature regarding the interest rate instrument. The first is the notion that central banks are merely engaged in short-term interest rate "smoothing", where this has usually been taken to mean the elimination of seasonal fluctuations in money market interest rates. Whilst this notion is not incorrect, it is seriously incomplete since it ignores the crucial fact that the short-term nominal interest rate is the monetary instrument, whose level is chosen to attain some policy target. It is thus not just seasonal fluctuations, but *all* fluctuations from this desired *level* which are "smoothed" by central banks. What is more, seasonality in money market interest rates is then *not* prima facie evidence that interest rates are not being smoothed, since central banks may be using their interest rate instrument in a seasonal fashion.

The second notion is that central banks should not in fact be using interest rates as their instrument at all but should, instead, be using some quantity variable, such as the money base. Even though this argument has been shown to be extremely problematic[4], especially in terms of implementation, it is indeed a hardy perennial. For this reason, a recent proposal for money base control by McCallum (1989) is examined and its operational flaws are pointed out.

With a sharper concept of the interest rate instrument, both in history

[4]See for example Foot, Goodhart and Hotson (1980) or Goodhart (1989).

and in theory, chapter 3 develops a more analytical picture of the relationship between banks, the money market and the central bank and explores how this relationship permits central banks to influence, fairly precisely, the level of money market interest rates. The foundation of this picture is bank reserve management, in which the money market plays a fundamental role by allowing banks to borrow and lend reserves to meet their settlement obligations. The money market itself, however, cannot equilibrate the demand and supply of reserves at the aggregate level, providing a role for the central bank to manage aggregate reserves and, in the process, influence money market interest rates.

Part II comprises four more or less self-contained chapters, all of which specifically concern the money market operations of the Bank of England and the consequent behaviour of interest rates within the sterling money market. Chapter 4 sets the scene by describing these operations in some detail, as well as other relevant features of the sterling money market, such as the instruments which are traded and how these trades are settled. This avoids repetition and duplication in subsequent sections.

Chapter 5 traces the effects of the *maturity* of the Bank's money market lending on the term structure of sterling money market rates. The daily practice of the Bank is to invite institutions to rediscount eligible bills whose unexpired maturity is under one month[5] at the Bank's official discount rate. When this official rate is expected to change, institutions are not indifferent to the maturity of the bills which they rediscount with the Bank, since this affects their implied cost of funds. This may cause very short term money market interest rates to move in the *opposite* direction to the anticipated change in the official rate. The chapter first derives this interest rate behaviour – which may be described as term structure "pivoting" – in a simple model, and then verifies the behaviour using daily money market interest rate data.

Chapter 6 continues to deal with the Bank's use of the bill market for

[5]Alternatively, the Bank invites institutions to sell bills under repurchase agreement whose maturity is also typically less than one month.

conducting its money market intervention, although the focus here is on the *yield on eligible bills.* The fact that the Bank stands ready to buy these instruments has caused their yield to be below equivalent money market rates, for example those on unsecured wholesale sterling loans (i.e. LIBOR). Whilst this is perfectly consistent with the existence of a liquidity premium on bills, this premium has not remained constant over time and has tended to widen significantly over recent years. This suggests the presence of constraints on the supply of acceptances, which is how bills are created. The chapter examines various possible constraints, and finds that the Bank's own regulation of the bill market may be a primary candidate.

One of the problems associated with an excessive liquidity premium on bills is that these instruments become costly to hold, either because they are difficult to fund profitably or because they involve a capital loss when rediscounted with the Bank. These costs create incentives for those institutions which hold bills – predominantly the clearing banks – to engage in strategic behaviour in order to offset these costs. In particular, this strategic behaviour on the part of the clearing banks aims to influence overnight interest rates in a manner which allows these banks, on average, to borrow overnight funds more cheaply. To be successful, however, this behaviour must not be predictable, with the result that strategic behaviour may increase the volatility of the overnight interest rate. Chapter 7 thus examines the behaviour of the sterling overnight rate and looks at the factors which influence both the mean and the variance of this rate.

Chapter 1

The Interest Rate Instrument in History

Although the Sveriges Riksbank was established before the Bank of England – in 1668 – it took another two centuries for the former bank to achieve the status of a fully fledged central bank[1]. Indeed, it was not until 1897 that the Riksbank obtained the monopoly note issue in Sweden. The Bank of England (the Bank), on the other hand, assumed the role of a monopoly note issuer relatively early[2] and developed many of the functions associated with this role well before the other central banks which were established in Europe during the nineteenth century. One of these functions was to influence the level of short-term interest rates – which is the focus of this thesis – and hence this chapter traces the origins and development of the interest rate instrument beginning, in section 1, with the early operations of the Bank of England. The Bank Charter Act of 1844, which placed strict limits on the fiduciary issue of the Bank, represents a natural dividing line through the Bank's early activities as regards interest rates. Up until the Act, the Bank's

[1]This chapter borrows extensively from Capie, Goodhart and Schnadt (1994), which contains summaries of, and provides references to, the development of central banking in England, Europe, the Americas and elsewhere.

[2]Officially in 1844, although Bank of England notes were more widely used than other bank notes long before this.

discount rate – known as Bank rate – was not actively deployed by the Bank in response to gold outflows, although by 1844 a well-developed system had evolved whereby the banking system could obtain reserves from the Bank. From 1844 onwards, and until the start of the first world war in 1914, this system of reserve provision continued to function as before, but the Bank's discount rate was actively applied to protecting its gold reserve. Furthermore, during this latter period several techniques were developed, now summarised by the term open-market operations, for making the Bank's discount rate effective in the London money market.

Section 2 looks at some of the early parallels to the Bank's use of Bank rate, particularly in Europe, where central banks were also beginning to use their official discount rates to protect their gold reserves and – although much less frequently – undertake open-market operations. Section 3 briefly traces US banking experience before the founding of the Federal Reserve System (the Fed) in 1914. Like all other central banks before it, the Fed also maintained a discount facility (known as the discount window) through which it provided reserves to the banking system at a discount rate of its choosing. However, since the founding of the Fed coincided with the start of the first world war, which considerably altered monetary and economic conditions in the US and elsewhere, the discount policies of the Fed and other central banks reflected these changes. In particular, as discussed in section 4, these policies reflected the needs of government finance rather than the preservation of gold reserves. Moreover, discount policies were typically augmented with a host of direct measures for influencing capital flows and the availability of credit, which persisted until after the second world war. Thereafter, as described in section 5, the primary monetary instrument in most countries again became the official discount rate. By this time, however, central banks had developed a better understanding of the relationship between bank reserves and money market interest rates, and also had the ability to manipulate this relationship through open-market operations. As these operations became increasingly sophisticated, so too did central banks' control of money market interest rates.

1.1 Origins – The Bank of England

1694-1844

The Bank of England was established in 1694 as the first chartered bank in England, and was authorised to issue bank notes (redeemable in gold), trade in bills of exchange and bullion, and accept deposits[3]. Its position as a monopoly note issuer was virtually guaranteed from the outset due to legislation which prevented the establishment of other joint stock banks (Smith 1936). Thus, while numerous small, private banks arose during the eighteenth century, none competed seriously with the Bank in the issue of notes. By 1770 most London bankers had ceased to issue notes altogether and used Bank of England notes for inter-bank settlement in what had become a well-developed clearing system.

The restrictions on joint stock banking meant that the early banking system which emerged in England comprised a large number of small banks, without any significant branch network. This in turn provided a role for a market in bills of exchange (short-term debt instruments) as a mechanism for redistributing reserves (specie or Bank notes) amongst banks. By purchasing bills, surplus banks could employ otherwise idle balances to acquire interest-bearing assets whilst at the same time providing reserves to banks (and other borrowers) in a geographically separate area. Specialist bill brokers quickly developed to facilitate this trade in bills and an active discount market was in place in London by the end of the eighteenth century (King 1936).

Given the role of the discount market in allocating reserves amongst banks, the London market discount rate responded to the aggregate demand for reserves. An increase in the demand for reserves, for example, would cause bills to be sold into the discount market, thereby driving up the market discount rate. Whilst this would encourage surplus banks to lend reserves by buying these higher-yielding assets, the aggregate supply of reserves was clearly constrained by the fact that banks were unwilling to let their re-

[3]For an early history of the Bank see Andreades (1909), Bowman (1937), Clapham (1944) or Giuseppi (1966).

serve holdings fall below some lower limit for prudential reasons. Increased currency withdrawals, usually due to seasonal increases in the demand for currency, would thus result in sharp increases in the market discount rate (King 1936 and Scammel 1968).

As the largest bank, the Bank of England routinely rediscounted (purchased) bills in the discount market, although concerns about credit risk meant that the Bank would consider as eligible only those bills which carried at least two good London names, and which had an unexpired maturity not exceeding 65 days. Whilst the Bank's discount rate – Bank rate – could be altered at the discretion of the Bank, the absence of any well-articulated linkages between Bank rate and the wider economy meant that its directors would have seen little reason to change this rate during the Bank's early years.

In any event, Bank rate, along with all other interest rates, was subject to a legal maximum of 5% under Usury Laws which had been in force since 1714. Consequently, Bank rate was held at 4% for most of the eighteenth century (it was raised to its legal maximum from 1773 onwards; see Hawtrey 1938).

The Bank's willingness to rediscount bills placed a ceiling on the market discount rate, since banks and bill brokers would simply sell bills to the Bank whenever the market discount rate threatened to rise above Bank rate. The Bank thus found that its purchases of bills would increase dramatically at certain times of the year – when cash demands were high – and then fall again once the market discount rate fell below Bank rate. In this way, the Bank's discount facility provided an automatic mechanism for increasing the supply of reserves, which on most occasions prevented seasonal fluctuations in the demand for currency from developing into liquidity crises. Concerns about the level of its reserves, however, sometimes led the Bank to reduce, or attempt to curtail altogether, its discounting activities. This only accentuated the liquidity crisis, with the result that the Bank invariably resumed its lending to avoid a full-blown panic (see chapter 2).

One such crisis occurred in 1825, precipitated in part by the Bank's refusal

to discount bills. Although the Bank subsequently reversed this decision when the crisis worsened (a detailed account is given in King 1936, pp.35-70), many country banks had, by this time, failed. This prompted a significant reorganisation of the banking system, both from within and on the part of the Bank of England. The changes from within the banking system stemmed from the hesitancy that was shown by the Bank in discounting bills. To reduce their reliance upon the Bank, most country banks chose to stop discounting bills with the Bank, and held liquidity in the form of call deposits with the bill brokers instead. The bill brokers – who secured these deposits by holding bills – continued to discount with the Bank of England, and so became the regular counterparties in the Bank's bill transactions. Since they now held considerable bill portfolios, and usually stood ready to buy or sell bills, these bill brokers were transformed into principals and became known as discount houses.

The 1825 crisis also highlighted the weaknesses of the banking system that had evolved under the restrictions on joint stock banking. These restrictions were thus lifted the following year, albeit only for those banks operating outside London. London banks had to wait until 1833, when the Bank's charter was renewed, to gain the privilege of operating on a joint stock basis. At the same time, the Bank obtained exemption from the Usury Laws, and for the first time could raise Bank rate above 5%. However, it would be another ten years before the Bank would begin to utilise its freedom to vary Bank rate: between 1833 and 1844 Bank rate was only raised above 5% once (during a crisis in 1839) and was changed only eight times (Hawtrey 1938).

During the early nineteenth century the Bank began to show an increasing awareness of the relationship between fluctuations in the reserves of the banking system and the market rate of discount. Seasonal changes in the demand for currency were one source of such fluctuations. Another source of fluctuation was related to the Bank's role as banker to the government. This function meant that the Bank held the accounts into which government receipts flowed and out of which disbursements were made. Since taxes were paid into these accounts every quarter, just before dividends on government stocks were paid out, the discount market would experience a periodic

shortage of reserves, which drove the market discount rate up to Bank rate, followed by a surplus of reserves, which drove the market rate below Bank rate again. This led the Bank to introduce, in 1829, a system of regular advances (against security) just prior to the four dividend dates in the year (see Clapham 1944 or King 1936). These advances were usually made at, or near, the prevailing market rate: this gave institutions an incentive to apply for advances in anticipation of higher rates, thereby smoothing – to some extent – fluctuations in market rates.

By 1830, therefore, a well-developed – but indirect – system was in place whereby the wider banking could obtain reserves from the Bank of England. In the event of a reserve shortfall, banks would withdraw call loans from the discount houses. The houses, having secured these loans by purchases of bills, would rediscount bills at the Bank of England. This mechanism has remained virtually unchanged until the present day and has shown a remarkable continuity against the numerous innovations which have changed the appearance of the sterling money market[4].

1844-1914

The Bank Charter Act of 1844 marks the beginning of a period during which Bank rate was actively used as a monetary instrument. As noted by Hawtrey (1938, p.23):

> "The primary purpose of the Act of 1844 had been to secure the convertibility of the Bank of England note by a limitation of the issue. If the system relied on for that purpose [i.e. simply refusing to discount] was to be suspended, some other safeguard was essential. The desired alternative was found in a high Bank rate, the deterrent effect of which ... had been gaining recognition for a number of years".

Since closing the discount facility in order to uphold the reserve minimum stipulated in the 1844 Act was not feasible, Bank rate gradually became seen as the instrument whereby the gold reserve of the Bank could be protected in

[4]Obviously the Bank's operations have not remained totally unchanged. Chapter 4 provides a detailed description of the Bank's current operations. Chapter 6 also charts the latter evolution of the discount market and the bank's operations.

the face of sustained gold outflows, either for internal or for external reasons. Figure 1.1 illustrates this clearly: Bank rate remained virtually constant before 1844, whereafter it was changed frequently, with more than five changes per year becoming the norm.

Whilst the Bank had clearly taken a step in the direction of becoming a central bank in 1844, commercial considerations remained important to its directors, who realised that maintaining Bank rate at high levels would result in the attrition of the Bank's portfolio of earning assets (bills). Discount houses had little incentive to discount bills with the Bank when Bank rate rose above market rates. This steadily eroded the Bank's bill portfolio as its existing holdings matured. As illustrated by figure 1.2, such commercial considerations meant Bank rate began to exhibit an increasingly seasonal pattern over time, as the rate was lowered during the periods of so-called "easy money" which followed periods of reserve stringency. This pattern became especially clear after the adoption of the gold standard by many other countries after 1870, for reasons which are discussed again below.

Periods of "easy money", when a surplus of reserves pushed market discount rates below Bank rate, posed another problem for the Bank of England. Low rates made it less profitable for country banks to invest their surplus reserves in the London market, which then aggravated the rise in rates when activity revived. They also encouraged expansion at times when the Bank was expecting a stringency of gold in the near future, which meant Bank rate would have to be raised to a higher level to sustain reserves Occasionally, therefore, the Bank wished to raise market rates to the level of Bank rate so as to make its rate "effective", and the directors of the Bank realised that this could be achieved by withdrawing reserves from the banking system. But how was this to be done?

Selling bills would have been an obvious solution, were it not for the fact that the Bank's bill portfolio was small at precisely those times when market rates were low relative to Bank rate. Similarly, borrowing reserves would have done the job, but conflicted with the Bank's avowed practice of not paying interest on its deposits. This left outright sales of Consols (government perpetuities), which the Bank held in sizeable quantities. However, this may

have imposed losses on the Bank since Consols would have to be sold when longer-term rates were likely to be above short-term rates (ease in the discount market) and purchased again when longer-term rates were below short-term rates (stringency in the money market).

These constraints led the Bank to engage in repurchase agreements against Consols: that is, Consols were sold spot and simultaneously repurchased forward. This transaction effectively meant that the Bank borrowed short-term funds, but avoided the capital losses which could have been incurred on selling Consols outright and buying them back at a subsequent date. According to Clapham (1944, p.295) such repurchase agreements were first employed during the 1830s, but they did not become a regular feature of the Bank's operations until after 1870. Even then, problems remained with this technique, because the Stock Exchange participants to whom the Bank sold Consols were themselves insufficiently reliant upon the discount market for funds. This led to instances where the Bank sold Consols, only to find that the institutions who bought these securities were in turn funding these purchases with advances from the Bank. Thus the Bank's operations would leave bank reserves unchanged and so too the market discount rate. To counteract this problem the Bank sometimes took the more direct step of borrowing funds (via repurchase agreements) directly from large lenders who wished to lend funds in the discount market (Sayers 1957). The Bank also sometimes sold Consols directly to the large joint stock banks instead of on the Stock Exchange.

1.2 Early Parallels – Europe

Whilst the Bank of England had been virtually the only nascent central bank during the eighteenth century, the nineteenth century saw central banks (or their precursors) being established in many European countries. Amongst others, these included the Banque de France (1800), the Nederlandsche Bank (1814), the Prussian Bank (1846, which became the Reichsbank in 1876) and the Banque Nationale de Belgique (1850) (see Capie, Goodhart and Schnadt 1994). In each case an important function of these banks was to discount

bills, usually of a good quality and a short maturity. The result was that they all provided a discount facility similar to that of the Bank of England. As in England, Usury Laws (or government decrees) typically prevented these banks from raising their official discount rates beyond 5%. The Banque de France, for instance, maintained its rate at 4% between 1820 and 1847, whilst the precursors to the Reichsbank hardly ever changed their discount rates from 4%. As in England – and more than likely *because* of the more active use of Bank rate there – the European central banks also began to vary official discount rates more frequently after the middle of the century.

Data on the official discount rates of the European central banks between 1844 and 1900 in Palgrave (1903) illustrates this development clearly. Before 1850, virtually no changes are recorded in official discount rates. Between 1850 and 1900, however, the official discount rate was changed 109 times in France, 155 times in Germany, 160 times in Belgium and 173 times in Holland[5].

The parallels do not end there. The emergence of the seasonal pattern in Bank rate described in the previous section also occurs in the discount rates of the continental central banks, especially after 1875 (tables are provided in Palgrave 1903). By this time nearly all the countries considered were on a gold standard. The reasons for this seasonal variation are the same as before: the need to protect gold reserves against gold outflows caused official rates to be raised, whilst the accumulation of excess reserves and the desire to invest these in interest-bearing assets caused rates to be lowered again. The President of the Reichsbank in 1880, when questioned why the Reichsbank had begun to discount at a preferential rate below its official discount rate, gave the following reply:

> "I can assure you that at no time and under no circumstances has the management of the bank taken an action for the purpose of paying the highest possible dividends .... But, gentlemen, I cannot be completely indifferent as to whether funds of the bank are employed or lie sterile in its vaults ... As I see it, the Reichsbank should not look for business, but neither should

[5]This is still less than in England, where Bank rate was changed no fewer than 392 times.

> it decline a proper opportunity to invest its funds in safe, bankable paper" (cited in Bopp, 1956, p.38).

During the era of the international gold standard, then, many central banks followed the lead taken by the Bank of England and employed their official discount rates to regulate and protect their gold reserves. This did not mean, however, that discount policy operated according to rigid rules, free from judgement or discretion. To the contrary, to quote another President of the Reichsbank in 1890:

> "In determining the discount policy of a central bank, account must be taken not only of the size but also of the character of the monetary demands made of it. An export of gold requires sharper countermeasures than a temporary withdrawal of gold into domestic circulation. A monetary demand based on overspeculation and overproduction requires sharper restrictions than a periodic increase in monetary demand at the turn of the quarter of the year which is grounded in the whole structure of the normal process of payments. These differences cannot be expressed in numbers and therefore discount policy cannot be regulated according to a purely mechanical principle" (cited in Bopp, 1953, p.30).

During the operation of the international gold standard some of the continental central banks also began to develop similar techniques to those of the Bank of England for making their discount rates effective. The Reichsbank is the most prominent example, and began to use open-market sales of Treasury bills in 1901 to withdraw surplus reserves from the banking system (Bopp 1956).

1.3 Late Followers – The Federal Reserve

Throughout the nineteenth century, and until 1914, the United States had no institution which resembled the central banks that had been established elsewhere. Two national banks had been in operation over the periods 1791-1811 (the First Bank of the United States) and 1817-1836 (the Second Bank of the United States), but these were distinguishable from other banks mainly because they acted as depositories for public funds (see Timberlake 1993). Limited central banking functions were also carried by the Independent Treasury, formed in 1841 (see Taus 1943), including what may be seen as the first open-market operations which were carried out "with a view towards relieving pressure upon the money market of New York" (cited in Timberlake 1993, p.79). However, these operations remained isolated instances rather than a coherent strategy for reserve provision.

Since 1863, US banking operated according to the so-called National Banking System, under which all note-issuing banks were compelled to cover their note issues by holdings of Treasury debt (see Friedman and Schwartz 1963). Whilst this arrangement effectively created a uniform national currency, it also made the supply of notes inelastic, since banks first had to obtain government debt before they could expand their note issue. This problem was compounded further by national restrictions on branch banking, as well reserve requirements. All these features resulted in a banking system which was extremely poor at dealing even with seasonal (and hence fairly predictable) variations in the demand for cash, let alone unexpected changes in cash demand. As in England, where the London discount market had developed for purposes of redistributing reserves amongst geographically isolated country banks, the New York money market arose to recycle reserves from surplus to deficit banks. It thus became commonplace for country banks to place their surplus reserves with New York banks in the form of call loans (see Goodhart 1972). As in England, however, *aggregate* variations in the demand for reserves caused sharp fluctuations in New York money rates and, in such circumstances,

> "... call loans proved to be among the most illiquid of assets. Experience

> taught the country banks that their correspondents in New York found it difficult to cash their deposits in such a situation. Consequently, immediately the slightest indication of a defect occurred, there was a scramble by the country banks to withdraw their balances *en bloc* from New York .... The New York banks were driven to suspend payment, and those country banks who had not withdrawn their funds in time had to suspend also" (Smith 1936, p.139).

In the absence of any mechanism for accommodating changing demands for currency (reserves), a role which was played by the discount facilities of the central banks described above, it is therefore unsurprising that the National Banking System was prone to frequent banking crises, with severe panics occurring in 1873, 1884, 1890, 1893 and 1907 (Mishkin 1991 and Calomiris and Gorton 1991).

Tallman and Moen (1993), and more recently Donaldson (1994), have confirmed that seasonal liquidity factors were a significant driving force behind these banking panics, a view which was held by many early commentators on the subject (e.g. Sprague 1910 and Kemmerer 1910). Indeed, without a central bank to provide reserves, money market interest rates would rise to extremely high levels as banks attempted to obtain reserves in the New York money market. Weekly data examined by Donaldson indicates that during the crises of 1873, 1884 and 1907, short-term interest rates reached levels of 161%, 66% and 50% respectively, levels that were far in excess of anything seen in London or on the continent. According to a report by the Senate Banking and Currency Committee in 1907:

> "... during the year 1907 the range of interest for call money was from 2 to 45 percent in January, from 2 to 25 percent in March, from 5 to 125 percent in October, from 3 to 75 percent in November, and from 2 to 25 percent in December ..... The blighting effect of these violent fluctuations of the interest rates is demonstrated by the rate charged for 90-day time loans, which during November and December, 1907, were running as high as 12 to 16 percent, with no business done in time loans of a longer period during the entire month of November and no business being done at times on prime commercial bills during the same months" (quoted in Burgess, 1946, p.196).

The severity of the crisis of 1907 prompted the formation of the National

Monetary Commission to study the domestic and international monetary system and to seek ways of increasing the "elasticity" of the US currency supply.

It was relatively clear that the outcome of the Commission would involve the establishment of a central bank, although the size of the US as well political concerns about the governance of such a bank militated against the formation of a single "United Reserve Bank" as proposed by Paul Warburg (1911). Instead, the Commission's Report, which was published in 1910, laid the foundation for a Federal Reserve *System* (or Fed), This began operations in 1914, and was comprised of 12 Federal Reserve banks, who were the owners of the Fed, acting under the umbrella of the Federal Reserve Board[6].

Aside from its ownership features, the Fed was not unlike most other central banks. Like them, the individual Federal Reserve banks were authorised to operate a discount facility – dubbed the "discount window" – under which they could purchase bills of exchange (known as banker's acceptances in the US) at a rate chosen by them, thereby providing an automatic mechanism for providing member banks with reserves. According to Burgess (1946):

> "The Reserve Banks ordinarily stand ready to buy the eligible banker's acceptances which may be offered to them. There have indeed been times when eligible acceptances have been refused; but generally speaking the reserve banks have been passive rather than active in their specific purchases of acceptances and *have expressed their policy by changes in the rates at which purchases were made*" (p.43; emphasis added).

Initially, each Federal Reserve bank posted its own discount rate (or set of rates, where more than one type of instrument was rediscounted). Although discount rates could differ between Reserve banks until the second world war, changes in these rates had to be approved by the Board (Meulendyke 1989). Furthermore, open-market operations did occur, although these initially took the form of outright *purchases* of securities, usually with a view to increasing the Reserve banks' interest-bearing assets, or to foster the market in banker's

[6]For details on the structure and early operations of the Fed see Wicker (1966) or Beckhart (1972).

acceptances. They were clearly not undertaken to withdraw reserves in order to make discount rates effective. It was only some years later, in the early 1920s, that the effect of these individual purchases on short-term money market interest rates, and in turn on credit conditions, was recognised[7]. The Governor of the New York Fed at the time, Benjamin Strong, thus proposed that discount policy and open-market operations should become coordinated, leading to the formation of the Open Market Investment Committee in 1923[8].

1.4 Wartime Disruption

The Fed had not even begun operations when hostilities broke out in Europe in July 1914. In England, this led to a massive surge in discounting at the Bank of England. Although Bank rate was increased, the demands for currency remained so high that (emergency) legislation was passed which permitted the Bank to increase its note issue beyond any limit prescribed in law. It was not long before the currency which had gone into circulation began to return to the banking system, swelling banks' reserves. This increase in reserves was amplified by the expansion of government spending, much of which was financed by short-term borrowing in the form Treasury bills and so-called Ways and Means advances provided by the Bank of England. The net result was the creation of a large surplus of reserves which, in the absence of action by the Bank, caused short-term market rates to fall. Thus an innovation was called for:

> "The Bank proceeded to borrow at call in the discount market all sums offering at a fixed rate of interest ... At the same time, the Government, instead of putting Treasury bills up to competitive tender offered them without limit of total amount to the market at a fixed rate of discount. The effect was to prevent the rates quoted in the discount market for call money and bills falling below these fixed rates. If the rates offered failed to attract as much money as the Government needed, the deficiency was made good

[7]See chapter 3 for further discussion on this issue.

[8]Which eventually – in 1935 – became the Federal Open Market Committee. See Wicker (1966).

> by advances from the Bank of England ... which injected a fresh supply of money into the market and made it more disposed to lend. Thus the market was really governed by the Treasury bill rate, and Bank rate was practically superseded by it. This system continued ... throughout the war and till April 1921" (Hawtrey 1938, p.130).

Notice that, under such conditions seasonal fluctuations in currency use became irrelevant as far as money market interest rates were concerned: large reserve surpluses, together with large holdings of Treasury bills, meant that banks were never in a position where they had to meet increased demands for reserves by discounting paper in the market or at the Bank (this point is taken up again in chapter 2).

In the US the picture was different, although the implications for seasonal fluctuations were the same. Rapid expansion in economic activity led to an influx of gold and to a massive increase in borrowing from the Fed's discount window. Discounts and advances to member banks rose from average levels of around $30 million in 1915-1916 to a peak of $2500 million in 1920. Over the same period the average number of member banks accommodated through the discount window rose from around 650 to just under 5000 (Burgess 1946). Not surprisingly, therefore, seasonal fluctuations in currency use were accommodated by seasonal variation in discount window borrowing, with short-term market interest rates staying at or near the level of the discount rate. Levels of borrowing remained positive throughout the 1920s, although they were reduced somewhat by gold inflows and Reserve bank purchases of securities. Interestingly, this situation was totally reversed during the 1930s, when inflows of gold became so large that bank's held unprecedented levels of excess reserves (see Morrison 1966). The Fed consequently used its power, acquired under the Banking Act of 1935, to double banks' reserve requirements, thereby effectively withdrawing reserves.

Variants of the UK and US experience doubtless occurred in many countries, with banks either holding a large surplus of reserves, or becoming indebted to their central bank (see Eichengreen 1992). In each case, however, central banks sought to maintain short-term interest rates at the levels chosen by the monetary authorities, although this level increasingly came under

the influence of governments during and between the wars. As Sayers (1957) put it:

> "Under the impulse of war finance, the technique of the central bank is manipulated so as to conform to events rather than to control them. Central bankers become the slaves of deficit finance" (p.20).

This meant that interest rates – sometimes both short and long, as in the US – were maintained at extremely low levels for long periods of time, as governments pursued so-called "cheap money" policies to reduce the cost of government borrowing and to stimulate capital formation. Bank rate, for example, was held at 2% for most of the 1930s, with market rates usually remaining even lower. This state of affairs persisted throughout the second world war and until 1951[9]. Similarly, the US discount rate remained at very low levels, and was usually ineffective even at these levels. During the war the US Treasury bill rate was officially pegged by the Fed, at $\frac{3}{8}$%, as were longer term government bond rates (Timberlake 1993).

Deficit finance, and the concomitant drive for cheap money, had another implication for the interest rate instrument:

> "... experience has taught us that the interest rate ... is a blunt instrument: it works too slowly in encouraging or discouraging in the right direction .... So we have moved towards the idea that the interest rate should be held pretty stable, and on the low side. And if this is done, it follows that the desired changes in activity must be encouraged, and the undesired changes discouraged, by other weapons" (Sayers 1957, p.31).

These "weapons" comprised direct controls of every imaginable kind. Exchange controls, credit ceilings, interest rate ceilings, and liquidity requirements all became a common feature of the financial landscape in many countries during and between the wars. It is important to note, however, that these controls were meant to *augment* the interest rate instrument rather than to replace it. Central banks everywhere continued to operate as they

[9]As before, it was the Treasury bill rate which became the relevant official rate during the war. See Sayers (1957).

had previously, namely to offer a discount facility under which certain eligible institutions could borrow reserves by selling or pledging certain eligible assets to the central bank. Irrespective of the other monetary weapons that were deployed by governments and central banks, therefore, discount windows were never closed.

1.5 Open-Market Operations

In one respect, deficit finance created an advantage for interest rate management. Almost everywhere, central banks found that their portfolios of short- and long-term government securities were much larger as a result of the government debt issues which occurred during the wars (and the Great Depression). When conditions normalised, this naturally improved their ability to undertake open-market sales or purchases of assets in order to make their discount rates effective. Unwanted variations in short-term money market interest rates could now be eliminated from *both* directions. Increases in the demand for reserves could be accommodated either via automatic discounts (or advances) or via open-market purchases, both of which prevented rates in the market from rising. Alternatively, decreases in the demand for reserves could be "mopped up" by an automatic decrease in discounts or, if necessary, by an open-market sale of assets (either on an outright or a repurchase basis), thus preventing rates in the market from falling. Government war finance, together with the evolution of broader and more liquid money markets, thus paved the way for a more precise and active role for (money) market operations as a means of achieving a desired level for short-term interest rates.

Here, the Federal Reserve took the lead. In 1951 it ceased fixing the prices of longer-term government bonds, and began to operate according to a "bills-only" technique: that is, the Fed would buy or sell Treasury bills to increase or decrease bank reserves (see Meulendyke 1990). The counterparties in these transactions were a relatively small group of specialist institutions – known as primary dealers – who undertook to make markets in government

securities. The Fed also frequently augmented these operations with repurchase agreements involving government securities and banker's acceptances, although these transactions were, until the 1970s, only undertaken with non-bank primary dealers. With the development of the federal funds market in New York in the mid-1950s[10], the federal funds rate emerged as an indicator of reserve availability within the banking system. If reserves were scarce, banks would be borrowing from the discount window, and the federal funds rate would be driven to equality with the Fed's discount rate. If reserves were plentiful, the federal funds rate would begin to fall, suggesting that the Fed should undertake open-market sales (i.e. borrow reserves).

These operations were only slightly more sophisticated than those of the Bank of England, whose techniques remained much the same as before. An increase in the demand for reserves meant that the discount houses would approach the Bank to sell bills (or obtain advances). Reserve surpluses were removed through the sale of Treasury bills. As already noted, this mode of operation remained basically unchanged until 1980 when the Bank, like the Fed, resorted to the more frequent use of open-market operations as a mechanism for supplying reserves. These changes, and the Bank's current money market operations, are presented in detail in chapter 4.

Unlike the Fed and the Bank of England, most European (and other) central banks did not initially undertake open-market operations to vary the level of bank reserves in the post-war period. Instead, they continued to rely upon their discount facilities to provide reserves until the 1970s and 1980s (see Kneeshaw and Van den Bergh 1989). Changes in reserve requirements were also often used to achieve the same effect as open-market operations, with Germany being a particularly good example (see Deutsche Bundesbank 1989). Raising the level of reserve requirements has exactly the same result as an open-market sale of securities[11]: *both reduce the existing level of reserves*

[10]This market was one in which member banks could lend their reserve balances at the Fed to other member banks on an overnight (or longer) basis. Consequently, "the" federal funds rate refers to an overnight interest rate (see Minsky 1957, Boughton 1972 or Goodfriend and Whelpley 1986).

[11]Of course, higher reserve requirements impose a higher tax on banks, and from this

relative to the level of required reserves and, by increasing the scarcity of reserves, potentially force banks "into" the discount window.

The use of reserve requirements in this way (i.e. as a mechanism to support the central bank's interest rate instrument) is a far cry from the conventional interpretation of reserve requirements, which treats them as a constraint on the capacity of the banking system to expand its balance sheet. Although this interpretation of reserve requirements is presented in most monetary textbooks, it is invalid since a necessary condition for its validity is that the supply of reserves is *quantitatively constrained* by the central bank. The historical description up to this point indicates that this has *never* been the case, at least not as long as there have been central banks[12]. The persistence of this interpretation, then, may be ascribed either to the belief that central banks have sought to place absolute limits on the level of bank reserves (which is clearly counterfactual) or that central banks *should* place absolute limits on the level of bank reserves. In order to maintain continuity at this stage, this latter issue is taken up again in chapter 2.

A useful summary of more recent changes in the manner in which central banks have determined the level of short-term nominal interest rates is provided by Kneeshaw and Van den Bergh (1989), who note that:

> " ... recent changes in central banks' money market operating techniques and procedures display several common tendencies. *Firstly*, steps have been taken in many countries to restrict or discourage recourse by the banks to conventional open-ended central bank credit facilities .... *Secondly*, central banks in many countries have in recent years made increasing use of flexible market operations ... for counteracting influences on reserves over which central banks have little direct control. *Thirdly*, in many countries the burden of reserve requirements has been lowered in recent years, and only limited use is now made of changes in reserve requirements as a means of making adjustments in banks' reserve positions" (pp.7-8).

Nonetheless, closer observation reveals many differences, which may involve the frequency of operations, the official counterparties, the type of

perspective are not the same as a sale of securities.

[12]For a discussion of these and other issues related to reserve requirements see Goodfriend and Hargreaves (1983) and Stevens (1991).

transactions, and limits on the access to the discount facility. These differences, in turn, have implications for the behaviour of short-term interest rates and are examined again in chapter 3 and also throughout *Part II*, in the context of the sterling money market.

1.6 Figures

Figure 1.1.
Bank Rate 1800-1914:
annual minimum, maximum and number of changes

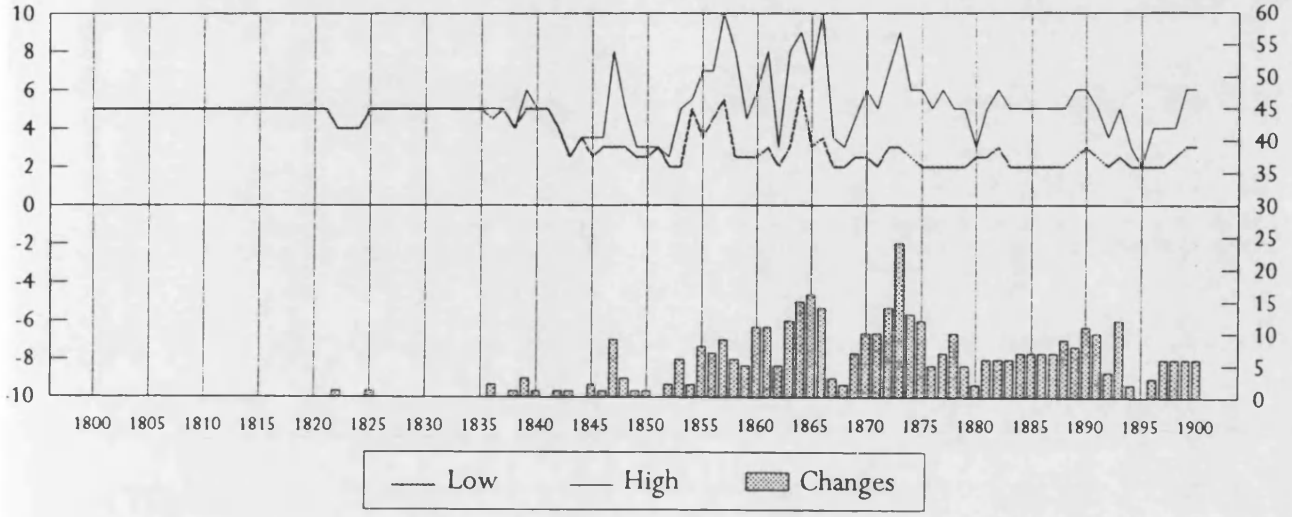
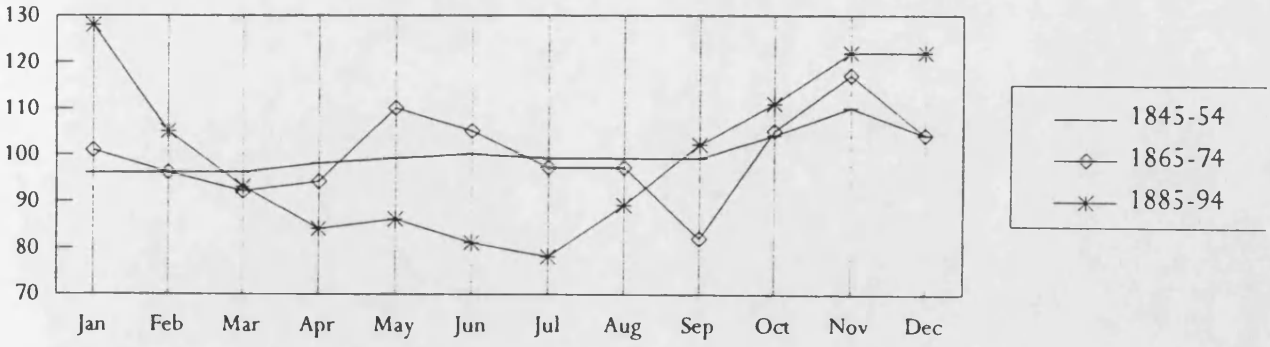


Figure 1.2.
Bank Rate in the nineteenth century:
seasonal fluctuations



Chapter 2

The Interest Rate Instrument in Theory

The historical sketch presented in the previous chapter has illustrated the continuity in the operations of the Bank of England and those of other central banks. It has argued that these operations have been consistent with the interest rate as the monetary instrument. Although this argument is strongly suggested by the historical record, the academic literature on this subject has been more ambivalent. It is the purpose of this chapter to highlight – and bring into question – two prominent sources of this ambivalence.

First, section 1 explores the idea that central banks are engaged in interest-rate smoothing. This has usually been taken to mean that central banks seek to eliminate seasonal fluctuations in money market interest rates. Whilst not incorrect, this notion is seriously incomplete, since it ignores the crucial fact that short-term interest rates are the *instrument* of monetary policy, whose level is chosen by the central bank to achieve some policy target.

Second, section 2 questions the idea that the monetary instrument can, and possibly even should, be some quantity variable, like the money base, instead of a price variable, i.e. a short-term interest rate. This has always been rejected by central bankers, as well as by academics who stress the institutional aspects of monetary policy, usually because of implementation

problems. Nonetheless, many academics persist in their belief that a (simple) choice exists over the monetary instrument. To dispel this notion, a recent proposal for money base control is examined, and its operational problems are highlighted.

2.1 Interest Rate Smoothing

2.1.1 The disappearance of interest rate seasonality

Sayers (1976) provides perhaps the most concise statement of central bank interest rate management as it was practised by the Bank of England (and other central banks) in the years leading up to 1914:

> "... the Americans had at this time a rather inelastic monetary system, and one of the biggest sources of demand for gold from London was to meet variations in the circulation requirements of the American economy. This was an economy still substantially geared to annual cropping, and it was no coincidence that the Bank had more often than not to face an autumnal drain of gold to New York. Was such a seasonal movement ... to be ignored by the Bank? .... Discretion was ... expected, whatever the Act of 1844 and its prophets had said. It was perhaps within narrow limits, but judgement was undoubtedly called for, in deciding whether the week's conditions (or indeed the day's) necessitated a movement of Bank rate; in this sense, *the Bank was already in its manipulation of Bank rate engaged in central banking*" (emphasis added).

After 1914, this scenario changed, especially as regards the behaviour of interest rates. Seasonal fluctuations in short-term nominal market interest rates virtually disappeared, both in the US and in all European countries, as did seasonal fluctuations in the official interest rates of the Bank of England and the other continental central banks. Moreover, interest rate seasonals have never subsequently reappeared to any significant extent (see Mankiw and Miron 1986).

The most frequent explanation for this disappearance of seasonality is the founding of the Fed in 1914. Whilst this is certainly not an incorrect

explanation, it is argued here that it omits an important part of the story. Furthermore, this explanation has led to a somewhat ambiguous characterisation of the interest rate instrument in the literature, which this section seeks to clarify.

Miron (1986) was the first to argue formally that it was the creation of the Fed which finally furnished an elastic currency in the US, thereby accommodating the reserve fluctuations which drove these interest rate seasonals. Clark (1986) challenged this argument on the grounds that it inadequately accounts for the *worldwide* disappearance of seasonality, and suggests that the dissolution of the international gold standard may have been a key factor. Taking up this challenge, Barsky, Mankiw, Miron and Weill (1988) proposed an explanation in which the creation of the Fed was reinstated as the primary causal factor: by eliminating the need for seasonal gold flows from other countries, they argued, the establishment of the Fed also eliminated the seasonal interest rate reaction of other central banks to these gold flows.

The historical account of the previous chapter suggests that, if anything, *both* the suspension of the gold standard and the founding of the Fed appear to be relevant in explaining the disappearance of seasonality. However, a third factor, rarely mentioned in the literature, also seems essential.

Recall that by 1914 the Bank of England was the only central bank which sought to make its official discount rate effective with any regularity, and did not even do so all the time. Elsewhere, central banks usually let market rates fall below official rates or, if anything, lowered their discount rates down to market rates in order to maintain their stock of earning assets. The newly formed Federal Reserve banks would have behaved this way. Without a significant portfolio of financial assets, they could not have countered any seasonal fall in market rates below the discount rate. This implies that seasonal variations in market rates, whilst perhaps not as extreme as before, would have persisted even *with* the Fed and *without* the gold standard.

Some evidence for this may be found by going back to the experience of the Bank of England in the first part of the nineteenth century, when the banking system experienced seasonal fluctuations in reserves – as it does to

a limited extent even today – due to variations in domestic currency use. Although Bank rate did not vary, market discount rates did: seasonal reserve ease caused rates to fall below Bank rate in the absence of any operations by the Bank to withdraw reserves. The establishment of a central bank in the US will, similarly, have placed a *ceiling* on market interest rates. Rates would have been free to fall below this ceiling, however, and would have done so on a seasonal basis unless member banks for some reason remained permanently indebted to the Fed, or the Fed undertook open-market operations to make its discount rate effective.

The disappearance of seasonality, then, was not only due to the founding of the Fed or the suspension of gold convertibility. It was also a direct result of the first world war. In the UK, after the initial liquidity shock resulting from the outbreak of war had been weathered, the banking system became flooded with reserves. Seasonal fluctuations in reserves, therefore, had little impact on interest rates as long as the Bank of England stood ready to borrow reserve surpluses from the banking system. In the US, banks relied so heavily on the Fed's discount window that seasonal variations in reserves also had little impact on interest rates. These liquidity disturbances, which left banks more or less permanently in a position of aggregate surplus or shortage, continued to plague many banking systems until after the second world war. By this time, however, central banks were better equipped for keeping market rates in line with official rates. Thus interest rate seasonals did not reappear to any significant extent.

2.1.2 The meaning of interest rate smoothing

The notion that it was the founding of the Fed rather than other factors which resulted in the disappearance of interest rate seasonality has also led to the view that interest rate management by central banks essentially amounts to the "smoothing" of seasonal variations in short-term interest rates. Mankiw and Miron (1991), for example, state that "since the Federal Reserve began operations in 1914, its policy has been to smooth nominal interest rates over the year", which it achieves "by expanding the money supply during those

times of the year when the demand for goods and services is high". Consequently, they refer to the policy of the Fed as "seasonal monetary policy". This perception also appears elsewhere in the literature (see Walsh 1990, pp.18-24) and, whilst it is not incorrect, is a somewhat ambiguous description of the interest rate instrument as it was (and still is) employed by the Fed and by other central banks. What is more, it is potentially misleading, as the following two examples indicate.

First, the emphasis on interest rate policy as smoothing has led at least one prominent researcher to develop a theory, based on otherwise plausible assumptions, which "implies that the Fed is concerned with the predictability of interest rates rather than with their level" (Cukierman 1991, p.113). This is an unfortunate theoretical implication, since it is so obviously counterfactual: central banks and participants in financial markets are *extremely* concerned with the level of short-term nominal rates.

Secondly, it has led to the view that central banks were somehow behaving differently before 1914, as is exemplified by the following statement by Clark (1986, pp.89-90):

> "Central banks existed in Great Britain, France, Germany, and other countries in the nineteenth century. Jevons ... claimed to have identified a dangerous association between the seasonal peaks of interest rates and the outbreak of financial panics, and, as early as the 1860s, he called on the Bank of England to take actions to eliminate interest rate seasonals. If interest rate seasonals were dangerous and central banks possessed the power to eliminate them, it is curious that all central banks waited until after world war one to exercise these powers.

The problem with the definition of interest rate smoothing as the elimination of seasonal fluctuations *per se* is that it ignores the role of the short-term interest rate as the monetary *instrument*, whose level is chosen to attain some policy target. Interest rate smoothing, therefore, if it is to be used as a descriptive term about central bank policy, should be defined as follows: it is the undertaking, by central banks, of transactions which keep (certain) market interest rates at their *desired* level. This definition – which will be the assumed definition from this point on – is preferred because it stresses that it

is not just seasonal fluctuations in market interest rates that concern central banks, but *all* fluctuations which would undermine the achievement of its extant policy target(s).

What is more, this definition does not preclude the possibility that the interest rate instrument may itself be employed in a seasonal fashion, and thus easily resolves Clark's puzzle about the behaviour of central banks before 1914. Far from wanting to eliminate interest rate seasonals, central banks *relied* upon them to counteract outflows of gold. That is, they raised their official rates when their reserves were threatened, since their policy target was a specified (minimum) level of reserves. This does *not* mean that central banks were not smoothing interest rates, or that they were not concerned with eliminating financial panics. They were, since without their discount facilities interest rates would have risen *further* and panics would have been more *frequent*, as is evidenced by the US before 1914. After 1914 *and the suspension of the gold standard*, central banks – a group which now included the Fed – no longer had to apply their instrument in a seasonal fashion: as their intermediate and final policy targets changed, so too did the manner in which they employed their interest rate instrument.

To sum up, interest rates have nearly always been the monetary instrument of central banks. This was the central point of Hawtrey's (1938) book *A Century of Bank Rate*:

> When the gold standard is suspended or abandoned, that does not mean that monetary contraction or expansion need no longer be applied. The alternative dangers of inflation and deflation remain, and if the gold reserve is not employed as an indicator, other symptoms must be observed, such as the state of employment, the level of profits, or a rise or fall of the price level. And when the threat of inflation or deflation is detected, it is in Bank rate that the essential corrective is found" (p.260).

Making the same point in modern terminology, one might say that central banks – *qua* central banks – have always engaged in interest rate smoothing, as it has been (re)defined above. Clearly central banks have improved their interest rate smoothing over time, in the sense that they are now much better at preventing undesirable fluctuations in short-term interest rates than they

were a hundred years ago. Even today, though, interest rate smoothing is not perfect. The aim of *Part II* of this thesis is to examine in greater detail the techniques, or operating procedures, which the Bank of England employs to smooth interest rates.

The essays therein will be concerned with the operational aspects of these techniques, rather than with how policy is derived. The implicit assumption throughout will be that the Bank of England has a policy target, and that it attempts to hit this target by appropriate choices of the short-term interest rate over time. In light of the discussion so far, this is hardly an unrealistic assumption:

> The original adoption of Bank rate as an instrument of monetary regulation was founded on theoretical reasoning. But the practical application of how it evolved in the nineteenth century was empirical. Starting from the postulate that a rise in the rate of discount must be a deterrent on the creation of credit, the Bank raised the rate whenever *its reserves seemed insufficient*, and went on raising it step by step until the desired effect was produced. When *the reserve began to increase*, the Bank lowered the rate step by step... " (Hawtrey 1946 p.274).

Most economists would agree that replacing the emphasised phrases in this description with the phrases "inflation seemed too high" and "inflation began to decrease", respectively, would not be an inaccurate description of the Bank's current mode of operation.

There are, however, those who have questioned the *necessity* for the Bank of England – indeed, all central banks – to operate in this way. It has frequently been proposed, for example, that central banks *should* impose quantitative limits on the amount of the currency and bank reserves (otherwise referred to as the money 'base') which they are prepared to supply, and let money market interest rates vary according to market forces. Poole (1970), in a widely cited paper, has characterised this issue as one of "optimal instrument choice", implying that a choice really can be made about the central banks' instrument. Central bankers, on the other hand, have usually denied that such a choice exists *in practice*. In addition, academics have been far more successful at making a case for some assumed (and often undefined)

form of money base control than they have at deriving operationally feasible proposals for such a form of control. Or analysing how these would operate in practice. Thus, the next section briefly takes up this issue.

2.2 The Money Base as an Instrument

2.2.1 Setting quantitative limits: money base control

Poole (1970) has defined an instrument of monetary policy as "a policy-controlled variable which can be set exactly for all practical purposes" and adds that "it is, for example, a straightforward matter to use the approach of this paper to treat the monetary base as an instrument" (p.198). This reflects the attitude of many academics who acknowledge that central banks could not use as their instrument any measure of money which included bank deposits, since this would not be directly controllable, but who nonetheless believe that the monetary base could, and indeed should, constitute the monetary instrument.

The single most influential proponent of this view is Milton Friedman who, in a series of lectures in 1959 entitled *A Program for Monetary Stability*, called for US monetary policy to be implemented according to a system resembling money base control. It is crucial to bear in mind that Friedman's proposal was motivated by the notion that a monetary aggregate, however defined, could be rendered subject to a *rule*, which could then be used to inhibit the discretion of the monetary authorities:

> "There is little to be said in theory for the rule that the money supply should grow at a constant rate. The case for it is entirely that it would work in practice. There are persuasive theoretical grounds for desiring to vary the rate of growth to offset other factors. The difficulty is that, in practice, we do not know when to do so and by how much. In practice, therefore, deviations from the simple rule have been destabilizing rather than the reverse" (Friedman 1959, p.98).

This has remained the primary attraction of money base control, although its proponents have been less forthcoming both with specific rules and with

operational suggestions of how these rules would actually be implemented. It is noteworthy that Friedman himself argued for the introduction of 100% reserve banking (although reserves would pay interest) as an important adjunct to his proposal, which probably decreased its overall acceptance significantly.

McCallum (1987, 1988 and 1989) thus seems to be a welcome exception. Not only does he suggest that the monetary base should be the instrument of policy, but also that this instrument should be set according to a *rule* which "specifies quarterly settings for the variable Δb_t, where $b_t = log B_t$ and B_t denotes the average value during quarter t of the monetary base" (1989, p.340). After stating a concrete example of such a rule, McCallum then seeks "to determine how nominal GNP would behave in an actual economy ... if its monetary authority were to conduct policy by setting the monetary base in accordance with the policy rule [just] proposed" (p.343). This is done by estimating a simple model of GNP determination, in which current GNP depends only on its past values and on the growth of the monetary base, and then using this estimated relationship to simulate the path which nominal GNP *would* have followed had the money base rule been applied. The conclusion is that "the rule would have worked quite well according to [the] model; the simulated models stay quite close to the target path [for nominal GNP]" (p.345). Having made what seems a plausible case for the adoption of a money base rule, McCallum then asks why such a rule has not been adopted in practice. The answer? "... [T]he feeling of self-importance of hundreds of [central bank] employees might be diminished substantially if policy were conducted in accordance with an automatic rule" (p.349).

It is useful, then, to evaluate McCallum's specific proposal, not only to see whether it has the desirable properties he claims, but also for highlighting the problems associated with money base control more generally. Assume, therefore, that a central bank were to implement McCallum's proposal. It is obvious that, by making the appropriate open-market operations, the central bank could always make the money base grow at exactly the rate required by the rule, say 1% per quarter. This part of the proposal is therefore trivial. The real problem, which is hardly ever addressed by the proponents of money base control, concerns the behaviour of the banking system when open-market

operations aim to make the money base, which is the sum of currency and bank reserves, grow at a particular rate. This introduces a potential conflict between the *actual* growth in bank reserves, which is affected by the behaviour of currency holders, and the *target* level of reserves, which banks need in order to satisfy their reserve requirements.

All banking systems are subject to a reserve requirement, which demands that banks, by law, hold at least a specified average amount of reserves over a specified period known as a maintenance period. Of course, it must be the case that the amount of reserves which would satisfy the requirement is determinate, otherwise it becomes impossible to establish whether banks have in fact satisfied them[1]. Note that the requirement could easily be zero, as it is in the UK[2]. More normally the reserve requirement is positive, and derived as a proportion of banks' average liabilities over a specified period known as a computation period. There are clearly many different types of reserve requirements which may be imposed and these vary considerably in practice. For illustrative purposes, however, assume that the central bank (which will implement the money base rule above) specifies that banks must hold reserves whose average over any quarter is some positive fraction of their average deposit liabilities over the previous quarter. This means that banks must manage their reserve holdings in the present quarter to meet a

[1]This means that the maintenance period cannot be infinite, and also that reserve carry-over provisions cannot be so loose as to render the maintenance period effectively infinite. See the discussion on reserve requirements in chapter 3.

[2]It may seem puzzling how a bank's reserves could ever become negative, and thus why a reserve requirement of zero should be called a requirement at all. Payment arrangements in most banking systems, however, do not impose a reserve-in-advance constraint on banks, with the result that banks can, and do, make payments without knowing their exact reserve levels over time. This reserve level is only revealed when payments are settled, usually at the end of the day. Therefore, if a bank has borrowed less than it has lent during the day, it may well be the case that it ends the day with negative reserves, i.e. is receiving an implicit loan from the central bank. In order to satisfy a zero reserve requirement, the bank may then have to borrow reserves from another bank which has a surplus (this is what happens in the UK), or the implicit loan by the central bank may be converted into an explicit loan, in which case the bank may be asked to pledge collateral to the central bank as security for the loan.

predetermined requirement[3].

It is easily demonstrated that the ability of banks to meet their reserve requirements under the money base rule is contingent upon the behaviour of currency holders. The average value of the money base per quarter is defined as the sum of the average value of the currency stock and the average value of bank reserves per quarter, i.e.

$$B_t = C_t + R_t .$$

Satisfaction of a money base rule implies that the central bank uses open-market operations to make $B_t = \bar{B}$. In order to satisfy their reserve requirements, banks must ensure that their actual reserves meet some target level, $R_t = \bar{R}$. Thus banks' reserves relative to their target is given by

$$R_t - \bar{R} = \bar{B} - \bar{R} - C_t ,$$

which clearly depends upon the preferences of currency holders. If, for example, currency in circulation were to increase by more than $\bar{B} - \bar{R}$ during the quarter, banks would become unable to meet their reserve targets, i.e. $R_t < \bar{R}$. Anticipating this violation of their requirement, banks will attempt to raise their reserve holdings by, in the first instance, borrowing on the money market. Of course, additional reserves cannot be obtained in this manner for banks in aggregate, and some banks will always remain below their reserve target. For banks to achieve their reserve targets, interest rates must be driven so high that one of two things will happen. Either currency holders are induced, directly, to deposit currency with the banking system, or nominal activity (or the price level) falls sufficiently to lower the transactions demand for currency, thus inducing this outcome indirectly. These,

[3]The argument which follows could just as easily be made by assuming that banks must hold reserves which average over any quarter at least some positive fraction of their average deposit liabilities over the *forthcoming* quarter. This arrangement would simply mean that banks would have to manage their deposit liabilities in the present quarter to meet a predetermined reserve holding. Since it is easier for banks to adjust their reserves than their deposits, which are largely demand-determined, this form of reserve requirement has never been imposed (this has not prevented academics from proposing them, however; see Laurent 1979).

then, are the implicit forces through which money base control will influence the economy. Before discussing the more plausible scenario, which is that the central bank's concern with the stability of the banking system will not allow interest rates to attain the level which is necessary for these forces to work, a less likely scenario is first examined. Namely, that these forces do actually work within the relevant time frame.

Although it is extremely unlikely, it may be that interest rates do actually set off the forces just described, and economic agents begin to reduce their currency holdings. Now the opposite problem may occur. Banks may find that they are accumulating reserves too fast, which will cause them to exceed their reserve requirements. Given that reserves are unremunerated, banks will seek to lend these, in the first instance, on the money market. As before, however, currency holdings need to adjust via the above forces. Even if these forces worked quickly enough, therefore, which is *a priori* doubtful, it would leave the problem of potentially dramatic swings in nominal interest rates[4].

Such interest rate behaviour, or indeed *any* interest rate behaviour, is totally ignored in the simple model adopted by McCallum to "test" the performance of money base control. As noted above, this model takes current nominal GNP to be a function of lagged nominal GNP and current growth in the money base on the grounds that both of these variables are statistically significant in an OLS regression estimated on actual US data over the period 1954-85. It hardly needs pointing out that the coefficients obtained in this regression will reflect the fact that the Federal Reserve did *not* implement money base control over this period. Instead, it was using short-term nominal interest rates as its instrument, and was accommodating any and all changes in the demand for the money base. McCallum's procedure thus falls squarely into the category of policy evaluation procedures criticised by Lucas (1976), since the coefficients in the estimated relationship are clearly not invariant under the change to a money base policy. Using these coefficients to simulate how nominal GNP would have behaved under a different policy does not therefore constitute a valid test of the performance of such

[4]This has been the primary criticism of money base control; see, for example, Foot, Goodhart and Hotson (1980) or Goodhart (1989a).

an alternative policy.

Instead, to test this performance, proponents of money base control will have to use models of the economy which are explicit, firstly, about the channels whereby the money base will influence prices (or nominal GNP) and, secondly, about the constraints placed on money base control by the presence of reserve requirements. These constraints were shown to imply potentially large fluctuations in short-term interest rates, which poses an even more compelling argument against money base control.

2.2.2 Why the money base has not been controllable in practice

Solvency constraints on the banking system clearly impose limits on the levels to which interest rates can rise without the stability of the banking system being drawn into question. When rates start to approach this level, therefore, as they might under the example presented above, the central bank will be faced with a dilemma: should it violate its money base rule by making reserves available? Doing so will avert a potential banking crisis in the current quarter, but it will also increase the expectation on the part of banks that it will again violate its money base rule in future quarters. This reduces the resolve of banks to hit their reserve targets in the future, however, and so too the extent to which they will bid up the interest rate to obtain reserves. It also lowers the chances that the equilibrating forces, inducing agents to redeposit currency, will actually work.

The problem with a money base rule, then, is the following. Either the central bank never reneges on the rule, in which case the rule is credible but causes interest rates to fluctuate considerably: this may or may not be consistent with stability in prices or nominal GNP, and will almost certainly cause occasional banking crises. Or, the central bank sometimes reneges on the rule, in which case the rule begins to lose credibility and causes interest rates to fluctuate less over time. Even Poole (1991), some twenty years after his paper on optimal instrument choice, now takes this view:

> "Central banks have not been successful ... in defining a credible policy based on short-run control of a monetary aggregate and possessed of an emergency escape clause that permits a quick response to a liquidity crisis. A liquidity crisis is inherently cumulative and is best nipped in the bud. How can a monetary quantity rule be defined that provides extra liquidity when necessary? " (p.39)

This was precisely the difficulty which faced the Bank of England in the eighteenth and nineteenth centuries, and which has faced all other central banks subsequently. There is a striking parallel between proposals for money base control which, due to the presence of reserve requirements, are seen as a means of constraining excessive deposit (money) creation by banks, and the minimum specie ratios imposed upon central banks under the gold standard, which were seen as providing a limit on the fiduciary issue of these banks. Whilst periodic wars, and the attendant need for governments to increase their expenditure, obviously caused this fiduciary limit to be exceeded from time to time, it was the development of banking, and in particular the development of money markets, which was ultimately responsible for central banks' inability to use the quantity rather than price of reserves as their monetary instrument.

Originally, banks resembled cash warehouses, holding most of the specie deposited with them in the form of vault cash. As the opportunity costs of these reserve holdings became apparent, and the nature of their deposit and withdrawal flows became understood, banks gradually increased their interest-earning assets relative to their non-interest bearing currency reserves. Moreover, as economic agents found it increasingly convenient to use bank deposits as a means of saving and making payments, so banks began to find that an increasing part of their reserve needs were for purposes of settlement with other banks. This led to several innovations, whose impact was to reduce the reserves needed by banks to meet their obligations. Instead of holding physical stocks of reserves, banks began to use *reserve accounts* at a common, or central, bank, since this decreased the costs and risks associated with actual transfers of physical reserves. The fact that payments by one bank were often payments to another bank not only made cancellation (or netting) possible, but also created the incentive to establish *interbank funds*

markets since a (temporary) reserve deficit in one bank would correspond with a (temporary) reserve surplus in another bank. The ability to borrow reserves in the event of a reserve drain, therefore, meant that banks could reduce their reserve inventories and yet remain confident of meeting their obligations (see chapter 3).

Whilst the development of interbank markets may have been perceived by banks individually as a form of insurance against the random daily fluctuations in their reserves, this was not true for banks in aggregate. Indeed, the opposite was true: by causing banks to reduce their reliance upon reserves as a buffer against random cash fluctuations, interbank markets caused banking systems to become less liquid overall, i.e. less able to cope with fluctuations in the *aggregate* level of reserves. As noted by Bhattacharya and Gale (1988), interbank markets are associated with a "free-rider" problem:

> "individual intermediaries will rely on the ex-post market to provide them with liquidity, and will underinvest in the liquid asset [i.e. reserves]. In the aggregate, however, the liquidity of the market is limited by the investments of individual intermediaries in the liquid asset. For this reason we should not expect an interbank market to perform very well" (p. 74).

Banking history amply demonstrates this: liquidity crises were a regular feature in the development of banking systems during the eighteenth and nineteenth century, periods in which interbank markets had made it increasingly attractive for banks to invest their surplus reserves with other (deficit) banks. That such markets were associated with a free-rider problem – i.e. that banks did not take into account the aggregate availability of reserves when making their own reserve decisions – was exemplified by the fact that even highly predictable, seasonal fluctuations in the aggregate demand for reserves were a frequent source of distress.

One measure of this distress was the short-term interest rate. Faced with the prospects of being unable to satisfy their creditors, banks obviously became willing to pay extremely high interest rates in order to obtain reserves and avert this outcome. Constraints on the aggregate supply of reserves, however, meant that these reserves would not be forthcoming, *no matter how*

much some banks were prepared to pay for them, i.e. no matter how high the interest rate. The last thing currency holders would do in a crisis was part with their currency, with the result that ordinarily liquid assets became, suddenly, extraordinarily illiquid.

Central banks (or, more correctly, those institutions which would become central banks) were well aware that by extending loans (rediscounting bills and other paper) during a crisis they would eventually threaten their own reserve positions and so too the convertibility of their obligations. Sometimes, therefore, they chose not to lend, or to curtail their lending after some critical point, which invariably meant that banks failed. Consequently, trade, which was heavily dependent on the credit provided by bills of exchange, was disrupted. Consider this description, by Baring (1797), of a banking crisis in England in 1793:

> "At first the Bank accommodated themselves to the circumstances and furnished large supplies [of specie]; but unfortunately the Directors caught the panic; their nerves could not support the daily and constant demand for guineas; and for the purpose of checking that demand, they curtailed their discounts to a point never before experienced, and which placed every part of the commerce of the country in a considerable degree of danger" (p.21).

This reluctance on the part of the Bank was repeated again in 1825, with similar consequences:

> "By every means [the Bank] tried to restrict its advances. The reserve being very small, it endeavoured to protect that reserve by lending as little as possible. The result was a period of frantic and almost inconceivable violence .... The country was ... within twenty four hours of barter" (Bagehot 1876, p.190).

On some other occasions, however, the Bank was prepared to lend, and it became apparent that the crisis sometimes subsided without any significant loans actually having to be made. In other words, the confidence generated by the Bank's willingness to lend was itself sufficient to prevent a crisis from degenerating into a widespread panic. Whilst this feature of currency "runs" is well understood now, it took some time before the Bank began to understand it during the nineteenth century. Walter Bagehot, writing during

the 1860's and 1870s, was one of the first to articulate clearly the nature of such panics and how the Bank should deal with them. His prescription has become well known: the Bank should raise its discount rate, but should discount freely at this rate. The latter, however, was the key. Subsequent commentators have often misinterpreted Bagehot's advice as stating that the Bank should lend to illiquid but not to insolvent institutions (e.g. Summers 1991). This is not the case, and there is nothing in *Lombard Street* dealing with the "bailout" of individual banks, solvent or insolvent. Instead, Bagehot was concerned that the Bank should not limit its advances to any comers, and thus should not place too many restrictions over the type of security it would purchase in a crisis:

> The *amount* of the advance is the main consideration for the Bank of England, and not the nature of the security on which the advance is made, always assuming the security to be good. An idea prevails ... at the Bank of England that they ought not to advance during a panic on any kind of security on which they do not commonly advance. But if bankers for the most part do advance on such security in common times, and if that security is indisputably good, the ordinary practice of the Bank is immaterial. In ordinary times the Bank is only one of many lenders, whereas in a panic it is the only lender (Bagehot 1876, p.196).

This advice – not to place limits on the provision of reserves – was taken by the Bank of England, and in turn by all other central banks. Indeed, it is an integral part of how these banks came to be recognised as central banks in the first place. Their willingness to provide reserves (or currency), especially to avert a panic, is widely recognised as one of the fundamental functions of a central bank. Yet this is precisely what (some) proponents of money base control suggest central banks should refrain from doing. They argue that, by withholding reserves, central banks will encourage banks to hold greater reserves as insurance against liquidity problems. This change in behaviour is in turn supposed to reduce the possibility that fluctuations in reserves threaten banks' liquidity and cause money market interest rates to vary so dramatically. Thus, for example, Artis and Lewis (1981) argue,

> "Once banks are forced to make up reserve shortages by borrowing interbank at a 'penalty cost' or by selling securities at a loss, they are likely to

> exercise much greater care in future when granting facilities and open-credit lines. Unused facilities are a valuable source of liquidity to customers, and banks might, in different circumstances, be expected to vary the 'price' for this service. There would also be an incentive for banks to refrain from lending and build up reserves when reserve shortages are anticipated" (p.124).

This completely ignores the fact that variations in demand deposits, over which banks have virtually no control, are also a major source of fluctuation in banks' reserves. It also flies in the face of historical experience, because it ignores the free-rider problem raised above. In the presence of a money market, banks face a constant incentive to lower their own reserve holdings on the assumption that they can obtain reserves, when necessary, in the money market. This problem does not diminish over time: on the contrary, higher precautionary reserve holdings, by reducing the incidence of liquidity crises, will in turn validate the reductions in reserves that tend to raise the probability of another liquidity crisis. The historical record shows that periodic liquidity crises did persist, even when central banks could credibly withhold reserves from the banking system. Early central banks tended to act in their own interest during a banking crisis, which meant that they could credibly uphold their minimum reserve ratios even when interest rates rose to astronomical levels and banks failed. Indeed, their credibility could hardly have been higher, and yet this was *not* sufficient to induce banks to hold larger reserves to avoid these periodic crises.

The resolve of central banks not to supply reserves during a crisis was undoubtedly made weaker over time by the increasing economic cost of bank failures. As economic agents became more reliant upon the banking system for making payments, and held a greater share of their monetary wealth with the banking system, bank failures became more disruptive and less tolerable. Furthermore, banking crises became better understood as being, in the main, temporary phenomena, deriving from a lack of confidence on the part of an imperfectly informed public, rather than from any deep problems of widespread bank insolvency due to imprudent lending. Thus the notion of lending freely became more palatable to central banks. Especially since they could always counter any longer-term problems associated with this surge in

bank reserves by raising the rate at which these were provided. As Hawtrey (1938) notes:

> ".... Bagehot's prescription for a panic is really just a restatement of the policy of relying on a high Bank rate in place of a refusal to lend or a rationing of discounts. It takes for granted that the high Bank rate will have a sufficiently deterrent effect " (pp.226-7).

And herein lies the rub. On occasion, central banks will be called upon to discharge their obligation as lenders of last resort, and will be expected to provide reserves in excess of whatever money base rule they may have adopted. At this point, however, their instrument ceases to be the money base, and they are forced instead to choose the rate at which they will provide reserves. At least some of the time, therefore, their instrument must be an interest rate. Unfortunately for the proponents of money base control, as could be seen from McCallum's example above, this will be rather frequent, given the difficulties of banks actually satisfying their reserve requirements under such as system.

Even Milton Friedman has recognised this difficulty:

> "If rediscounting were eliminated, one minor function now performed by the discount rate would need to be provided for in some other way .... [S]ome discrepancies between required and actual reserves are unavoidable, yet some penalty must be imposed on such discrepancies to enforce the reserve requirements .... The simplest alternative would be a fixed rate of 'fine'. To avoid discrepancies becoming an indirect form of discounting, the 'fine' should be large enough to make it well above likely market rates of interest. The fine would then become the equivalent of a truly 'penalty' discount rate ... except that no collateral, or eligibility requirements, or the like would be involved" (1959, p.44-5).

Friedman unfortunately neglects to take this "minor function" to its logical conclusion. If the banking system were indeed below its required reserve target, money market rates would be bid up to the supposed penalty rate, which would then cease to be a penalty. Indeed, the rate at which the central bank made reserves available would act only as penalty if banks did not anticipate this shortage of aggregate reserves (see section 2 in the following

chapter). In practice, of course, banks would have expectations about the likely evolution of aggregate reserves over the maintenance period, and would choose their reserve strategies accordingly. If a reserve shortfall was expected, the cost of reserves (i.e. short-term interest rates) would be bid up to the central bank's lending rate. If a reserve surplus was expected, the cost of reserves would be driven towards zero.

Such a system, unfortunately, would be a far cry from one in which the money base was the monetary instrument as defined above. Every period in which there occurred a shortfall in actual reserves relative to required reserves, the central bank would be forced to provide this difference. Its choice variable, over which it would have complete discretion, would be the rate at which it supplied these reserves. Far from providing a system of monetary ease or restraint via choices of the money base, the central bank would, instead, *have to resort to appropriate choices of this interest rate if it wished to influence macroeconomic variables in the direction dictated by its policy goals* . This system, however, would come unerringly close to central bank operating procedures as they looked before 1914 (i.e. interest rate smoothing in a rather primitive form). As such, it could easily be improved by using open-market operations to reduce unwanted variations in money market interest rates. In sum, therefore, short-term interest rates would remain the monetary instrument.

Chapter 3

The Determination of Money Market Interest Rates

The previous two chapters have consistently emphasised the connection between the reserves of the banking system and money market interest rates. They have argued that it is the potential for this connection to disrupt the banking system and, in turn, economic activity, which provides a plausible account of why central banks operate as they do. But what, exactly, are the connections between banks, the money market and the central bank? More specifically, how can the supply and demand for bank reserves – and hence the central bank as the supplier of reserves – play such an important role in determining money market interest rates, when it would seem that the supply and demand for credit is a far more obvious candidate[1]?

This chapter draws out the relationship between banks, reserves and money market interest rates in more detail, showing how this relationship permits central banks to influence, fairly precisely, the level of money market interest rates. To fix ideas, the discussion starts by outlining, in section 1, the functional relationship between money markets and banks, drawing particularly on the sterling money market as an example. Thereafter, in section

[1]Dow and Saville (1988), for instance, stress the importance of the latter in determining the level of short-term interest rates.

2, a more analytical approach is taken to develop these ideas. In particular, a model of bank reserve management is presented, which then serves as the basis for an aggregate model of bank money market borrowing wherein it is shown how the central bank can determine the money market interest rate. Since the model is somewhat stylised and incorporates actual institutional features only in a minimal fashion, a final section addresses the relevance of two such features – reserve requirements and the discount window – for the determination of interest rates in practice.

3.1 Banks and the money market

3.1.1 Credit allocation

By most accounts, bills of exchange were the first money market instruments (see chapter 6). Although bills became legal in England in 1697, three years after the formation of the Bank of England, they had been extensively used before this to finance trade in England and throughout Western Europe. The primary function of a bill was to enable a seller (or exporter) to obtain cash as soon as possible after the dispatch of goods. At the same it enabled the buyer (or importer), who would typically only receive the goods after some time, to delay payment. By creating a negotiable instrument, i.e. a bill which was payable to the bearer, a third party could finance the trade deal, becoming a lender by buying the bill and holding it to maturity[2].

Negotiability was an important characteristic, since it meant that the initial lender (i.e the buyer of the bill) had the flexibility of selling the bill before maturity, thus transferring the loan to another creditor, who in turn also had this flexibility. The creation of a negotiable instrument was thus a convenient mechanism for resolving the different maturity preferences of borrowers and lenders, since its tradability on a secondary market implied

[2]The lender would receive a return on the bill by buying it at a price below its face value, i.e. at a discount.

that lenders could liquidate their loans without disrupting borrowers[3].

The fact that bills were negotiable, however, did not make them readily tradable. Information about the credit-worthiness of borrowers was clearly important to the buyers of bills, since it would determine whether these bills would be honoured at maturity. In order to enhance the tradability of bills, it became common practice for local financiers or banks, whose credit-worthiness was more widely known, to "accept" these bills, i.e. agree to honour them at maturity[4]. Since bank bills were rated more highly than ordinary trade bills, they could be traded more easily on a secondary market.

Bills of exchange (acceptances) provide an early example of how the presence of banks assisted and reinforced the development and liquidity of money market instruments. This relationship is still very relevant today, as has been argued recently by Garber and Weisbrod (1992): bank credit lines underpin the liquidity of many money market securities, a primary example being commercial paper. The relationship, however, also runs in the other direction, since the presence of money market instruments in turn assisted and reinforced the development of a banks.

Early banks were much more localised than their modern counterparts, and evolved only slowly into an integrated banking "system". This evolution was often inhibited by banking regulations, such as the restrictions on joint stock banking in the UK and on interstate banking in the US, which effectively prevented banks from developing branch networks. Consequently, banks in one area (e.g. country banks) will have found they had a surplus of investible funds, whilst those in another area (e.g. city banks) found they had a shortage of funds. The development of negotiable instruments proved to be an important means whereby funds could be redistributed from surplus to deficit banks. For example, in the UK, banks with lending opportunities could accept bills and then sell them into the London bill market, where they were purchased by banks which had funds to invest. Later (after 1825),

[3]This benefit of secondary markets has been formally examined recently by Bencivenga, Smith and Starr (1992).

[4]Hence they also became known as bankers acceptances, which is what bills are called in the US. Banks naturally charged an acceptance fee for the credit risk which they assumed.

surplus funds were primarily lent to discount houses (bill brokers) on a call basis, in which case it was the houses that purchased bills (King 1936). In the US the pattern was similar: banks could lend to borrowers in other states by purchasing commercial paper. Alternatively, they could lend funds on a call basis to the so-called "money center" banks.

The development of money markets, then, was an integral part of the process whereby individual banks became linked into a wider banking *system* or *network* which could allocate credit on a national level. This establishment of a network of banks, linked through a money market, was also crucial to another important function undertaken by banks, namely the provision of payment services.

3.1.2 Payments

The widespread use of physical currency in the exchange process came about because it reduced the resource costs of matching individuals with their preferred consumption bundles (Kiyotaki and Wright 1989). Consequently, currency became the most liquid asset in a monetary economy, with the liquidity of all other assets being defined in terms of the time taken (or cost incurred) to exchange them for currency (Lippman and McCall 1986). The use of currency, however, was not the final step in the development of payment media. The most prevalent means of payment in most (advanced) monetary economies today is not physical currency but bank deposits[5]. This is because organising payments through bank deposits could reduce the costs of monetised exchange even further.

This cost reduction was the result of two innovations: the centralisation of bank reserves and the development of money markets. Transfers of physical currency involved transportation costs and risks of theft, which could be eliminated by utilising a centralised system of book entries across a set of reserve accounts at a clearinghouse or central bank. Furthermore, the use of such a set of reserve accounts allowed mutual obligations to be offset (netted),

[5]As the next section argues, this does *not* mean that currency is eliminated as a settlement asset.

reducing the amount of reserves needed to settle a given volume of payments (Goodfriend 1990). This reduction was desirable because reserves were non-interest bearing, which meant that reserve inventories held for settlement purposes incurred an opportunity cost. This cost will have become more apparent to banks once a market had emerged on which they could borrow and lend reserves, such as the early bill market in the UK. This market also provided an avenue for reducing these costs even further than was permitted by netting.

In the absence of either a sustained increase or decrease in deposits, a bank's reserves will fluctuate randomly over time, being either temporarily high or temporarily low due to the payments made by customers. By purchasing negotiable instruments, such as bank bills, temporary reserve surpluses could be invested in interest bearing assets without comprising a bank's liquidity. This was because bills had a short maturity and could be sold before maturity if reserves were needed even sooner. What is more, reserve fluctuations were not independent across banks. The operation of a payment network meant that reserve outflows from some banks would inevitably result in reserve inflows at other banks. This made it more likely that a bank could easily sell bills when its reserves became temporarily low, since this would usually mean that surplus reserves would be accumulating at some other banks, who would be wanting to invest these in bills. The provision of payment services, therefore, could be achieved more cheaply when banks had access to a money market as interbank settlement obligations resulting from payment flows could be offset by the appropriate money market transactions, obviating the need for costly reserve inventories.

In addition to permitting a wider allocation of credit through the banking system, early money markets facilitated the redistribution of reserves amongst banks, and became vital to their provision of payment services. Indeed, this function of money markets, more than any other, is probably what has spurred their development over time, evidenced by the fact that banks have always been the most active participants within money markets. It is the redistribution of reserves, moreover, which plays a key role in the determination of money market interest rates.

3.1.3 Movements in bank reserves

The widespread use of bank deposits as a means of payment did *not* eliminate the role of currency as the ultimate settlement asset. Rather, it meant that a much smaller amount of currency supported a larger volume of payments. The impetus for this development was that currency did not bear interest and hence involved an opportunity cost. Banks, using a set of reserve accounts, and having access to a money market, significantly reduced this cost. Nevertheless, banks still use currency to settle their obligations, albeit in the form of reserves (book entries) rather than physical currency (notes and coin). Except for their location, reserves are the same as vault cash: they comprise currency held at the central bank rather than in the bank's own vaults[6]. Reserves are held because of the associated reduction in the costs of transferring currency to other banks for settlement purposes.

Note that the common practice of counting vault cash as part of bank reserves is not pursued here, since this becomes confusing in the analysis which follows. Throughout this and subsequent chapters, therefore, the term "reserves" is distinct from, and does not include, banks' vault cash.

The demand for reserves

Whilst it is obvious that banks must always hold enough vault cash to satisfy the expected demands of their customers for currency, it is less clear what level of reserves they need to hold in order to satisfy their interbank obligations. This will depend, in general, upon three factors: the settlement arrangements between banks, the extent of development of the money market, and the level of reserve requirements.

Historically, settlement arrangements between banks have meant that payments are settled at discrete intervals (e.g. weekly, daily), with the time between successive settlements gradually becoming shorter as clearing meth-

[6]This statement is strictly true only for fiat currency. During the gold standard, for example, the ability of banks to obtain *gold* (rather than central bank notes) in exchange for their reserves (book entries) was contingent upon the central bank actually upholding its convertibility pledge. Once central bank notes became fiat currency, however, reserves could *always* be exchanged at par for physical currency.

ods became more streamlined. Discrete settlement meant that a number of payment instructions (e.g. cheques) were accumulated before they were netted by a clearinghouse and the net obligations then settled via the transfer of reserves. This sort of arrangement, known as net settlement, is still widespread today, and implies that payment instructions are made (and received) *before* they are settled, i.e. before reserves are actually transferred[7]. Whilst this settlement lag exposes banks to risks[8] – since they make outgoing payments on the basis of *assumed* incoming payments – it also means that banks do not actually have to hold many reserves in order to meet their settlement obligations. Instead, they simply need to anticipate their net settlement obligation, and then try to offset this via the appropriate money market transactions.

Consider, for example, a bank whose customers have made a larger than average number of payments, and assume that the bank receives this information before settlement. Although the bank will be expecting to pay out reserves when these payments are settled (at the end of the day), it does not need a reserve inventory to meet this obligation. Instead, it can borrow reserves in the money market by, for instance, selling some bank bills. This increases the settlement obligation of other banks and, correspondingly, reduces the borrowing banks' *net* obligation. The borrowing bank can be reasonably sure to find banks which are willing to lend reserves since its own expected reserve need will be matched, elsewhere, by expected reserve surpluses. Since these surpluses are costly in terms of foregone interest, banks will have an incentive to reduce these by attempting to lend funds in the

[7]Prominent examples of such net settlement systems are SAGITTAIRE in France, EAF in Germany, CHAPS in the UK and CHIPS in the US, although these are now all electronic rather than paper based. For a comprehensive survey of these and other payment systems see BIS (1993).

[8]These risks have grown with payment traffic, which is now enormous. Currently, therefore, reforms are underway in many industrialised countries to reduce and manage these risks, including a move to alternative settlement arrangement known as gross, or continuous, settlement (see Borio and Van den Bergh 1993). The latter would mean that banks would need reserves *before* they could make payments, and would thus only really alter the *timing* of their money market borrowing for this purpose.

money market (for instance by buying bank bills).

In practice, therefore, the net demands by banks for reserves for settlement purposes will be virtually zero, given appropriate interbank settlement arrangements and well-functioning money markets. Although banks have no direct control over the reserve flows resulting from the activities of their customers, they can offset these flows indirectly through transacting in the money market. The fact that banks are often observed to hold fairly substantial reserves, even in the presence of well developed money markets, is thus largely due to reserve requirements.

The power to impose and vary the reserve requirements of banks is written into the statute of virtually every central bank. This power has been exercised by most of them either periodically or continuously since it was acquired (see Capie, Goodhart and Schnadt 1994). Reserve requirements are usually calculated as a fraction of some measure of banks' average deposit liabilities over a specified *computation* period. To satisfy the requirement, banks must ensure that their average reserve balances over a specified *maintenance* period equal or exceed the computed level. Since the failure to satisfy the requirement will attract a penalty[9], banks will ensure that they do not violate their requirements. Essentially, reserve requirements increase the overall demand for reserves to exactly the level which is implied by the requirement. Banks will not, however, seek to hold reserves in *excess* of this requirement: the demand for reserves by banks for other purposes, such as interbank settlement, remains virtually zero when there is a well developed money market.

This is clearly illustrated by table 3.1. Whilst total bank reserves differ widely between countries, this difference vanishes once reserve requirements are taken into account. Excess reserves – the level of reserves banks *choose* to hold – are extremely low everywhere, especially when measured in relation to the size of banks' call deposits. In the UK, for example, reserves represented less than *one thousandth* of banks' aggregate sight deposits to the UK private sector during the last six months of 1992[10].

[9]The exact nature of this penalty is discussed below, in section 3.2.5.

[10]See *Bank of England Quarterly Bulletin*, February 1993, Table 3, p.124.

The supply of reserves

The argument in the preceding discussion was based on the fact that interbank settlement obligations, when aggregated across all banks, cancel to zero. In other words, interbank transactions do not alter the aggregate supply of reserves. This is why a money market is sufficient for reducing the aggregate demand for reserves for interbank settlement purposes (virtually) to zero, causing the aggregate demand for reserves to be determined, for the most part, by reserve requirements. But what determines the aggregate supply of reserves? And, more importantly, what equilibrates the aggregate supply of reserves with the aggregate demand for reserves?

The aggregate supply of reserves at any point in time is simply the sum of the amounts of currency which are held by individual banks with the central bank. Consequently, there are two, and only two, sources of variation in the aggregate supply of reserves: (i) net currency withdrawals from the banking system and (ii) transactions by the central bank.

When a customer withdraws currency from (deposits currency with) a bank, its vault cash falls (rises). Banks pay close attention to these withdrawals and deposits, and ensure that their vault cash is maintained at fairly constant levels so that they do not fall short of cash or hold unnecessarily high stocks of cash. Variations in vault cash, therefore, will directly affect banks' reserves[11] as they will, in the first instance, replenish unduly low vault cash by reducing reserves (i.e. obtaining cash from the central bank), or reduce unduly high vault cash by increasing reserves (i.e. depositing cash with the central bank). Since there are no corresponding effects of these transactions on other banks' reserve accounts, the aggregate level (supply) of reserves is altered by these changes in banks' vault cash.

Similarly, transactions by the central bank also alter the aggregate supply of bank reserves. When the central bank acquires assets (or, for example, pays the salaries of its employees) the transactions are settled by a transfer of reserves to the payees' bank(s), thereby increasing aggregate reserves. Conversely, when the central bank sells an asset (or when a private sector

[11]Recall that vault cash is not assumed to be part of reserves here.

obligation held by the central bank matures) the transaction is settled by a transfer of reserves to the central bank, thereby reducing the aggregate stock of reserves. Often, these transactions by the central bank will reflect its role as the government's bank, since this usually means that some of the accounts of the government are held with the central bank. Thus tax payments to the government (or debt issues by the government) will reduce aggregate reserves, whilst any government expenditures will increase aggregate reserves[12].

Consider the effect of a decrease in the supply of bank reserves due (say) to increased withdrawals of currency into circulation. The banks experiencing this withdrawal will attempt to replenish their lower reserves by borrowing on the money market. Since the *aggregate* supply (stock) of reserves has fallen, however, at least one bank will have a lower than desired[13] level of reserves after settlement. Whilst this reserve shortfall may be passed to other banks on successive days through interbank borrowing and lending, it cannot be eliminated without a subsequent increase in the supply of reserves, either due to currency being redeposited with banks, or due to central bank transactions. Without this increase, at least one bank would violate its reserve requirement by the end of the maintenance period, thereby incurring a penalty. Over the remainder of the maintenance period, therefore, banks in aggregate would persistently be trying to borrow more reserves than they were willing to lend. This behaviour would tend to drive up money market interest rates, at least on those instruments which banks regularly used to adjust their reserves. Had the initial disturbance in the supply of reserves been an increase rather than a decrease, the opposite forces would pertain, with money market interest rates tending to fall.

For any given demand for reserves, therefore, changes in the aggregate

[12]Note that these factors are *not* perfectly predictable on a daily basis, making their net impact on the supply of reserves uncertain.

[13]In a modern banking system, and with ready access to a money market, this is likely to be determined by the level of reserve requirements. In the more distant past, before statutory reserve requirements, banks will still have had their own "target" level of reserves, based on their previous withdrawal and clearing needs. For a discussion of the UK case, where reserve requirements are zero, see Howard (1982).

supply of reserves will produce movements in money market interest rates, because banks will attempt to correct these changes by borrowing or lending in the money market. Crucially, however, this does not give rise to any significant forces which would act to reduce the excess demand for, or supply of, reserves. Offsetting changes in the demand for currency would be one such channel, but currency demand is driven primarily by transaction needs and is virtually unresponsive to temporary movements in short-term interest rates. Indeed, it is *because* of changes in net currency withdrawals for transactions purposes that money market rates change in the first place. This leaves offsetting central bank transactions as the only other means whereby aggregate reserve fluctuations, and hence movements in money market interest rates, could be damped.

3.1.4 The central bank

Previous chapters have indicated that seasonal movements in currency use often had dramatic effects on the banking system, particularly currency withdrawals. As banks attempted to borrow reserves, money market interest rates rose sharply, and ordinarily liquid assets suddenly became very illiquid. This caused considerable disruption to trade and other economic activity which was dependent upon short-term finance. Even worse, if seasonal currency demands persisted, these forces would *cumulate*, thus introducing the possibility of a widespread increase in currency demand, this time due to waning confidence in banks' ability to redeem their obligations. This would threaten to disrupt the entire banking system, and its by-product, the payment system.

Given this potential for adverse movements in bank reserves to disrupt the normal operation of money markets and the banking system, the revealed preference of all central banks over time was to offset such reserve movements. This was done passively at first and in a *unidirectional* manner. That is, variations in aggregate reserves were allowed to "reveal" themselves through variations in money market interest rates. When rates rose, indicating reserve stringency, central banks accommodated this reserve need by discounting eligible securities, preventing rates from rising further. When

rates fell because banks had excess reserves, central banks initially did nothing. This meant that money market rates often fell to extremely low levels for prolonged periods. Over time, central banks devised methods whereby they could reduce the reserves of the banking system, thus keeping money market rates at levels which they felt were more consistent with their policy targets. As noted in chapter 1, these included open-market operations, which acted on the aggregate supply of reserves, and reserve requirements, which acted on the aggregate demand for reserves.

In sum, central banks have developed increasingly sophisticated ways in which to "manage" the aggregate level of reserves, improving their ability to prevent any unwanted fluctuations in money market interest rates. They now devote considerable resources to predicting disturbances in aggregate reserves *before* these are allowed to reveal themselves through movements in money market rates, and have thus become proactive rather than reactive in their influence over rates. Consequently, by anticipating the impact of factors such as net currency withdrawals or tax payments to the government, and undertaking more frequent open-market operations, central banks now rely less on facilities such as a discount window. Given that these factors are not perfectly predictable, however, central banks have retained their discount windows as a safeguard against unexpected reserve stringency and the attendant upward movement in interest rates.

The next section now develops these ideas in an analytical framework, based upon a model of bank reserve management. This model is then aggregated to explain the determination of money market interest rates by the central bank.

3.2 An analytical approach

3.2.1 A model of bank reserve management

According to an early theorist of banking:

> "Probability is the foundation of banking. The solvency and the profits of the banker depend on the probability that he will not be called upon to meet at once more than a certain amount of his liabilities" (Edgeworth 1888, p.133).

This characterisation forms the basis of a classic choice problem for the bank manager. Holding reserves is costly, since they are non-interest bearing, and implies that it would be sub-optimal to ensure a bank's liquidity by holding assets in the form of reserves only. Conversely, holding some illiquid assets exposes the bank to the risk that it is unable to meet its obligations. A recent textbook on the management of financial institutions describes the problem thus:

> The management of a financial institution's [reserve] position is something of a knife-edge situation because holding too many [reserves] penalises a bank's earnings and, thus, its stockholders. A ... manager who holds excessive amounts of [reserves] is unlikely to survive for long. Similarly, a manger who excessively undershoots the reserve target faces enhanced risks of liquidity crises and regulatory intervention. Again, such a manger's tenure ... may be relatively short-lived " (Saunders, 1994, p.578).

Clearly, therefore, the "optimal" level of reserves which a bank will wish to hold will fall somewhere between 0% and 100% of the bank's liabilities. But how should this level be chosen?

The most common approach to this situation is to treat it as a static inventory problem in which the bank manager must choose the optimal inventory of reserves, denoted by r^*, under some assumed uncertainty about the demand for reserves[14]. To simplify matters, the uncertainty facing the

[14]The first presentation of such a model in this context was by Morrison (1966), who cites Arrow, Harris and Marshak (1951) and Dvoretsky, Kiefer and Wolfovitz (1952) as

bank is summarised by a single random variable, δ, which represents the net obligation of the bank at settlement (i.e. $\delta < 0$ represents an expected payment of reserves). This settlement obligation is assumed to follow a *known* distribution, described by the distribution function $F(\delta)$ (density $f(\delta)$). Thus the situation being modelled is not unrealistic, and can be thought of as the determination of the reserve demand of a bank which faces a (random) daily settlement obligation under a net settlement arrangement[15].

To capture the tradeoff facing the bank, assume that when it holds reserves r which exceed its obligations, i.e. $r > -\delta$, it incurs a proportional surplus cost c_s, representing, for example, the foregone interest on its surplus reserves. Alternatively, if the bank holds insufficient reserves to meet its obligations, i.e. $r < -\delta$, assume that a proportional deficiency cost c_d is incurred, representing the cost of an emergency loan from the other settlement members (or, in the extreme, bankruptcy). Under this characterisation, the optimal inventory of reserves per time period is chosen to minimise these ex-post costs, and it is easily recognised that this choice is subject to a trade-off. Choosing a higher inventory of reserves reduces the probability of incurring c_d, but increases the probability of incurring c_s.

Formally, the solution to the bank's choice problem described in this simple scenario is found by the bank by solving the following cost minimising program:

$$\min_r \quad C = c_s \int_{-r}^{+\infty} (r+\delta)\, f(\delta)\, d\delta \;+\; c_d \int_{-\infty}^{-r} -(r+\delta)\, f(\delta)\, d\delta\,,$$

where the first integral gives the expected value of a reserve surplus, and the second integral the expected value of a reserve deficiency, for each choice of

---

the inventory-theoretic forerunners of this approach. Subsequent papers using essentially the same static model, although with slightly different features, are Poole (1968) and Frost (1971). Others, such as Knobel (1977) and Chen and Mazumdar (1992) have extended the model into a dynamic context. Baltensperger (1980), Santomero (1984) and more recently Pulli (1992) provide useful discussion and bibliographies of these and related models.

[15]Adjustment costs are assumed to be zero, although they need not be zero (see Frost 1971).

reserves r. This has the first order condition

$$c_s \int_{-r}^{+\infty} f(\delta)\, d\delta = c_d \int_{-\infty}^{-r} f(\delta)\, d\delta ,$$

which suggests that the bank chooses the optimal inventory of reserves so as to equate the expected marginal cost of a reserve surplus with the expected marginal cost of a reserve deficiency. This condition can be rearranged to derive the optimal inventory of reserves, r^*, as

$$F(-r^*) = \frac{c_s}{c_s + c_d} . \tag{3.1}$$

From (3.1) it can be seen that the optimal inventory of reserves depends upon the level of c_d *relative* to the level of c_s and on the properties of the distribution of δ[16]. More specifically, the optimal reserve inventory is chosen to equate the probability of avoiding a reserve deficit (or surplus) with the relative cost of such an outcome. This accords with the intuition that, if a reserve deficit is costly relative to a reserve surplus, a bank would rationally choose to hold a higher reserve inventory to reduce the probability of the relatively more costly outcome.

3.2.2 An example

It is illustrative to examine some of the properties of this simple model of reserve management by considering the plausible case where δ has a *symmetric distribution.* Since the settlement obligations of a bank will, in reality, fall within a finite interval, the continuous uniform distribution provides a reasonable and tractable[17] example. Assume, then, that δ is distributed uniformly over the closed interval $[m - \epsilon, m + \epsilon]$. The mean and variance of δ are therefore $\mu_\delta = m$ and $\sigma_\delta = \frac{1}{3}\epsilon^2$ respectively, and $-\delta$ has the following

[16]Note that the reserve inventory enters this distribution function with a negative sign, which implies that (1) traces the *mirror image* of the distribution function.

[17]A truncated distribution would also be reasonable, but is generally less tractable. Pulli (1992), for example, uses a truncated normal distribution, which he later approximates using an exponential distribution.

distribution function (see Johnson and Kotz 1970):

$$F(-\delta) = \begin{cases} 1 & \text{if } \delta < (m-\epsilon) \\ (2\epsilon)^{-1}(m+\epsilon-\delta) & \text{if } (m-\epsilon) \leq \delta \leq (m+\epsilon) \\ 0 & \text{if } \delta > (m+\epsilon)\,. \end{cases}$$

Using this distribution function, (3) can then be used to derive the optimal reserve inventory of the representative bank

$$\begin{aligned} r^* &\leq (m-\epsilon) && \text{if } c_d = 0 \\ r^* &= (m+\epsilon) - 2\epsilon\left(\tfrac{c_s}{c_s+c_d}\right) && \text{if } c_d, c_s > 0 \\ r^* &\geq (m+\epsilon) && \text{if } c_s = 0\,, \end{aligned} \tag{3.2}$$

This linear (but discontinuous) function, which is depicted in figure 3.1, traces out the *optimal reserve inventory* for each combination of costs $\left(\frac{c_s}{c_s+c_d}\right)$. It has several noteworthy features.

First, (3.2) is horizontal when either $c_d = 0$ or $c_s = 0$. This implies that the bank is then indifferent over the inventory of reserves that it holds, as long as this is at *most* $m-\epsilon$ (when $c_d = 0$) or at *least* $m+\epsilon$ (when $c_s = 0$). When both costs are positive, (3.2) indicates that the optimal reserve inventory is unique and increasing in the cost ratio $\left(\frac{c_s}{c_s+c_d}\right)$. In general, however, the reserve inventory does not equal the bank's expected reserve need m: only when the respective costs are exactly equal, i.e. when $c_d = c_s$, will the bank choose to hold reserves equal to its expected reserve need.

Secondly, due to the symmetry assumption, the *variance* of the random disturbance δ has a predictable impact on the optimal level of reserves, although this impact will depend on the cost ratio. A change in the variance of δ in this example of a uniform distribution is represented by a change in the parameter ϵ, and its impact on r^* is given by:

$$\frac{\partial r^*}{\partial \epsilon} = 1 - 2\left(\frac{c_s}{c_s+c_d}\right)\,.$$

This indicates that, if $c_s > c_d$, the optimal level of reserves varies inversely with changes in the variance of δ, whilst the converse obtains if $c_d > c_s$.

When $c_s = c_d$, the optimal level of reserves is invariant with respect to the variance of δ. The impact of an increase in the variance of δ is, therefore, a "flattening" of the optimal reserve inventory, as is shown in figure 3.1.

3.2.3 Money market borrowing

Whilst the above model provides an intuitive analysis of how an individual bank will choose its optimal reserve inventory in the presence of various costs[18], it is clearly unsatisfactory insofar as it has left out several essential elements of the environment in which all banks now find themselves. First, banks have access to a money market, which implies that they will not hold a reserve inventory to meet their random settlement obligations, but will borrow (or lend) reserves on the money market instead. Secondly, banks face a reserve requirement, which determines the reserve inventories which banks actually carry. Finally, central banks typically supply reserves to the banking system by means of a discount window.

These three features have been incorporated into a model of the type presented above by Poole (1968). As before, the bank faces an uncertain net settlement obligation δ, but knows the distribution $F(\delta)$. Instead of the optimal reserve inventory, r, however, the choice variable is now the optimal level of *money market borrowing*, denoted in what follows by b. Not surprisingly, this does not change the bank's choice problem in any fundamental way. Assuming the bank has some opening reserve balance ($r_0 \geq 0$), the bank can borrow too much, i.e. $b > -(\delta + r_0)$, in which case it will emerge from the end-of-day settlement with positive reserves. In the absence of an ex-post money market, these reserves must be held overnight, and incur a proportional opportunity cost equal to the interest rate i at which these reserves could have been lent in the money market, i.e. $c_s = i$. If, on the other hand,

[18]One obvious way in which this simple static model may be extended is through the introduction of reserve adjustment costs, which then imply that the bank could no longer costlessly adjust its reserves to the desired optimal level (see Frost 1971). This would not change the properties of the basic model, however, except that it would no longer be optimal for the bank always to adjust its reserves to some desired level.

the bank borrows too little, i.e. $b < -(\delta + r_0)$, its net settlement obligation cannot be met. In this case, the bank must approach the central bank's discount window to borrow its reserve shortfall, incurring a proportional cost d equal to the discount rate. The cost of a reserve deficiency, therefore, is the additional cost which is incurred by borrowing at the discount window, i.e. $c_d = d - i$.

Reserve requirements (and the bank's opening reserve balance) play no fundamental role in the model. They do have to be defined, however, otherwise the bank can trivially meet its settlement obligations. The role of reserve requirements is simply to determine the reserve level around which the bank makes its borrowing decision and can, without loss of generality, be assumed to be zero. Similarly, the bank's opening reserve balance only serves to alter the level of borrowing, but does not change the bank's choice problem, which is to borrow just enough reserves to meet its settlement obligations.

Poole's (1968) model, then, is one in which a representative bank chooses the optimal amount of money market borrowing b^* to meet a random settlement obligation, subject to known costs of over- or under-borrowing. It is worth noting that the *timing convention* adopted in this model is, for analytical reasons, a significant departure from reality. Banks are assumed to base their desired borrowing on a money market interest rate which is the same for all banks (i.e. no credit risk). Banks then each trade *once*, after which time the money market is closed, and no further trading takes place. Thus money market trading is assumed to be centralised and simultaneous. In practice, of course, money markets tend to be decentralised, with prices being established and bargains being struck between two counterparties. Furthermore, trading takes place over an entire day, and banks will incrementally approach their desired borrowing level through a succession of different trades. Clearly such behaviour is exceedingly difficult to model in a simple fashion, which is why a more stylised approach to trading is usually adopted. This is not without its drawbacks, however, which are discussed below.

From the first order condition (3.1) the optimal level of money market

borrowing under this richer institutional structure can be written as

$$F(-b^*) = \frac{i}{d}. \tag{3.3}$$

As in (3.1) the decision – in this case the decision to borrow – depends on only two factors: the cost of a reserve surplus relative to the cost of a reserve deficiency, and the properties of the distribution $F(\delta)$. These costs now have a well-defined interpretation, and it is clear that the level of the market interest rate i relative to the central bank's discount rate is the key determinant of the bank's borrowing decision. Clearly, there are many different money market interest rates, which makes the interpretation of "the" money market rate i somewhat difficult. In practice, however, banks tend to use some instruments far more than others for reserve management purposes, with the preferred instrument nearly everywhere being unsecured loans of very short maturity, typically overnight. Thus i is simply the overnight rate[19]. Before the development of a money market in unsecured funds in the UK, banks would use the bill market for reserve adjustment purposes, in which case i could be thought of as the interest rate on short-term bills.

Although Poole assumes that the δ is a normally distributed random variable, it is hardly plausible that the bank should experience reserve outflows which exceed its own deposits, or inflows which exceed the deposit of other banks. Continuing, therefore, with the simple example developed above, which took the distribution of δ to be uniform over the interval $[m-\epsilon, m+\epsilon]$, the optimal level of borrowing may now be written, using (3.3), as

$$\begin{aligned} b^* &\leq (m-\epsilon) && \text{if } i = d \\ b^* &= (m+\epsilon) - 2\epsilon\left(\tfrac{i}{d}\right) && \text{if } 0 < i < d \\ b^* &\geq (m+\epsilon) && \text{if } i = 0\,. \end{aligned} \tag{3.4}$$

Instead of an optimal reserve inventory, (3.4) traces out the *optimal borrowing (lending) decision* of a representative bank. As before, this is a linear (but discontinuous) function, which in this case depends on the cost ratio $\frac{i}{d}$.

[19]In the US the overnight funds market is known as the federal funds market, and thus i may be thought of as the federal funds rate.

When either $i = d$ or $i = 0$, this function is horizontal, implying that the bank becomes indifferent over the amount of borrowing, as long as this is below $m - \epsilon$ or above $m + \epsilon$, respectively[20]. When $0 < i < d$, however, the borrowing decision is uniquely determined by $\frac{i}{d}$. If the money market interest rate i is "high" (i.e. close to d), the opportunity cost of a reserve surplus is high and the bank chooses to borrow less than its expected settlement obligation (i.e. $b^* < m$), preferring to make up its expected reserve deficiency by borrowing at the central bank. Conversely, if the money market interest rate i is "low" (i.e. close to zero), the bank chooses to borrow more than its expected settlement obligation (i.e. $b^* > m$) because the cost of surplus reserves is low as compared to the cost of a reserve deficit.

Graphically, then, (3.4) exactly resembles the function shown in figure 3.1. The interpretation of (3.4) as a money market borrowing function, however, obviously implies that the horizontal axis now measures borrowing (rather than reserve inventory), whilst the vertical axis measures the level of the money market rate relative to the central bank's discount rate $\frac{i}{d}$ (which is now the relevant cost ratio).

3.2.4 The money market interest rate

How appropriate is the above model as a basis for explaining the determination of money market interest rates, which clearly depends on banks' *aggregate* borrowing behaviour? Poole (1968) has noted that "it is not possible to jump easily from the analysis of the individual bank to the banking system as a whole since reserve flows into or out of a particular bank are not likely to be independent of the reserve flows of other banks" (p.781). In other instances, however, he argues that the money market interest rate will reflect the reserve position of the banking system as a whole. Thus, for example, "a

[20]This is plausible, since (for example) $i = d$ implies that there is no additional cost to borrowing funds from the central bank. To ensure that it must borrow from the discount window, a bank will borrow less than its *minimum* expected settlement obligation in the money market. Conversely, $i = 0$ implies no cost of borrowing reserves in the market, causing the bank to borrow more than its *maximum* expected obligation.

relatively high [money market] rate suggests that *banks in general* are short of reserves [and] .... may indicate that reserves are flowing out of the banks ..." (p.776-7; emphasis added). How, then, can this model be used to derive the money market rate?

Following the discussion in section 1, the expected settlement obligation δ anticipated by an individual bank arises from two distinct sources: (i) interbank payments and, (ii) changes in net currency use and central bank transactions. It is the second of these which is relevant to the determination of the money market interest rate.

Recall that obligations arising from interbank payments (e.g. customers writing cheques) create equal and offsetting obligations in other banks, and are thus not independent across banks. Indeed, they are totally interdependent, and leave the aggregate supply of reserves unaffected. For this reason such obligations, even though they are uncertain ex-ante, could easily be accommodated *without* recourse to a central bank. By establishing an *ex-post* money market, i.e. an opportunity for borrowing and lending funds after settlement obligations have been revealed, banks could easily offset the reserve surpluses or deficits which resulted from periodic settlement. Indeed, they face incentives to do this. By lending their excess reserves in the ex-post market at some rate $i' > 0$, surplus banks can reduce their opportunity costs to $c_s = (i - i') < i$ (recall that i is the interest rate in the ex-ante money market). Similarly, by borrowing reserves in the ex-post market at some rate $i' < d$, deficit banks can reduce their cost of obtaining reserves to $c_d = (i' - i) < (d - i)$ (recall that d is the central bank's discount rate).

Whilst the determination of the interest rates i and i' will not be pursued any further here, two important points emerge from the discussion. First, if reserve flows are generated purely by interbank payments, then the central bank plays no essential role in the determination of the market interest rate, since the ex-post money market can always clear without banks having to borrow from the central bank. Second, this would imply that considerations *other* than those pertaining to bank reserve management as detailed so far were necessary to obtain a determinate money market interest rate: the ex-ante and ex-post money markets could clear at *any* rate $i = i'$, and a model

of reserve management does not suggest which one.

The settlement obligations which are relevant to the determination of the money market rate, therefore, are those of the second type, namely net currency withdrawals and central bank transactions, which affect the *aggregate stock of reserves.* This second category of obligations makes a central bank necessary, since these obligations will not, in general, cancel out across banks. What is more, these obligations are likely to be much less interdependent than those resulting from interbank payments.

3.2.5 An aggregate model

To examine the impact of these factors on the money market interest rate, an aggregate version of the reserve management model above can be developed with $j = 1, \ldots, n$ more or less identical banks.

Uncertainty, as before, is summarised by a single variable δ_j, which represents the random daily settlement obligation of each bank. It is assumed that these settlement obligations do not, in general, cancel out across banks in aggregate. Specifically, it is assumed – not unrealistically – that these obligations arise because the government maintains its accounts with the central bank. Each period, tax payments are made into this account via cheques written on banks, who then seek to offset their expected (negative) settlement obligations by borrowing in the money market. Similarly, the government also makes daily disbursements out of this account via cheques drawn on the central bank. When these are deposited with banks, they seek to offset their expected (positive) obligations by lending in the money market.

Thus, as before, each bank faces a uniformly distributed settlement obligation δ_j on the interval $[m_j - \epsilon, m_j + \epsilon]$, which here is the net outcome of payments to and from the government, and differs across banks[21]. Since banks are identical except for their expected settlement obligation, the bor-

[21]It is assumed, for simplicity, that the variance of each banks' obligation is the same, i.e. $\epsilon_j = \epsilon$.

rowing function of each bank $j = 1, \ldots, n$ can be written, from (3.4), as

$$\begin{aligned} b_j^* &\leq (m_j - \epsilon) && \text{if } i = d \\ b_j^* &= (m_j - \epsilon) - 2\epsilon\left(\tfrac{i}{d}\right) && \text{if } 0 < i < d \\ b_j^* &\geq (m_j + \epsilon) && \text{if } i = 0\,. \end{aligned} \tag{3.5}$$

Each bank, individually, takes the money market interest rate as given when making its optimal borrowing decision. However, aggregate borrowing demand, denoted by $B^d \equiv \sum_{j=1}^{n} b_j^*$, determines the market clearing interest rate. It can be seen from (3.5) that this will depend on – amongst other things – the *aggregate* expected settlement obligation, denoted by $M \equiv \sum_{j=1}^{n} m_j$. This aggregate net obligation M may usefully be referred to as the expected money market shortage ($M > 0$) or surplus ($M < 0$), and arises in this simplified setting from net disbursements from the government's account held at the central bank. In reality a number of other factors, most notably currency withdrawals and central bank transactions, will also contribute to the money market shortage or surplus. These are ignored here to keep the exposition simple.

Aggregating (3.5) gives

$$\begin{aligned} B^d &\leq (M - n\epsilon) && \text{if } i = d \\ B^d &= (M - n\epsilon) - 2n\epsilon\left(\tfrac{i}{d}\right) && \text{if } 0 < i < d\ , \\ B^d &\geq (M + n\epsilon) && \text{if } i = 0 \end{aligned} \tag{3.6}$$

which defines the *aggregate demand for money market borrowing*, and is depicted in figure 3.2. As expected, this looks similar to the individual borrowing function (3.4), and is again a linear, discontinuous function of $\frac{i}{d}$. Notice, however, that changes in M cause a one-for-one change in aggregate borrowing demand in the money market. As M rises above zero the money market shortage increases, and hence aggregate money market borrowing increases proportionately at any given market interest rate. Similarly when M falls below zero the money market surplus increases, and hence aggregate money market lending increases proportionately. In short, M is a *shift* parameter in the aggregate money market borrowing function.

Whilst M (and some other factors) determines the aggregate demand to borrow in the money market, the aggregate supply of funds is, by definition, constrained by the condition

$$B^s = 0\,.$$

This is depicted as a vertical line at B = 0 in figure 3.2. Thus the money market interest rate i must adjust to clear the money market, i.e. achieve $B^d = B^s$. By imposing this condition upon the borrowing function in (3.6) it is possible to solve for the market clearing rate i^* as

$$\begin{aligned} i^* &= d && \text{if } M \geq n\epsilon \\ i^* &= d(2n\epsilon)^{-1}(M + n\epsilon) && \text{if } -n\epsilon < M < n\epsilon \\ i^* &= 0 && \text{if } M \leq -n\epsilon\ , \end{aligned} \tag{3.7}$$

which clearly depends on the money market shortage (surplus) M. When M rises above zero (shortage), banks in aggregate will want to borrow more than they can, i.e. $B^d > B^s$. Thus the money market rate must rise, since a higher interest rate is consistent with a lower borrowing demand for each individual bank. To clear the money market i must rise by enough to reduce aggregate borrowing demand to $B^d = 0$. For large enough M (large shortages), the money market rate becomes equal to the central bank's discount rate (see figure 3.2). This clears the money market because banks then become *indifferent* as to whether they borrow reserves from the central bank or in the money market, and also prevents the market rate from rising above the discount rate.

Conversely, when M falls below zero (surplus), banks in aggregate want to lend more than they can, i.e. $B^d < B^s$. Thus the money market rate must fall to encourage institutions, individually, to borrow more (i.e. to the point where $B^d = 0$). When the surplus becomes sufficiently large, however, the money market rate falls to zero, which is consistent with market clearing since banks then become indifferent between holding excess reserves and lending these on the money market (see figure 3.2).

In this model, then, the aggregate expected settlement obligation – the money market shortage or surplus – is the key variable that determines the

money market interest rate. Given that reserve requirements are zero, the shortage (surplus) corresponds to the expected aggregate supply of reserves, and so the money market interest rate responds to the expected aggregate availability of reserves. This is consistent with the discussion in section 1, where it was asserted that the behaviour of the money market interest rate was the outcome of profit-maximising decisions of individual banks, each attempting to manage their reserves around some target level in the face of random fluctuations[22].

3.2.6 Can the discount rate be a "penalty" rate?

It is often asserted that the discount rate is, or could be, a "penalty" interest rate, levied by the central bank on those banks who commit reserve management "errors" and who have to make good their shortfall by borrowing from the discount window. In this model, and indeed in any model which incorporates reserve management, it cannot be the case that the discount rate is a penalty rate. While it is true that, ex-post, some banks may have to borrow from the discount window, possibly at a higher rate than they could have borrowed funds in the money market, this simply reflects the fact that, under uncertainty, it is unlikely that expected costs coincide with actual costs. Ex-ante, however, banks have full information about the discount rate, and are free to choose their money market borrowing accordingly. Whilst they could choose to borrow in the money market so as to be (virtually) certain that they would not have to approach the central bank discount window, this is generally sub-optimal since they will then typically end up holding excess reserves, which is also costly. Instead of a penalty, then, the discount rate (and the non-payment of interest on reserves) should be seen as *cost parameters* given by the institutional framework in which banks operate. Reserve

[22]Models of money market interest rate determination, also based on bank reserve management, have been proposed by Ho and Saunders (1985) – who use a mean-variance framework – and by Spindt and Hoffmeister (1988). In each case the basic result is the same: the money market rate depends on the aggregate reserve position of the banking system. See equation [18] in Ho and Saunders (p.983) and equation [24] in Spindt and Hoffmeister (p.411).

management represents the attempt, by banks, to minimise these costs.

A clearer indication that the discount rate could not be a penalty rate in this framework is that banks' optimal money market borrowing choices are *invariant* to the absolute level of the discount rate. It would not be possible for the central bank, even if it wanted to do so, to somehow penalise banks by raising the discount rate. Consider the effect of an increase in the discount rate from d_0 to d_1. Given M, and the existing market clearing interest rate i_0^*, this lowers the *relative cost* of funds in the money market from $\frac{i_0^*}{d_0}$ to $\frac{i_0^*}{d_1}$), thereby raising the optimal borrowing demand of all banks. Of course, the aggregate supply of funds is given, with the result that the money market rate will *rise* to clear the market. Specifically, the money market rate will rise to the point where the relative cost of funds between the discount window and the money market is restored to its original level, i.e. $\frac{i_1^*}{d_1} = \frac{i_0^*}{d_0}$.

The outcome of an increase in the central bank's discount rate is, *ceteris paribus*, a rise in the money market rate. This is hardly surprising, and illustrates the fundamental point that *a money market shortage is sufficient for the central bank to influence, fairly precisely, the money market rate.* By the appropriate choices of the discount rate, the central bank can then raise or lower the level of money market interest rates in such a way as to achieve its intermediate and ultimate policy objectives. A higher discount rate, however, does *not* penalise banks in the sense that a reserve shortfall is now somehow more costly than before[23]

This result obtains because the relative cost of funds $\frac{i^*}{d}$ in (3.7) is, for any given discount rate, uniquely determined by the *magnitude* of the money market shortage M. It is in this respect, however, that the aggregate model is also somewhat misleading, and reflects the overly simplified manner in which uncertainty and money market trading are incorporated into the model.

[23]Of course, a higher discount rate, by raising money market rates, would penalise those banks which are exposed to interest-rate risk, e.g. those banks who had borrowed funds on a floating rate basis to fund fixed rate loans.

3.2.7 Should the size of the shortage matter?

The fact that the size of the money market shortage plays a role in this model is problematic for the following reason. When there is a money market surplus, for example, banks in aggregate already have more than enough reserves to meet their settlement obligations. Whilst this makes it unavoidable that some banks must end the day with excess reserves, these reserve holdings will be higher than they need to be, since those banks which have a reserve deficit after settlement will borrow from the discount window. Similarly, although a money market shortage makes it unavoidable that some banks must borrow from the central bank to meet their settlement obligations, this borrowing will be higher than it need be, since some banks will be left holding surplus reserves after settlement.

These outcomes are clearly counterfactual, since one would not expect some banks to be holding surplus reserves whilst others were borrowing from the central bank. The function of the money market, after all, is to redistribute reserves, which would tend to mitigate such an outcome. In practice, therefore, one would expect the *sign* of M (i.e. a shortage $M > 0$ versus a surplus $M < 0$) to play a role in the determination of money market rates, but not the size of M. In other words, the slope of the aggregate money market borrowing function would be much steeper in practice than that derived in the model, since it would quickly become apparent, through banks borrowing behaviour in the money market, whether there was a money market surplus or shortage. This would lead, in the case of an aggregate surplus, to the money market rate falling towards zero, since there would be banks who, based on their expected settlement obligations, would still be willing to lend when all other banks had met their borrowing targets. Similarly, in the case of an aggregate shortage, the money market rate would rise to equal the discount rate, since some banks would always remain unable to meet their borrowing targets by borrowing funds in the money market.

This information revelation function, which is an integral part of all financial markets, is basically assumed away in the model as a consequence of invoking the extreme trading assumption that banks only have a single oppor-

tunity to trade in the money market before all uncertainty is revealed. Thus although trading in the money market reveals the sign of M (i.e. whether there is an expected surplus or shortage) through the level of i – from (3.7) $i > \frac{d}{2}$ implies an expected shortage, and vice versa – this information is no longer of any use to banks without another opportunity to trade in the money market. It is then hardly surprising that some banks will end up with a reserve surplus even when there is a money market shortage, and also that some banks will end up with a reserve deficit even when there is a money market surplus.

But what if banks could trade in an ex-post market before they borrowed from the central bank? Then, banks with surplus reserves could lend these to banks which still needed reserves. If, however, $i^* < \frac{d}{2}$, revealing an expected aggregate surplus, competition between lenders would drive the ex-post money market rate close to zero, since this is the return they would receive if they didn't manage to lend their excess reserves. If, on the other hand, $i^* > \frac{d}{2}$, revealing an expected aggregate shortage, competition between borrowers would immediately drive the ex-post money market rate very close to the discount rate, since this is the cost of funds if banks didn't manage to cover their reserve deficit.

Relaxing the trading assumption, therefore, would permit the money market to redistribute reserves more effectively than in the existing model, and would also make the money market rate (at least in the ex-post market) much more sensitive to the aggregate reserve position of the banking system. At the same time, however, the simple model of reserve management assumed above would become potentially more complicated, since the expected money market shortage would also presumably become relevant to the representative bank's ex-ante borrowing decision. Such a more complicated model will not be pursued here. In any event, the existing model can provide a good first approximation if it is assumed that banks expect their individual random settlement obligations δ to be distributed on a narrow interval, i.e. ϵ is "small". If this is the case, the aggregate money market borrowing function becomes steep, and only a very small shortage (surplus) is necessary to drive the market clearing interest rate up to the central bank's discount rate (down

to zero).

It is then also much more obvious that the discount rate cannot, in any sense, be considered as a penalty rate. A money market shortage would (almost) always drive the market rate to equality with the discount rate, eradicating even the ex-post cost of those banks which had to borrow from the discount window on account of the shortage. A money market surplus, on the other hand, would (almost) always mean that the discount rate was irrelevant, since market interest rates would fall to zero and banks would not have to borrow from the discount window on account of the surplus.

3.3 Refinements

In the model just described, the money market interest rate is prevented from rising above the discount rate in the event of a money market shortage because the central bank is prepared to lend reserves through its discount window[24]. Should the money market move into surplus, however, the money market rate will fall below the discount rate and, in exceptional circumstances, may fall to zero. Simply maintaining an open discount window – as was done by central banks during the nineteenth century – cannot prevent money market rates from falling. To achieve this, one of two actions is necessary by the central bank.

The central bank could undertake open-market operations (such as the sale of financial securities, for example) to reduce aggregate reserves (i.e. increase M). This would increase aggregate borrowing demand and cause the market rate to rise. Alternatively, the central bank could extend a *borrowing* facility that was analogous to the discount window except that banks could buy securities from, rather than sell securities to, the central bank at some announced rate, say d'. This would mean, in the above model, that the cost of a reserve surplus would become $c_s = i - d'$, with the result that (3.3), the

[24]This assumes, of course, that there are no *limits* on, or *non-pecuniary* costs attached to, borrowing at the window; these features are addressed below.

borrowing function of the individual bank, could be rewritten as

$$F(-b^*) = \frac{i - d'}{d - d'} .$$

The effect of d' is thus to impose a floor on the money market interest rate since, when the rate reaches this level, banks become indifferent between lending reserves in the money market and lending them to the central bank. In the analysis above it was assumed that such a facility did not exist, i.e. that $d' = 0$, which meant that the market rate could fall to zero in the event of a (large) money market surplus. With $d' > 0$, however, the central bank could prevent this fall in money market rates. Although there are several examples of central banks employing such a facility[25] for this purpose, central banks have generally preferred to utilise the first channel, namely open market operations.

Through open-market operations, central banks can always make their discount rates effective, since the effect of such operations is to change M, the money market shortage. Through the appropriate choices of M, therefore, central banks can always cause the money market interest rate to remain at or near their discount rates. As noted earlier, this amounts to central banks engaging in reserve management at the aggregate level to "control" the level of some money market interest rate. The view of many academics notwithstanding, this is also what central banks perceive themselves to be doing. Thus:

> "Open market operations allow the Trading Desk at the New York Reserve Bank to adjust the volume of reserves in the [banking] system before depository institutions borrow at the Federal Reserve - that is, to manage nonborrowed reserves. In this way the Fed can offset the reserve swings caused by the public's demand for cash and by other factors. *By managing non-borrowed reserves in relation to estimated reserve requirements, the Fed can ... affect the interest rate institutions pay each other when borrowing overnight* - the federal funds rate." (Meek 1982, p.10; emphasis added)

[25]The Bank of England offered such a facility to deal with the enormous money market surpluses generated by government spending during the first world war (see chapter one). The Bundesbank currently operates such a facility, under which banks are invited to buy very short-term securities at a stated rate (see Deutsche Bundesbank 1989).

In order to influence money market interest rates with a reasonable degree of precision, however, central banks have to take into account several features of the money market environment and of their own operating procedures which were only partially taken into account in the model above. These are the structure of reserve requirements and the discount window, each of which is discussed in turn.

3.3.1 Reserve requirements

In the model derived in the preceding section it was assumed that the level of reserve requirements which banks had to meet was zero, and that this requirement had to be met on a daily basis (i.e. the maintenance period was a single day). Whilst this is an accurate description of actual reserve requirements in force in the UK, it is less accurate when applied to the US or several European countries, where reserve requirements are positive and met over a longer maintenance period. Thus, whilst banks remain concerned about the level of aggregate reserves over the entire maintenance period, day-to-day fluctuations will be of somewhat less concern, since on all but the last day of the maintenance period banks face an *average* reserve target.

This means that banks face a different reserve management problem on all but the last day of the maintenance period. In particular, holding reserves in excess of the reserve target carries a lower opportunity cost than otherwise, since these reserves contribute towards the bank's average reserve holdings and thus potentially compensate the bank against reserve drains on subsequent days. Similarly, when reserves fall below target, the bank can delay borrowing from the discount window because it might receive reserve inflows on subsequent days. In short, before the last day of the maintenance period banks will not necessarily respond to reserve flows by borrowing or lending in the money market, since these flows might be temporary. Only when they expect reserve flows to persist, will they react in the manner suggested by the model in section 2. The ability by banks to undertake *reserve averaging*, therefore, implies that the relationship between aggregate reserve movements and the money market interest rate is not likely to be constant

over the maintenance period: specifically, this relationship is likely to become more sensitive as the end of the maintenance period approaches. Several authors have noted this phenomenon, notably Poole (1968) himself who argued (and found evidence) that the reserve averaging caused the variance of the money market interest rate to increase over the maintenance period[26].

The ability for reserve accounting to affect banks' reserve management behaviour means that it is clearly important to central banks, who attempt to influence money market interest rates through the management of aggregate reserves. The nature of this concern is perhaps best illustrated by a debate which occurred during the 1960s between the Federal Reserve and several critics, notably Brunner and Meltzer (1964) and Cox and Leach (1964a), who argued that the structure of reserve requirements was problematic:

> "The present system of reserve settlement periods was adopted in the early years of the Federal Reserve as a compromise of widely varying practices among the district Reserve Banks .... These regulations are particularly objectionable in that they require a huge volume of 'defensive' open market operations, interfering unnecessarily with private security markets. Furthermore, the settlement procedures at times pose operating difficulties for banks. In addition to the fact that they magnify the impact of random deposit fluctuations, the current regulations ... destabilise trading in federal funds" (Cox and Leach 1964a, p.93).

To correct these perceived deficiencies, it was proposed that the Fed should *lengthen* the reserve maintenance period (from one week to one month) and, more important, should *stagger* these maintenance periods so that they overlapped with one another (i.e. a new maintenance period would begin before the existing one elapsed). These changes, it was argued, would permit banks to engage in reserve averaging to a greater extent, thus stabilising money market rates and requiring fewer offsetting operations by the Fed.

---

[26]Twenty years later, Spindt and Hoffmeister (1988) rediscovered this result, presenting virtually the same argument (and evidence) as Poole did in 1968. Campbell (1987) presents a more intriguing explanation as to why changes in money market rates may be predictable over the maintenance period, based on imperfectly observable shifts in banks' desired reserve holdings over the maintenance period.

The Fed's response, put forward by Sternlight (1964), was that greater reserve averaging on the part of banks would have two undesirable consequences. First, it was far from obvious that banks, given longer maintenance periods, would, or even could, redistribute current reserve disturbances in a smooth fashion throughout the remainder of the maintenance period. Instead, given uncertainty about future reserve disturbances, these could *cumulate*, eventually disrupting money markets to an even greater extent, and requiring even greater offsetting operations by the Fed. Second, and of greater concern to the Fed, was that staggered maintenance periods would leave banks with *too much flexibility* over their reserve positions, thereby reducing the ability of the Fed to influence money market interest rates in keeping with its policy objectives:

> "... it must be noted that what for the individual member bank is the burden of meeting reserve requirements or the incentive of putting excess reserve to work is, for the banking system as a whole, the backbone of central bank regulation of the credit base .... under the proposed plan, the [Fed] would not have nearly as good an idea as it does now of where the banking system stood at any particular time with respect to reserve availability in relation to requirements" (Sternlight 1964, p.94).

In other words, the Fed's influence on money market interest rates was premised, to a significant extent, on a *predictable* relationship between banks' borrowing behaviour in the money market and aggregate reserves. Staggered reserve maintenance periods would clearly loosen this relationship, interfering with the Fed's ability to use aggregate reserves to produce a desired money market interest rate[27].

When the Fed subsequently revised the structure of member banks' reserve requirements, it went some way towards satisfying the proposals for staggered settlement by introducing a carry-over provision. This permitted banks to deviate (within a specified margin) from their reserve targets dur-

[27]This concern was never properly addressed by the Fed's critics, who argued, somewhat paradoxically, that the Fed could always increase the magnitude of its offsetting operations if it felt that the banking system was too unresponsive in reflecting a desired policy stance (see Cox and Leach 1964b, p.535).

ing the current maintenance period, making up this deviation in the subsequent maintenance period. Furthermore, the Fed also eased the calculational aspects of its reserve requirements by moving from the existing system of contemporaneous reserve accounting (i.e. one in which the computation period coincided with the maintenance period) to a system of lagged reserve accounting (i.e. one in which the computation period preceded the maintenance period). This structure has been retained by the Fed, with some revisions[28], and there have been several examinations of the properties of these alternative arrangements[29].

The upshot is that different regimes of reserve requirements clearly have different implications for bank reserve management, and hence for central bank control of money market interest rates. As long as they are consistent with a determinate demand for reserves on the part of banks, however, they do not change the conclusions of the basic model presented in section 2 in any way. They may, however, be usefully included in such a model to provide greater insights about the precise relationship between aggregate reserves and money market rates.

3.3.2 The discount window

Although central banks have increasingly come to rely upon open-market operations as their primary means of influencing bank reserves and money market interest rates, the structure of their discount facilities is sometimes an important component of the manner in which they achieve this influence. Again, this is less true of the UK than it is of other countries, the primary example being the US. Nonetheless, it is useful to briefly examine the relevance of the restrictions imposed by central banks over their discount facility, and its impact on the relationship between aggregate reserves and money market

[28]In 1984 the computation period was adjusted to overlap with all but the last two days of the maintenance period, whilst in 1990 the level of requirements was dropped (see Feinman 1993b).

[29]See Cosimano (1987), Saunders and Urich (1988), Evanoff (1990) and Brunner and Lown (1993).

interest rates.

Without restrictions over the discount facility, the cost of funds from the facility is fully captured by the single parameter d. Under this condition, the model in section 2 indicates that the money market interest rate never exceeds the central bank's discount rate, since an institution would never prefer to borrow in the money market at a higher rate than the cost of funds at the discount window. This behaviour is certainly borne out in practice. In France or Germany, for example, where funds are available at a preannounced discount rate to any bank with the appropriate securities, very short-term money market rates never rise above the official discount rate[30].

Once restrictions are imposed on the discount facility, however, the discount rate ceases to capture the full cost of funds. Depending on the nature of the restrictions, the cost of funds will usually be *above* the discount rate and may *increase* with the amount of borrowing, the frequency of borrowing, etc. Naturally, banks will take these additional cost factors, which may be pecuniary or non-pecuniary, into consideration when making their money market borrowing decisions. And, whilst money market rates will never rise above the point at which banks become indifferent between the cost of funds at the discount facility and in the money market, this level may now be above the discount rate, since this no longer captures the full cost of discount window funds. Note, however, that the observed money market rate remains an *outcome* of the particular cost structure imposed on the discount facility by the central bank, which has two implications.

First, when the money market rate rises above the discount rate, it does *not* mean that there is then an arbitrage opportunity between the window and the market, or that those banks which are borrowing from the window are receiving a "subsidy" rate. Waller (1990), for example, believes this is the case:

> " ... a problem arises in that, whenever the market interest rate is above the central bank's lending rate, there is an arbitrage opportunity to exploit by borrowing from the central bank and relending at the higher market rate

[30]For evidence see Schnadt (1994a and 1994b).

> of interest. If all borrowing requests were granted, the market rate of interest would be driven down to the central bank's lending rate, which, in turn, may hinder the achievement of other monetary objectives. In order to prevent this, the central bank must engage in non-price rationing to allocate its credit" (p.273).

It is disturbing to find such a blatant confusion of cause and effect. Why would profit-maximising banks be prepared to pay more for funds in the money market than they would at the discount window? The answer is that they would not. Clearly, therefore, it is *because* of non-price rationing at the discount facility that the money market rate is above the discount rate in the first place[31].

Secondly, and more crucially, the central bank must take into account the manner in which its restrictions affect banks' behaviour, since this will affect the nature of the relationship between the aggregate reserves of the banking system and money market interest rates. To illustrate this, two types of restrictions – pecuniary and non-pecuniary – will briefly be contrasted, since these have a considerably different impact on bank behaviour.

The discount facility of the Riksbank of Sweden provides a good example of the first type of restriction, and represents what has been referred to as an "interest rate scale" (see Englund, Horngren and Viotti 1989 and Martensson 1992 for details). Basically, banks each face a *set* of interest rates at which they can borrow from the central bank, which is increasing in the amount borrowed in a manner which depends upon their capital. Furthermore, banks each face a similar set of rates at which they can deposit reserves at the central bank, which is decreasing in the amount deposited, and also depends upon their capital. This arrangement differs from that defined in the basic model of section 2 only in that a bank's costs of a reserve deficit or surplus are no longer constant, but comprise a number of "steps" depending on the amount of the deficit or surplus (i.e. the amount borrowed or lent at the discount facility).

[31]Waller also raises a "non-problem" when he maintains that equality between the discount rate and the money market rate could somehow "hinder the achievement of monetary objectives". The discount rate is a choice variable of the central bank, and thus can always be adjusted, along with money market rates, to reflect the objectives of the central bank.

This is then true for the aggregate money market borrowing function as well, with the result that the Riksbank can, by altering aggregate reserves (M), put the banking system at a particular point on the aggregate interest rate scale, which then determines the money market interest rate. By withdrawing reserves via open-market operations, for example, the Riksbank causes some banks to increase their borrowing at the discount facility. This pushes them, and thereby money market rates, to a higher point on the interest rate scale[32].

The discount window of the Federal Reserve (or Fed) is the most well known example of the second type of restriction, which involves non-pecuniary costs of borrowing. According to Mengle (1986):

> Nonprice rationing of Fed credit became firmly established as a matter of practice during the late 1920s .... For example, in 1926 the Board adopted a policy of discouraging continuous borrowing from the discount window. In 1928, it specifically stated that banks should not borrow from the window for profit. Since then, the Fed has emphasised nonprice measures along with the discount rate to control borrowing" (p.31).

Although the impact of such nonprice rationing on the relationship between aggregate reserves and money market interest rates is now well understood, it took some time before officials at the nascent Fed realised the impact of these restrictions. As explained by Burgess (1946) and Riefler (1930), if banks were discouraged from continuous borrowing at the window, they would attempt to borrow in the money market instead, thus driving money market rates above the discount rate. Furthermore, open market operations by the Fed to increase reserves would, under these circumstances, be employed by banks to repay their discount window borrowing, and would thus have no impact on the aggregate availability of reserves. They would, however, cause money market rates to fall, which in its turn might increase the availability of credit[33].

---

[32]Banks still have an incentive to use the money market since their marginal cost of funds from the discount facility is not the same. Banks with more capital face a somewhat flatter interest rate scale, which provides an incentive for banks with less capital to borrow from them, via the money market, before borrowing from the central bank.

[33]Wicker (1965, 1966 and 1992) gives an excellent account of the early development and

Subsequent events made the discount window, and its attendant costs, largely irrelevant until the 1950s (see chapter 1), when discount window regulations (Regulation A) were revised. At the same time, the development of the federal funds market meant that banks began to rely increasingly upon the money market for reserve management purposes (see Boughton 1972). Thus the federal funds rate began to reflect the aggregate scarcity of reserves to a much greater extent than previously, periodically rising above the discount rate when reserves were scarce (see Minsky 1965 and Mengle 1986). This spread between the fed funds rate and the discount rate became more or less permanent during the 1970s. Numerous econometric studies attempted to describe the relationship between the spread and borrowing, usually based on the implicit assumption that it was the incentive provided by this spread which *caused* borrowing to increase (see Goldfeld and Kane 1966 or Frost and Sargent 1970). As has been pointed out above, the causality runs in the other direction: it is because banks are reluctant to borrow from the window that the spread appears.

To some extent the issue of causality is irrelevant, since it is clear that the Fed determines – through its open market operations – the aggregate volume of borrowing from the discount window. What matters to the Fed, therefore, is the *stability* of banks' discount window borrowing behaviour, since the determination of any particular fed funds rate by the Fed is premised on some *assumed* reluctance on the part of banks to borrow from the window. Goodfriend (1983) has argued that the structure of the nonprice rationing employed by the Fed is itself a *source* of instability in this relationship since "borrowing demand depends in a potentially complicated way on lagged levels of borrowing and on expected future spreads" (p.354). Indeed, there is ample evidence from the Fed itself that such instability has made determination of its desired fed funds rate extremely difficult at times[34].

---

significance of these ideas regarding the relationships between open-market operations, discount window borrowing and money market rates.

[34]Each year, in the spring issue of its *Quarterly Review*, the New York Federal Reserve publishes a detailed summary of its money market operations in the previous year. Increasing bank reluctance to borrow, attributed to increased worries that such borrowing

Why, then, does the Fed choose to impose non-pecuniary costs over its discount facility in this manner? Although difficult to establish conclusively, the most plausible reason is probably a desire by the Fed to maintain the perception that it is *not* directly responsible for determining the level of the federal funds rate, thereby reducing the influence of political forces on its policy instrument. Indeed, several commentators (see, for example, Greider 1987) have presented a convincing case that the Fed's adoption of supposedly "new" operating procedures in October 1979 were motivated by the same concerns. By engineering large increases in member bank borrowing at the discount window under the guise of a supposed change in operating procedures, the Fed succeeded in raising the fed funds rate to unprecedented levels (which it needed to do to control inflation). These increases in the fed funds rate were a direct consequence of banks' reluctance to borrow from the window, which was thus usefully exploited by the Fed to cloak its own responsibility for such high interest rates. Raising its discount rate to this level would have achieved the same result, but would have attracted enormous political opposition. By arguing, instead, that the fed funds rate was rising as a result of market forces, it successfully shrugged off this pressure, and succeeded in bringing inflation down. Nonprice rationing, therefore, was not without its advantages.

In sum, the structure of the discount window, like the structure of reserve requirements, is an important determinant of the specific relationship between aggregate reserves and money market interest rates which will hold in practice (see Pearce 1993). They do not, however, alter the basic insight of the model developed in this chapter: central banks are – albeit through a variety of means – directly responsible for the observed level of money market interest rates.

---

might be perceived as a sign of liquidity problems, was reported in 1988, 1990 and 1991 (see Garfinkel 1990). Substantiating this, Persitiani (1991) notes the response by the Fed chairman to a proposal to open discount window borrowing to greater public scrutiny: this, according to the chairman, " could be expected to have a chilling effect on the use of the discount window" (p.29).

3.4 Figures and tables

Figure 3.1
Optimal reserve inventory

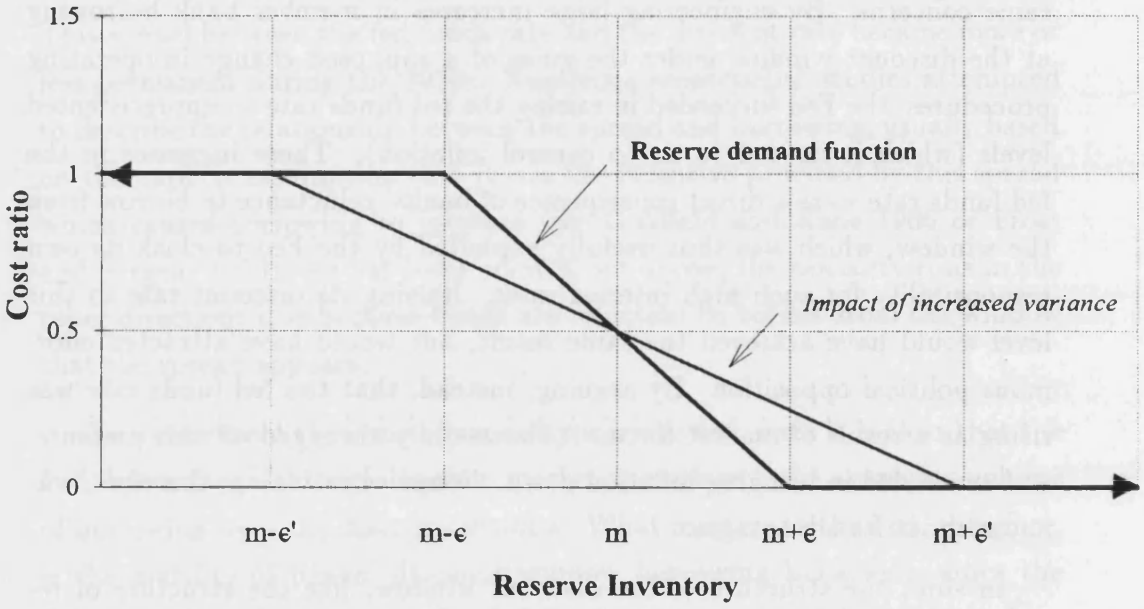


Figure 3.2
The determination of the money market interest rate

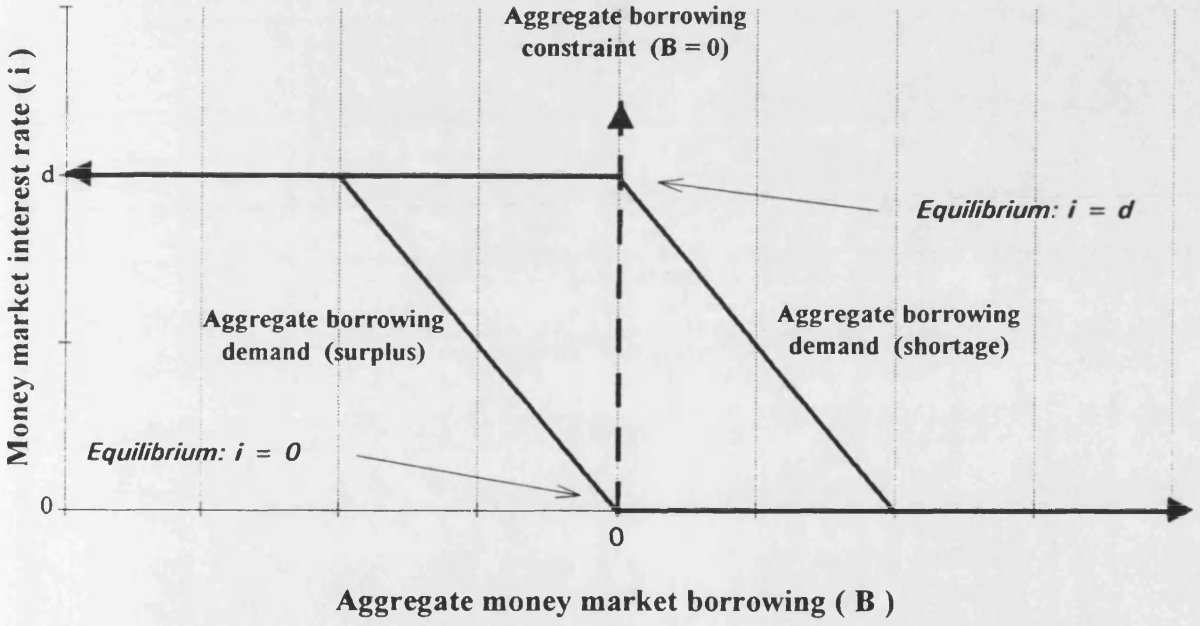


Table 3.1
Bank Reserves in Five Countries

Country	Bank Reserves in 1992 (£bn)	
	Total Reserves (incl. requirements)	Excess Reserves (excl. requirements)
France	2.17	0.06
Germany	35.41	0.49
Japan	14.89	0.02
U.K.	1.50	0.18
U.S.	20.30	0.57

Source: Bank for International Settlements (1993).

Chapter 4

The Sterling Money Market and Bank of England Operations

The purpose of this chapter is primarily descriptive, and it is split into four sections, each of which details a different aspect of the sterling money market [1]. Section 1 describes the financial instruments which are traded in the sterling money market, briefly covering their history as well as their associated issue and trading conventions. Section 2 then describes the transactions technology whereby trades in the sterling money market are settled, and also outlines some of the risks inherent in the settlement process. Section 3 describes the daily operations of the Bank of England in the sterling money market, focusing on the assets, counterparties and timing of these operations, and the impact of these operations on money market interest rates. Finally, section 4 describes two recent developments which have changed the longstanding role of the discount houses as the Bank's counterparties.

[1] This chapter is a shortened and revised version of two published papers by the author on the sterling money market. See Schnadt (1994a, 1994b).

4.1 The Sterling Money Market

The sterling money market is defined here as *the market for the issue and trading of sterling denominated wholesale debt claims with an original maturity of one year or less* and covers a class of assets which includes Treasury and commercial bills, certificates of deposit, unsecured and secured wholesale deposits and commercial paper[2]. Gross daily turnover in this market may easily approach £40 billion, which is second only to turnover in the London foreign exchange market[3]. The outstanding amount of borrowing in this market is much larger and is estimated, according to table 4.1, to approach £300 billion. As trading in the sterling money market is relatively unregulated, the role of conventions is particularly important in maintaining the integrity of the market and its participants. While these conventions are, to some extent self-imposed, the Bank of England has published a *London Code of Conduct* (May 1992) for principals and brokers in the sterling money market, which details the general standards for dealing in this market and contains a statement of best practice.

The transactions technology of the sterling money market has, for many years, been such that physical contact between participants is unnecessary. Closed circuit television services (e.g. Reuters) convey price information to a dispersed set of buyers and sellers, comprised of banks and other financial institutions, corporations and specialist broker and dealer firms. These prices are negotiated, and deals made, by telephone, and the transfers of money and securities which finally settle these deals are then achieved via electronic settlement systems.

The discussion which follows will divide the sterling money market into

[2]Several texts and articles contain good descriptions of the sterling money market (although at very different levels of generality), detailing aspects such as history, instruments, participants, market conventions and Bank of England operations. See, for example, MacKinnon and Dowding (1989), Wilson (1989), Harrington (1991), Walmsley (1992) or, more recently, Llewellyn and Tew (1993).

[3]Activity in the London equity market, for example, is small by comparison, being only about £1 billion daily.

three sub-markets: the discount market, the deposit market, and the commercial paper market. This subdivision is artificial, and is made for descriptive purposes only. In practice these sub-markets are totally integrated, in the sense that participants in the sterling money market are not restricted in any way from participating in any of these sub-markets[4]. Thus a corporation can raise short-term funds by drawing commercial bills, by issuing commercial paper, or by borrowing in the deposit market, and will pursue whichever option provides the cheapest cost of funds. Similarly, banks can, and do, invest their short-term funds by purchasing bills or certificates of deposit, or making secured or unsecured loans in the deposit market. Again, the instrument which promises the highest return will be chosen.

Since money market instruments are close substitutes, and because access to these instruments is unrestricted, arbitrage ensures that their yields, *for any given maturity*, will be driven to equality once characteristics such as liquidity, credit-risk and tax have been taken into consideration. If this were not the case, riskless profit opportunities would exist, which would be inconsistent with a market populated by competitive, profit-seeking institutions. This simple equality condition does *not* hold, however, for money market interest rates at different maturities. Instead, rates at longer maturities will reflect market participants' expectations about future short-term rates which, as the previous chapter has shown, will be dependent on the current and future decisions of the central bank about its monetary instrument. This issue, the determination of the term structure of money market interest rates, is taken up in detail in chapter 5 and will thus not be discussed further in the current chapter.

[4]This is not necessarily the case in other domestic money markets. Participation in the federal funds market in the US, for example, is limited to the member banks of the Federal Reserve System. See Schnadt (1994a).

4.1.1 The discount market

History

The discount market, as previous chapters have shown, is the earliest example of a money market in England and dates back to the eighteenth century. The central instrument in this market was, and still is, the commercial bill, which became legal in 1697 and whose issue continues to be governed by the Bill of Exchange Act of 1882[5]. This Act defines the bill of exchange as an unconditional order in writing requiring the drawee (the borrower) to pay a fixed sum of money to the bearer (the lender) at a specified future date. It was common practice, however, for banks to accept bills, i.e. assume liability for payment at maturity, thereby creating a more marketable instrument, known as a bank bill (or banker's acceptance), which bore *two* names.

The early discount market in London was one in which bill brokers brought together the buyers (lenders) and sellers (borrowers) of bills and, as already noted elsewhere (chapter 3), this market soon played an active role in redistributing funds between banks. Furthermore, the Bank of England considered certain bills as "eligible" for discount, which also meant that bills were sold to the Bank when the banking system needed reserves. A banking crisis in 1825 caused this arrangement to change, and created both a new set of financial institutions and a new financial instrument. Instead of holding bills for liquidity purposes, banks voluntarily began to place funds with the bill brokers on a secured but "callable" basis. The bill brokers in turn became principals, known as discount houses, holding bills (and government paper) as security against these call deposits.

These changes meant that the discount market became synonymous with a group of institutions, the discount houses, who were market makers in bills. Banks, therefore, did not borrow and lend short-term funds directly amongst themselves. Rather, they did so through the discount houses, selling bills to, or withdrawing call funds from, the houses when they needed funds and

[5]Holden (1955) traces the origins and development of negotiable instrument law in the UK, whilst Gillett Brothers (1956) contains valuable introductory material on bills of exchange.

buying bills from, or placing call funds with, the houses when they needed to lend funds. Banks also did not deal directly with the Bank of England when they needed reserves. Instead, the discount houses sold bills to the Bank and became the official counterparties of the Bank in its money market operations. This arrangement remained virtually unchanged until the early 1960s, with a single innovation occurring in 1877, when the government began to issue Treasury bills for the first time. Given the increased funding demands of the government during world war one and the attendant decrease in international trade, the issue of Treasury bills soon outstripped the issue of bank bills. This situation persisted throughout the interwar period[6] until 1965, when bank bills outstanding again became larger than the Treasury bill issue (see chapter 6).

Issue and trading conventions

Following the development of the wholesale deposit market in the 1960s, which is detailed below, the discount market has become a relatively small component of the sterling money market. Bank bills and secured loans represent less than 7.4% and 2.5% respectively of the total market and Treasury bills only 1.7% (table 4.1). Nevertheless, the discount market still occupies a pivotal position in the sterling money market and will continue to do so until the Bank changes the manner in which it supplies reserves to the banking system. This will become clear from section 2 below, which describes the Bank's operations, and indeed from chapter 6, which examines the bill market in detail.

Bills do not bear explicit interest and are consequently known as *discount* instruments since the lender purchases the bill at a discount to its face value. Since almost all bills are accepted before being sold into the market or to the Bank of England, the initial price of a (bank) bill is calculated on the basis of a discount rate, which is market determined, plus an acceptance fee, which is determined by the acceptor based upon the credit-standing of the borrower.

[6]Since trade was again disrupted by the Great Depression and government borrowing remained high throughout the interwar period and into the second world war.

This price (per £100 of bills) is calculated as follows[7]:

$$P_t = \left(1 - \frac{m_t\,(d_t + a)}{365}\right), \tag{4.1}$$

where a is the acceptance fee (in percent), d_t is the (annualised) discount rate (in percent), and m_t is the unexpired maturity (in days) of the bill. Naturally, when the bill is resold (rediscounted) in the secondary market, the price formula is exactly as in (4.1), but with $a = 0$. The evolution of the price of a bill between issue and maturity is shown (for two different discount rates) in figure 4.1, and can be seen to rise monotonically over the life of the instrument when the discount rate is held constant. A change in the discount rate, however, causes the current value (price) of a bill to change discontinuously, since the future payoff (the face value) is fixed.

The return, or yield, on such a bill when it is purchased and held to maturity is given by[8]

$$i_t = \frac{d_t}{1 - \frac{m_t\, d_t}{365}}, \tag{4.2}$$

which implies that the yield to maturity on a bill always *exceeds* the discount rate, by an amount which increases with the unexpired maturity m.

Currently, sterling bills are typically drawn in amounts (lots) of £5 million, although smaller sizes are issued depending on the specific needs of the customer. Acceptance fees average about $\frac{1}{2}$%, but will be as low as $\frac{1}{8}$% for the best borrowers and as high as 2% for the worst customers. Bid-ask spreads in the secondary market are usually $\frac{1}{8}$%.

It is conventional practice for the acceptor of a bill to insert a clause on the face of the bill, indicating the nature of the transaction for which the funds are being borrowed (e.g. a distributor of motor vehicles would issue bills claused as "current purchases of motor vehicles"). Since bills are often

[7]Note that the calculation of interest rates for sterling deals is done on a 365-day basis.

[8]For the derivation of this formula consult Stigum (1990) or Walmsley (1992).

issued to finance goods in transit, or inventories, they have a short maturity (or tenor), usually of one month or three months, although shorter and longer maturity bills are issued. For the same reason, the extension of bill finance is often said to take place against self- liquidating transactions, as the borrower can meet the liability of the bill (when it matures) from the proceeds of the subsequent resale of these goods. Not all bills, however, are drawn against an underlying transaction of this sort. Finance bills may be drawn by companies against a general charge over goods in transit (without reference to a specific shipment or set of invoices). Finance bills may also be issued to raise working capital or to finance goods which are sold on a hire-purchase basis.

In addition to their role as an instrument of finance, bills play an important part in the liquidity of banks in the UK. Since bills may be readily sold in the secondary market, they are a means whereby a financial institution can obtain cash quickly and at low cost. For this reason, bills are defined as primary liquid assets, and may be held by banks operating in the UK to satisfy the statutory liquidity requirements which are imposed by the Bank of England. Furthermore, bills are purchased by the Bank in its daily money market operations[9]. However, only specific bills, known as eligible bills, may be used by institutions to meet their liquidity requirements or offered to the Bank in its money market operations. In order to qualify as eligible, bills should[10]:

(i) identify (i.e. clause) the underlying transaction for which this form of financing is being arranged, which should be short-term, self-liquidating and not for capital purposes[11];

(ii) be drawn in sterling for an original maturity of no more than 187 days;

(iii) not be payable outside the UK;

[9]Note that while a bank may hold its own bills for purposes of rediscounting them with the Bank, only those bills accepted by another bank may serve as primary liquid assets.

[10]These eligibility criteria are specified in Bank of England (1988).

[11]In practice, however, this criterion is not strictly applied (see chapter 6).

(iv) not be drawn on, or accepted by, a bank which has a shareholding link with the drawer; and

(v) be accepted by an institution named on a list of eligible acceptors published by the Bank of England.

Before an institution can become an eligible acceptor, the Bank specifies that it must have a substantial and broadly based sterling acceptance business in the United Kingdom and that its bills must command the finest rates in the market. Of course, eligibility immediately confers extra utility upon a bill, i.e. it becomes a so-called primary liquid asset since it may be used to satisfy liquidity requirements and can be sold to the Bank of England. This means that bill rates will be *tiered*, with eligible bills carrying a finer rate of discount than non-eligible bills[12]. It also means that bills yields will be lower than the yields on equivalent assets which do not have this property, i.e. unsecured sterling deposits, reflecting a *liquidity premium*. Figure 4.2 depicts the yield of a one month bill as against the yield on a one month unsecured deposit and indicates, somewhat surprisingly, that this premium has fluctuated considerably. Chapter 6 therefore examines the recent behaviour of this premium in more detail and traces its effects on other short-term interest rates.

4.1.2 The deposit market

History

Up until the 1950s, the sterling money market retained its traditional form as a discount market. A long-standing cartel arrangement between the clearing banks prescribed common bases for the interest rates at which they accepted deposits and made advances. It also prescribed that they maintain a specified fraction of their assets in the form of cash and call money with the discount market. Consequently, there was little wholesale borrowing and

[12]Further rate differentiation, or tiering, between bills in each class may also occur, based on the perceived credit risks associated with each particular issuer/acceptor.

lending amongst banks and virtually no competition by banks for funds from corporate and other lenders.

In 1955, however, a decision by the Treasury that local authorities should raise some of their funding in the open market instead of from the central government (through the Public Works Loan Board) created a demand for short-term wholesale sterling funds. This was met, at the time, by foreign banks (see Bank of England 1966). It was not long before corporate lenders and non-clearing banks realised that funds could be lent and borrowed on a wholesale basis at more attractive rates than had previously been available from the clearing banks, thus providing the impetus for a so-called "parallel" money market. Hence another instrument was introduced to the sterling money market, namely unsecured, wholesale deposits of short maturity. A new group of institutions, money brokers, then arose to facilitate trade in this instrument.

The wholesale deposit market expanded rapidly throughout the 1960s and a further development occurred in 1968, when sterling negotiable certificates of deposit (CDs) were issued for the first time (see Bank of England 1972). A CD is a negotiable instrument issued by a bank (or building society), certifying that a deposit has been made with that bank which is repayable to the bearer, with interest, at maturity. Compared with an ordinary term deposit a CD has the advantage, from the borrowing bank's point of view, that funds will not be withdrawn before maturity. This can, and frequently does, occur with term deposits, when a customer unexpectedly needs their funds: by issuing a CD, on the other hand, a customer can simply sell the security in the secondary market if funds are needed.

Although the clearing banks and the discount houses were initially excluded from this new market, partly due to banking regulations and partly due to their own restrictive cartel agreements, this exclusion did not last very long. In 1971, following the introduction of new banking regulations (which became known as Competition and Credit Control; see Goodhart 1989), clearing banks became major issuers of CDs. At the same time, discount houses began to make markets in these instruments, thereby effectively integrating the discount and deposit markets. Over a period of fifteen years,

therefore, between 1955 and 1970, the sterling money market was radically transformed from a relatively small discount market, with few participants, to a much larger market. It spanned a wide array of instruments and discount houses, clearing banks, non-clearing banks, corporations and other financial institutions competed alongside one another to borrow and lend short-term, wholesale sterling funds. Not surprisingly, this had a profound effect on the traditional business of banking. Instead of gathering retail deposits and then attempting to find a profitable outlet for these funds, banks could now practice *liability management*: that is, they could enter into lending commitments, such as mortgage of hire-purchase finance, funded purely by borrowing in the wholesale money market. Access to a retail deposit base, therefore, was no longer necessary to compete as an intermediary.

There were also new risks associated with these changes, which were well illustrated in 1973-75 by the insolvency of a number of nascent secondary, or fringe, banks which had relied exclusively on the wholesale markets to finance their property loans. Following sharp increases in interest rates and steep falls in property values, these banks soon became unable to roll over their short-term funding, and several failed. Those that survived did so as the result of a forced recycling of deposits, coordinated by the Bank of England and consequently dubbed the "lifeboat" (see Reid 1982).

Issue and trading conventions

The sterling deposit market remains by far the largest component of the sterling money market. At the end of 1993, UK banks and building societies were borrowing some £60 billion via sterling CDs and held about £150 billon in wholesale short-term sterling loans to other banks and overseas borrowers (table 4.1). As noted above, trading in these instruments occurs on a bilateral basis, thus the interest rates which are established will vary between individual deals, being influenced by factors such as the credit-standing of the borrower and the size of the transaction. As in the bill market, bid-ask spreads are narrow and are typically about $\frac{1}{8}\%$.

Unlike bills, which are discount instruments, term deposits and CDs pay an explicit interest return, or yield to maturity, which is applied to the prin-

cipal amount of the deposit and paid at maturity. The price of a CD (per £100 of principal) is thus given by

$$P_t = \left(\frac{1 + \frac{i_0 m_0}{365}}{1 + \frac{i_t m_t}{365}} \right) \tag{4.3}$$

where i_0 and m_0 are the initial yield to maturity (in percent) and the initial number of days to maturity respectively (which serve to determine the payment made by the issuer at maturity), and i_t and m_t are the current yield and the current number of days to maturity respectively. Unlike bills, which are quoted on a discount basis, deposits and CDs are quoted on a yield (to maturity) basis. As with a bill, however, the price of a CD rises continuously as maturity approaches, unless the market yield changes, which then implies a discontinuous jump in the price of the CD in the secondary market (see figure 4.1). Since CDs have the advantage of being negotiable in the secondary market, their (credit-risk adjusted) yields will be marginally lower than those on unsecured deposits, reflecting a slight liquidity premium. As figure 4.3. shows, however, the yield spread between a one month CD and a one month deposit is extremely small, since they are otherwise perfect substitutes.

CDs may only be issued by institutions authorised under the Banking Act 1987 (i.e. banks) or the Building Societies Act 1986 (i.e. building societies), whilst any institution can borrow in the wholesale deposit market[13]. Typical maturities for CDs and term deposits are 12 months, 6 months, 3 months and 1 month, whilst term deposits are also commonly made for periods of seven days and 1 day (i.e. overnight). Activity in these so-called "short-dated" instruments is extremely large, mainly because they are used extensively by financial and other institutions for reserve (cash) management purposes as described in chapter 3. These rates – particularly overnight rates – will thus vary considerably throughout the day, reflecting the aggregate availability of very short term funds (reserves). This aggregate availability is influenced by, amongst other things, the daily money market operations of the Bank of England, which seek to offset underlying movements in aggregate reserves

[13]A detailed outline of the issue and trading conventions in the CD market may be found in the British Bankers' Association booklet *Certificates of Deposit on the London Market: Market Guidelines* (November 1990).

due to changes in the note circulation and other factors (see section 3). Since these operations occur at discrete intervals and do not necessarily imply that reserves immediately reach the wider market, overnight interest rates tend to be extremely volatile as compared with rates at longer maturities (e.g. 1 or 3 months). A number of other factors also contribute towards this volatility, which are examined in greater detail in chapter 7.

It is noteworthy that the interest rates on a number of sterling financial instruments are based upon interest rates determined in the unsecured sterling deposit market. All financial instruments with variable, or floating, interest rates, for example, require a commonly agreed reference rate, as do many derivative securities, such as interest rate swaps, forward rate agreements, and interest rate futures and options. In order to standardise the determination of these reference rates and provide a set of "generic" money market borrowing (and lending) rates, the British Bankers' Association (BBA) have designated twelve prime London banks whose quotations are used to calculate these rates. Known as LIBOR[14], these rates are calculated and published daily for each monthly deposit maturity between one and twelve months, and are fixed at 11h00 in the following way: the quotations of eight (of the twelve) reference banks are collected, the two highest and two lowest quotations are eliminated, and the *middle* four quotations are then averaged.

4.1.3 The commercial paper market

History

A more recent innovation in the sterling money market has been the development of sterling commercial paper (SCP), which could only be issued for the first time in 1986 (see Bank of England 1987a). Essentially, commercial paper is not unlike a bill of exchange or a CD, in that it is a short-term negotiable instrument. Unlike a bill, however, which is normally accepted by a bank and thus becomes the liability of the acceptor *and* the drawer (i.e.

[14]London Interbank Offered Rate, being the rate at which a first-class bank offers to lend unsecured funds to another first class bank in London. Conversely, LIBID is a borrowing rate.

two-name paper), SCP remains the liability of the borrower only (i.e. one name paper). This is attractive to borrowers whose credit-standing is well known by lenders, since they do not have to incur an acceptance fee when borrowing, which reduces their cost of funds. As with bills, it is normal market practice for SCP to be issued on a discount basis, although provision has been made for such paper to bear a stated rate of interest.

Issue and trading conventions

SCP is generally issued via an uncommitted, continuously offered programme to which the issuer appoints one or more banks or securities houses to act as dealers[15]. These dealers, usually acting as principals, buy SCP from the issuer for distribution to end investors. Alternatively, the issuer may issue SCP directly to investors, or through a tender panel comprising banks and securities houses, who then bid competitively for the paper. However, activity in the SCP market is largely investor driven and issues are generally initiated in response to specific investor demand identified by dealers.

Since SCP is single name paper (i.e. the liability of the issuer only) information about the credit-worthiness of issuers is of considerable importance to the wide acceptability of such paper. In order to facilitate adequate disclosure of such information, therefore, SCP issuers must meet certain requirements. Specifically, in order to be eligible to issue SCP, the issuer must be:

(i) a company with net assets of not less than £25 million and either:

  (a) have shares or debt listed on the International Stock Exchange or on the Unlisted Securities market or

  (b) be incorporated in the UK or have shares or debt listed on a recognised overseas exchange, or

(ii) an overseas government or public authority with debt listed in London or on a recognised overseas exchange, or

[15]A detailed outline of the issue and trading conventions for SCP may be found in the British Bankers' Association booklet *Commercial Paper: London Market Guidelines* (September 1992).

(iii) in any other case, the issuer must have the SCP guaranteed by either

(a) a guarantor with net assets of not less than £25 million, or

(b) an institution authorised under the Banking Act 1987.

For reasons which will become clear in chapter 6, growth in SCP has remained slow, with the outstanding volume of such paper averaging only about £3 billion between 1987 and 1992. Furthermore, SCP tends to be purchased and held until maturity and an active secondary market has yet to develop for this instrument.

4.2 Transactions technology

Lenders, having contracted to lend cash in the sterling money market, need to be able to transfer this cash from their account to the account of the borrower. Where this loan takes the form of a purchase of a negotiable sterling security, borrowers must then also be able to transfer the security to the lender. In order to settle the deals which are struck between counterparties in the sterling money market, therefore, a number of sophisticated funds and securities transfer systems have been developed.

4.2.1 Funds Transfer

There are currently two large-value funds transfer systems in operation in the UK: an older, paper-based system known as the Town Clearing and a newer, electronic system known as CHAPS (Clearing House Automated Payments System). Both are administered by an umbrella organisation known as the Association for Payment Clearing Services (APACs)[16].

The Town Clearing

The Town Clearing is a paper-based system for the same-day clearing and settlement of large-value sterling *debit* instruments (i.e. cheques) within

[16]The information below was obtained from BIS (1993), and the document *APACs: A Description*, published by APACs in December 1993.

London. Operating in its present form since 1946, the Town Clearing is the modern descendant of the original clearing system which has existed in London for over two centuries. The system has fourteen settlement members, including the Bank of England, and settlement occurs across members' operational (reserve) accounts held at the Bank of England. The cut-off time for normal business for same day value is 15h00, although items may be presented to paying banks at the settlement centre until 15h50.

CHAPS

CHAPS was introduced in 1984 and constitutes a nationwide electronic system for sending irrevocable sterling *credit* transfers. CHAPS has the same settlement members as the Town Clearing and payments are likewise settled across settlement members' operational accounts at the Bank of England[17]. Being electronic, however, CHAPS payments are sent directly from one settlement member to another and there is no clearing house. Each settlement member, moreover, has many branches throughout the UK, which are all capable of initiating and receiving CHAPS payments. Furthermore, there are over 400 financial institutions which, through agency agreements with settlement members, can initiate and receive CHAPS payments[18]. From November 1993, the cut-off for normal business for same day value has been 15h10, although the system remains open for a further twenty minutes for payments between settlement members to settle transactions entered into for the purpose of achieving the minimum required closing balance in their operational accounts (see section 2.3).

The average daily clearing volume and value handled by these two systems is substantial, and in the twelve months before October 1993, totalled 42,000 items and £96 billion respectively (table 4.2). In order to encourage the use

[17]Settlement members are required to join both payment systems. Two further banks have recently been accepted as settlement members on CHAPS, which will bring the total number of settlement members to sixteen. Applicants must meet a minimum clearing value criterion and must agree to pay an entry fee and a share of the operating costs based upon the proportion of payment volume handled.

[18]These are naturally subject to the terms of the agency agreement, which will typically include limits, imposed by the settlement bank, on incoming/outgoing payments.

of CHAPS, APACS has progressively raised the minimum transaction value on the Town Clearing, which now stands at £0.5 million, whilst there is no minimum limit on CHAPS. This has brought the proportion of the total average daily clearings handled by the Town Clearing to less than 5%.

4.2.2 Securities Transfer

Another comparatively recent development in the sterling money market has been the introduction of an electronic securities transfer system by the Bank of England. Following its successful introduction of a computerised settlement system for gilt-edged securities in 1986, known as the Central Gilts Office, or CGO, the Bank introduced a similar system for money market securities in 1990, known as the Central Moneymarkets Office, or CMO. By lodging securities such as bills, CDs and commercial paper with the Bank of England, and logging them onto CMO, they can then be transferred electronically between settlement members until they mature. This has the advantage of avoiding the expenses and risks associated with the physical transfer of securities between counterparties [19]. A further advantage is that negotiable securities can be dematerialised (i.e. rendered paperless), thus reducing the costs of printing and issuing such securities[20]. Most sterling money market securities, therefore, are now logged onto CMO and in 1992 some 247,000 transactions were made through CMO, representing a nominal value of over £3000 billion.

There are currently about fifty financial institutions which are settlement members on CMO, comprising mainly the discount houses, clearing banks, merchant banks and other specialist institutions like Stock Exchange Money Brokers (SEMBs). These members have a book-entry account in their own name and, if they are not a settlement bank on CHAPS, must make arrange-

[19]These risks were well illustrated by the theft of nearly £300 million CDs from a City messenger in 1989.

[20]The Bill of Exchange Act, which governs the issue of bills, states that these must be in writing, which has prevented them from being dematerialised. Provision was made in the 1992 Finance Act 1992 to dematerialise CDs and commercial paper.

ments with such a settlement bank to make and receive payments on their behalf for instruments transferred from and to other CMO members. Over two hundred additional firms then participate indirectly in CMO through agency arrangements with CMO members. When a security is purchased, the seller initiates its transfer to the purchaser and this transfer is then accompanied by a corresponding payment instruction to transfer funds from the purchaser to the seller. It is noteworthy that this payment instruction is *not* "assured", and payment may thus be refused by the purchaser's settlement bank (although this has not, as yet, happened in practice).

4.2.3 Settlement

In chapter 3 it was argued that, in the presence of a money market, banks do not have to hold many reserves for settlement purposes. Instead, they must estimate their expected settlement obligations and then borrow or lend funds in the money market so as to offset this obligation. This is precisely what settlement banks in the UK attempt to do.

Recall that, in the UK, settlement banks face a daily reserve requirement, under which the Bank of England requires settlement members (i.e. clearing banks) to achieve at least a zero balance in their operational account by the end of each working day. In fact, clearing banks attempt to meet a small, positive *target* reserve balance each day. The level of these target balances is negotiated on a bilateral basis between each clearing bank and the Bank of England. But it is extremely small when compared with the typical forecast errors made by a settlement bank in estimating its daily expected settlement obligation. A large clearing bank, for example, will only have to achieve a target balance of about £30 million, but may easily experience an error in its forecast settlement obligation of £150 million. This would imply that clearing banks regularly sustain either a negative closing reserve balance, thus violating their reserve requirement, or are left holding excess reserves at the end of the day, thus incurring an opportunity cost, neither of which is observed in practice. How, then, is it possible for settlement banks to ensure that they achieve their target closing reserve balance each day, given

that their true reserve balance becomes known with certainty only *after* their settlement obligations are revealed (i.e. after markets are closed)?

The answer is that CHAPS settlement members have access to an ex-post money market of the type described in chapter 3. Upon being notified of their net CHAPS obligation at about 15h15, settlement banks become virtually certain of their final reserve positions, since the only outstanding obligations are those of the Town Clearing, which only closes at 15h50. Since the Town Clearing now accounts for less than 5% of large-value payments, however, these obligations are relatively small as compared with those over CHAPS. Once they are informed of their reserve positions, those banks which find that their CHAPS obligations would leave them with a negative closing reserve balance will want to borrow from those banks which find that they would be left with excess reserves. The money market, however, is already closed and thus some other avenue for redistributing reserves must be available. This redistribution is thus achieved via the following informal agreement between the eight largest settlement banks.

Each week one settlement bank is designated as the "first lender" (otherwise referred to as having "the chair")[21]. This status entitles the bank to lend any excess reserves it may have to those banks whose operational accounts would be left in deficit at the end of the day. If, after the bank which is "first lender" has lent all its excess reserves, there remain banks which still need to borrow reserves, then the settlement bank with the largest excess reserve balance is designated as the next lender. This lending process continues until no bank will end the day with a negative reserve balance.

The rate of interest applied to these ex-post loans, which are extended on an overnight basis, is calculated as the *highest* rate at which at least two banks lent overnight funds during trading in the market that day. These ex-post transactions then constitute the final payment instructions over the Town Clearing, which still remains open until 15h50. In this way, therefore, all settlement members will typically be able to achieve their target balances.

[21]This privilege rotates amongst settlement banks on an alphabetical basis, i.e. Barclays, then Citibank, then Lloyds, etc.

Of course, it may sometimes occur that aggregate reserves are insufficient, in aggregate, for these target balances to be achieved, in which case settlement banks must approach the Bank of England for overnight funds (see section 3).

4.2.4 Risks

The large daily payment values over CHAPS and the Town Clearing imply that settlement members incur substantial intraday exposures to one another. This comes about because both the Town Clearing and CHAPS are *net* settlement systems, which means that payment instructions are made *before* funds are actually transferred. It is possible, therefore, that some settlement member may not be able to settle their obligation as a result of having insufficient reserves – after borrowing – in their operational account. This eventuality would expose the *recipients* of these (non)payments to losses, since these banks will themselves have made outgoing payments – which they must honour since these are irrevocable – on the assumption that their incoming payments would be settled. In the extreme case, non-settlement by one member may culminate in systemic settlement failure, as other members in turn become unable to settle their obligations.

To decrease the risk of such a systemic settlement failure, and to limit settlement exposures, CHAPS has introduced a number of short-term measures which include the imposition of bilateral net receiver limits and an agreement between settlement members that the settlement exposures in the system are equal to the net bilateral obligations (rather than gross bilateral obligations, which are significantly larger). In addition, settlement members observe self-imposed sender limits. To reduce settlement risks in the longer term, however, APACs has agreed to place CHAPS on to a real-time gross settlement (RTGS) basis by the end of 1995[22]. This will mean that payment instructions over CHAPS will be settled continuously or, rather, as soon as they are made, i.e. funds will be transferred into a recipient's operational

[22]The architecture of this proposed system is explained in detail in Bank of England (1994) and will not be presented here.

account at virtually the same time that notice of payment is received. In contrast to existing settlement arrangements, where payments are made without reference to the existing reserve balance of the sender, RTGS will mean that settlement members are subjected to a *reserve-in-advance* constraint, since they will not be able to make outgoing payments until they actually have sufficient reserves in their operational account to cover this payment.

The Bank of England, anticipating the possibility that RTGS may slow payment traffic when aggregate reserves become too low (possibly even to the point where payment traffic ceases altogether, i.e. "gridlock"), has agreed to provide reserves to settlement members on an intraday basis. The reserves provided in this way will not attract an explicit cost, but will only be provided against approved collateral[23] such as bills and government securities. Effectively, then, payment arrangements on CHAPS will remain much the same under RTGS, except that receiving banks will no longer face any settlement exposures. The *intraday* availability of reserves will become more important, however, and there has been some speculation that RTGS might lead to the emergence of an intraday market for reserves, e.g. a money market for loans with a maturity of several hours.

4.3 Bank of England Operations

Until 1980 the Bank of England's money market operations were, as detailed in chapter 1, primarily confined to open market bill transactions and lending via its discount facility. The Bank's sole counterparties were the discount houses, who regularly approached the Bank for funds, usually by selling eligible bills – which then comprised mainly Treasury and local authority bills, but also included bank bills – to the Bank, or by borrowing on a secured basis with bills as collateral. These funds were obtained, until 1972, either at existing market discount rates or at Bank rate, which was the Bank's announced discount (lending) rate. The amount of funds obtained by discount

[23]Pledging collateral may well involve an opportunity cost, although this is difficult to measure in practice.

houses in these operations, however, was not revealed to the wider market.

In October 1972 Bank rate was replaced by Minimum Lending Rate (MLR), whose determination was supposedly "market-determined", in that the official rate was linked, by a formula, to the average rate established at the most recent weekly tender for three month Treasury bills. This formula, which meant that MLR was typically about $\frac{1}{2}\%$ above the Treasury bill rate, gave interest-rate expectations an important – but not necessarily dominant – role in the determination of official rates and resulted, on occasion, in market and official interest rates being "ratcheted" upwards. For this reason, provision to override the formula was made, under which the Bank could make special changes in MLR. These changes were generally announced at midday on Thursdays (before the Friday auction of Treasury bills), with the new official rate established in this way becoming effective immediately and the formula suspended until market rates had moved into line with official rates[24]. Dissatisfaction with this mechanism led, in May 1978, to this formula being abolished, with MLR again being determined, as Bank rate had, by administrative decision. Changes in this rate were normally announced at 12h30 on Thursday, with the new rate becoming effective immediately.

The late 1970s also saw a period of intense debate over the monetary control arrangements of the Bank which culminated, in March 1980, with the publication of an official (joint Bank and Treasury) discussion paper (known as a Green Paper) on monetary control, which even considered the possibility of money base control (HMSO 1980). In the event, the latter was not seen as a strong alternative to the Bank's existing interest rate instrument[25] and was also vigorously countered by the Bank (see Foot, Goodhart and Hotson 1980), which pointed to the operational problems associated with such a system (these were discussed in chapter 2). The Bank did, however, announce in

[24]In March 1977 the operation of the formula was modified: in cases where market rates moved downwards, the Bank now reserved the right either (i) to leave MLR unchanged, or (ii) to move it downwards but by less than would normally have been implied by the formula.

[25]Consider paragraph 4.15 of the Green Paper: "These difficulties are such that we doubt whether a monetary base control system ... would produce the desired results".

November 1980 that it was considering changes to its operating procedures in the money market. These changes were to

> "... place greater emphasis on open market operations and less on discount window lending. It has been decided that these operations should continue to be conducted in the bill markets ... and in large part through the existing intermediaries [i.e. discount houses] ... to whom discount window facilities would remain confined" (Bank of England, 1980, p.428).

The Bank also envisaged keeping short-term money market rates within an officially determined, but *unpublished*, band, thereby ceasing to announce an official rate in the form of MLR. This was to be achieved by inviting discount houses to offer bills and then choosing a minimum discount rate at which it would accept these offers. Two subsequent announcements by the Bank in 1981 (Bank of England 1981a, 1981b) spelled out these changes in more detail (these also included modifications to banks' liquidity requirements and also to the bill market, which are discussed in chapter 6) and announced that these would be introduced in August that year. However, by March the Bank had ceased to quote an official discount rate for bills and had asked its counterparties, the discount houses, to approach the Bank with offers to sell bills in two maturity bands, namely two weeks and one month. Furthermore, by July the Bank had begun to circulate, via the main press agencies, details of its daily operations, including its forecast of the money market shortage (see below), the amount of its bill transactions with the houses and the rates applied to these transactions.

The following sections now examine these operations, as have they been implemented since 1982, in greater detail.

4.3.1 The Money Market Shortage

Each working day, at 9h45, the Bank of England publishes on Reuters its forecast of the money market shortage (or surplus) and its estimates of the various factors which will give rise to this shortage. Recall, from chapter 3, that the money market shortage (M) may be defined as the net expected deviation in aggregate reserves (R) below its target value ($\bar{R}$). In the UK

this aggregate target is simply the sum of banks' required daily target reserve balances (referred to as bankers' balances), which total about £180 million.

Basically, there are four main factors which reduce bank reserves, and thereby create a money market shortage; in particular, banks' aggregate reserves are reduced:

(i) when banks' closing reserve balances on the previous day were below target: i.e $\bar{R} - R_{t-1} > 0$,

(ii) when the amount of currency in circulation, C_t, increases: i.e. $C_t > 0$,

(iii) when net disbursements from the government's account at the Bank of England, G_t, which reflect items like tax payments to the government and government expenditure, are negative: i.e. $G_t < 0$, and

(iv) when net receipts by the Bank of England, A_t, due to previous transactions (such as a sale of gilts or foreign exchange), or due to maturing private sector assets held by the Bank, are positive: i.e. $A_t > 0$.

The actual deviation in aggregate bankers balances (reserves) from their target level is then determined by the cumulative effect of these factors, i.e.

$$M \equiv (\bar{R} - R_t) = (\bar{R} - R_{t-1}) + C_t - G_t + A_t \,, \tag{4.4}$$

which, if the left hand side is positive, is labelled as a money market shortage. Over the last fifteen years, this magnitude has almost always been positive (i.e. a shortage) and has been fairly large (i.e. almost always in excess of £100 million, often in excess of £1000 million), usually due to either the influence of factors (iii) or (iv)[26]. In the absence of offsetting operations by the Bank of England, therefore, settlement banks will not only have systematically violated their reserve targets, but would also have been unable to meet their settlement obligations, since aggregate reserves will have been consistently *negative*. To avert this state of affairs, the Bank has provided reserves via two channels, namely open market operations and its discount facility.

[26]The government's policy of overfunding, which is described in chapter 6, played an important role in creating such large shortages.

4.3.2 Open Market Operations

In order to offset the impact of the money market shortage on banker's balances (i.e. "remove" the shortage), the Bank of England must purchase financial assets from the private sector in an amount very close to the estimated shortage. At the same time as it publishes its forecast, therefore, the Bank will usually invite its counterparties – primarily the discount houses, although this group now also includes Stock Exchange money brokers and gilt-edged market makers – to offer eligible bills for outright sale and/or for subsequent repurchase against same day cash settlement.

In the case of a *repurchase agreement* (or repo) involving bills, the Bank specifies the rate which will apply to the transaction, as well as the terminal date (or dates), which is typically between 14 days and one month. In the case of an *outright sale*, however, the Bank does not specify a rate and will purchase only those bills with an unexpired maturity of between 3 and 33 days[27]. In addition, the Bank specifies that it will not purchase bills which were issued less than seven days previously.

The minimum discount rate which the Bank applies to its outright bill purchases is known as the *stop rate*: only offers at or above this rate are accepted. If the amount of acceptable offers exceeds the amount of the shortage, the Bank will scale these offers so as not to provide reserves in excess of the shortage. The total amount of offers accepted, which is referred to as the amount of *money market assistance*, is then immediately published on Reuters.

Since the money market shortage is not accurately predictable, in particular because net reserve flows into government accounts at the Bank of England are not perfectly known, the Bank will often revise its announced

[27]In fact, the Bank splits the maturity of its bill purchases into four different maturity "bands": band 1 (3-14 days), band 2 (15-33 days) band 3 (34-63 days) and band 4 (64-91 days). Although the Bank has regularly bought band 3 and 4 bills in the past, it has not done so from September 1992 onwards. This is possibly because small "technical" adjustments in its band 3 and 4 dealing rates, which were made from time to time, became seen as a signal of an impending change in the Bank's official rate.

shortage. Given this uncertainty surrounding its forecasts of the shortage, the Bank will repeat its morning bill operations (if they occurred[28]) at noon and again at 14h00, usually with the same menu of transactions.

In the rare event of a money market surplus, the Bank invites its official counterparties and the clearing banks to bid for Treasury bills of one or more maturities for same day settlement. The Bank then accepts bids at the *lowest* rate offered, and thus the stop rate in this instance represents an unpublished selling rate rather than a buying rate.

In order to gain a clearer impression of the magnitudes involved in the Banks' open market operations, consider the period between January 1992 and October 1993 as an example. During this period, the average forecast shortage was just over £1 billion. This was revised almost every working day, with these revisions (which normally occur in increments of £50 million) ranging between +£450 million and −£550 million. Since upward revisions were more frequent than downward revisions, the average revision was marginally positive and was about +£50 million.

Although bill repos have become increasingly prominent in recent years, outright bill purchases still form the bulk of the Bank's transactions. In the period under consideration, for instance, outright purchases comprised 50% or more of the Bank's total daily transactions on more than 75% of the days on which the Bank undertook open market operations.

A breakdown of the timing of the Bank's open-market transactions during this year revealed that, on most occasions (over 90%), more than half of the published shortage would remain outstanding after the Bank's 9h45 operations[29]. Conversely, on most occasions (over 90%) less than half of the revised shortage would remain outstanding after the Bank's 14h00 operations. Against this basic pattern, however, were instances when the entire shortage was removed during the morning operations (approximately 12% of

[28]In the event of a shortage below approximately £700 million, the Bank will not normally conduct operations at 9h45.

[29]This frequency will be biased upwards by the fact that the Bank does not always conduct morning operations.

the time) and instances when the entire shortage still remained after the afternoon operations (approximately 6% of the time). It is difficult, therefore, to define a normal, or expected, pattern of timing for the Bank's open market operations (this issue is discussed again in chapter 7).

4.3.3 Discount window operations

In addition to these open-market transactions, which are concluded at 14h00, the Bank also makes same-day funds available to its official counterparties via a standing facility, which may be accessed from 14h45. Under this facility the discount houses may borrow an amount of up to twice their capital against collateral of eligible bills and gilt-edged securities with a maturity of under five years. Gilt-edged market makers and stock-exchange money brokers may also borrow under this facility, but their borrowing is limited to a lower proportion of their capital than the houses. In total, therefore, the amount of borrowing under the 14h45 facility is limited to approximately £750 million.

Requests to borrow under this standing facility which are submitted at 14h45 carry a unpublished charge, although this is known to be at – or marginally above – the Bank's (band 1) stop rate. The amount of 14h45 borrowing, known as late assistance, is then published. Further requests for late borrowing may be made after 14h45, but these transactions are no longer published and will attract increasing rates of interest the later they are submitted.

The breakdown of the timing of the Bank's operations for the period between January 1992 and October 1993 indicated that late assistance is granted with considerable frequency. However, whilst late assistance is often positive, the average amount of such assistance was typically less than one quarter of the average forecast shortage (i.e. less than £250 million). Nonetheless, on some occasions, roughly 6% of the time, the entire shortage has to be taken out via late assistance. This may pose problems when the shortage is large (i.e. in excess of £1 billion), since the capacity of the Bank's official counterparties to borrow using late assistance has an upper bound of about £700 million. Effectively, therefore, the Bank's official counterparties

will then be borrowing funds from the Bank in excess of the stop rate. This begs a question: why is it ever the case that there is still a large shortage outstanding at 14h00 when there have been *three* opportunities for financial institutions to sell bills to (obtain reserves from) the Bank of England? The answer, it seems, is that institutions may face an incentive to delay the sale of bills to the Bank, and the next three chapters will each present an element of this incentive.

4.3.4 Money market interest rates

In chapter 3 it was shown how a money market shortage was sufficient to drive short-term money market interest rates to equality with the lending rate of the central bank. Does this happen in practice? Evidence presented in Llewellyn (1990) as well as in figure 4.5, suggest an answer in the affirmative.

Figure 4.4 shows the official discount rate of the Bank of England, which between 1972 and 1981 was known as MLR and was announced by the Bank. Since 1981, when the Bank's operating procedures were modified, this rate has been known as the stop rate and is an unpublished minimum discount rate. Given that the Bank operates in 4 (latterly only 2) maturity bands, it may apply a marginally different stop rate in each of these bands. Usually, however, the Bank chooses these rates so as to make the corresponding *yields* in each maturity band the same[30] and hence figure 4.4. shows only the band 1 stop rate. The current level of this rate is common knowledge amongst market participants, since the Bank publishes the rate at which it conducts its open-market operations each day. It is apparent from figure 4.4 that the Bank's official rate does *not* fluctuate on a daily basis and is held constant by the Bank until economic conditions warrant either a higher or a lower rate. These changes in the Bank's official rate are typically made in steps of $\frac{1}{2}\%$, although larger changes (1%) and smaller changes ($\frac{1}{4}\%$) are also observed.

Figure 4.5 plots the spread between the Bank's official discount rate (on a yield basis) and one month LIBOR, and indicates that the latter rate has,

[30]This can be seen, from equation (4.2), to be consistent with *lower* discount rates in the longer maturity bands.

between 1975 and 1993, usually remained within 1 percentage point or less of the stop rate. It is noteworthy that the spread between LIBOR and the stop rate becomes considerably narrower *after* 1981, when the Bank adopted its new operating procedures, suggesting that these permitted the Bank to "smooth" money market interest rates to a greater extent than before. Nevertheless, this smoothing is not perfect and, occasionally, sustained deviations of market rates from the Bank's official rate did occur. The following chapter therefore examines the transmission of official rates to market rates in greater detail.

4.4 Two recent developments

Since 1825 the Bank has not normally dealt directly with the clearing banks, although these banks do have a general invitation to sell Treasury bills and local authority bills directly to the Bank when open-market operations are being conducted. Since the stock of Treasury bills held by the clearing banks has been small in recent decades, and the rate at which they can sell these bills to the Bank has been marginally above the stop rate, this has occurred relatively infrequently.

In addition to these bill purchases the Bank has occasionally been willing to purchase gilt-edged stocks directly from banks for resale to them at a future date in the event of particularly large money market shortages[31]. On such occasions, the Bank has announced in advance that it was offering a temporary facility to banks and announced the terms of this facility.

Apart from these channels, then, the clearing banks would seem to have been excluded from any direct role in the operations of the Bank. This is not accidental, according to Coleby (1982):

> "An alternative means of putting cash into the system ... would have been for the Bank to lend directly into the interbank market. That method did not seem, on examination, to offer a satisfactory means of determining

[31] Recall, from chapter 1, that repurchase agreements against consols were also the Bank's earliest open-market operations.

> interest rates, because the highly concentrated structure of sterling deposit banking would have confronted the Bank with only a handful of large takers of funds - the clearing banks. Rate determination in those circumstances would have come close to a bilateral haggle. So the decision was taken to continue to provide cash through bill operations, dealing largely with the discount houses" (p.214).

When adopting its new operating arrangement in 1982, therefore, the Bank made it clear that it did not wish to deal directly with the clearing banks. Choosing to operate in the manner that it does, however, has not been sufficient to avoid this outcome, since the clearing banks now own most of the non-official stock of eligible bills. What is more, recent changes in the Bank's operations – namely the introduction of a permanent gilt repo facility – indicate that the clearing banks have begun to play an explicit role in these operations, which is likely to increase in the future.

4.4.1 Who owns sterling bills?

During the 1970s the Bank of England imposed statutory liquid asset requirements at a common level across all UK banks, requiring them to hold stipulated reserve assets in proportion to a particular measure of their sterling liabilities[32]. The reserve assets which had to be held to satisfy these requirements included eligible bills, gilts with a maturity of under one year and secured deposits with the discount houses. Designating the latter as reserve assets was naturally favourable to discount houses, since it ensured a sizable flow of short-term funds to the houses, for which they did not have to compete, and allowed them to maintain large portfolios of bills and other short-term assets. The position of the houses as the major holders of eligible bills was further reinforced by the structure of liquidity requirements, which imposed a *ceiling* on the amount of bills which banks could hold in order to satisfy their requirement.

[32]Prior to this, banks had "voluntarily" held a stable, sizable proportion of their assets in the form of liquid assets; see Bank of England (1962).

Amongst the changes in the Bank's operations in 1981 was the abandonment of these common liquidity requirements and their replacement by liquidity requirements which were negotiated on a bilateral basis between individual institutions and the Bank of England (see Bank of England 1981c). Thus although the definition of reserve assets didn't change, institutions now had greater scope for holding liquidity in a form which they preferred, i.e. bills rather than secured money. To minimise the effects of these changes on the position of the houses and to bolster their role as official counterparties, the Bank stipulated that all eligible accepting institutions had to maintain an average equivalent of 6% of their eligible sterling liabilities in the form of secured money with the discount houses (or with the Bank's other counterparties). This requirement, known as "club money", again guaranteed a sizeable flow of funds to the houses and maintained their position as the largest non-official holders of bills.

This position reached a turning point in 1987, however, when the "club money" requirement was abolished, leaving banks free to choose the proportion of their liquid assets which they wished to keep in the form of secured money with the discount houses. Although many banks chose to continue placing "callable" funds with the houses, the larger clearing banks chose, instead, to hold a greater proportion of their short-term liquid assets in the form of eligible bank bills. In the last few years this trend has been maintained, so that the clearing banks, as a group, are now the largest non-official holders of bills and now regularly own over 90% of non-official holdings. Within this group, the "big four" clearing banks are the major holders and account for over 80% of the groups' holdings. Note that this is a fairly recent development, with the major increase in clearing bank holdings occurring during 1990 (see chapter 6).

Discount houses have thus had to contend with a marked reduction in their traditional sources of funds, and a concomitant decrease in their assets, particularly eligible bills. This has not been without consequence for their role as the official counterparties in the Bank's operations. Instead of being active principals in these operations, which they have done for many years, discount houses are now virtually passive as counterparties, passing bills through to

the Bank of England on behalf of the larger clearing banks (for details see Schnadt 1994a). Thus it is the clearing banks who have now become the *de facto* counterparties to the Bank. And, although this has not resulted in the determination of short-term interest rates via a "bilateral haggle", it has had adverse consequences for the behaviour of money market interest rates, in particular very short-term rates. This issue, however, is addressed further in chapters 6 and 7.

4.4.2 Gilt repo

Although originally intended as a temporary measure, to be used only in the event of large money market shortages, the Bank's gilt repo facility has recently become a *permanent* feature of its open market operations. This development has its origins in the recent ERM crisis during which the Bank – on September 16, 1992 – purchased vast amounts of sterling in support of the currency. Naturally, this generated large money market shortages – so large, in fact, that a continuation of the Bank's normal bill purchases would have resulted in the Bank buying, on a single day, almost all eligible bills outstanding[33]. Consequently, the Bank announced that it would use its temporary facility to deal with this temporarily large shortage. Under this facility, all banks with eligible liabilities over £1.5 billion and the five largest building societies, were invited to offer gilt-edged securities to the Bank for subsequent repurchase one month later. This facility was subsequently rolled-over, usually on a monthly basis, each time that it matured, until the Bank proposed, in December 1993, to make this a permanent feature of its operations (see Bank of England 1993).

In particular, the Bank proposed that the repo facility be offered on a regular bimonthly timetable – on the Thursday following the first and third Monday of every month – with funds being provided either for two weeks or for one month, at the discretion of the counterparty[34]. These funds could be

[33]For an account of this episode, see *Bank of England Quarterly Bulletin*, November 1992, p.395.

[34]As in a conventional repurchase agreement, counterparties have to put up a cash

applied for on the relevant Wednesday between 10h00 and 11h45 – once the Bank had confirmed the periods for which funds would be available and had announced the rates of interest which would apply – by stating the amount of funds applied for at each maturity. The Bank would, by 12h45, announce on Reuters the total amount of funds provided at each maturity, and, if applications had been scaled back, the proportion by which applications at each maturity had been scaled.

The key innovation ushered in by these gilt repos, however, was the fact that the counterparties in these transactions could now be *any* bank or building society operating in the UK. For the first time, therefore, since 1825, the Bank had again elected to deal directly with banks. Should these repo facilities be extended in the future, as they might, to replace the Bank's daily bill operations, then there will no longer be a unique role for the Bank's official counterparties, the discount houses. Of course, the houses may retain their role as specialist financial institutions, and possibly even as market makers in money market securities. But their central role, as the conduit through which the wider banking system obtained reserves from the Bank of England, will have ceased, and with it one of the oldest traditions in the sterling money market.

---

margin, in this case 2.5% of the consideration, to guard against adverse movements in the price of the securities.

4.5 Figures and tables

Figure 4.1
The price of money market instruments:
Bills versus CDs

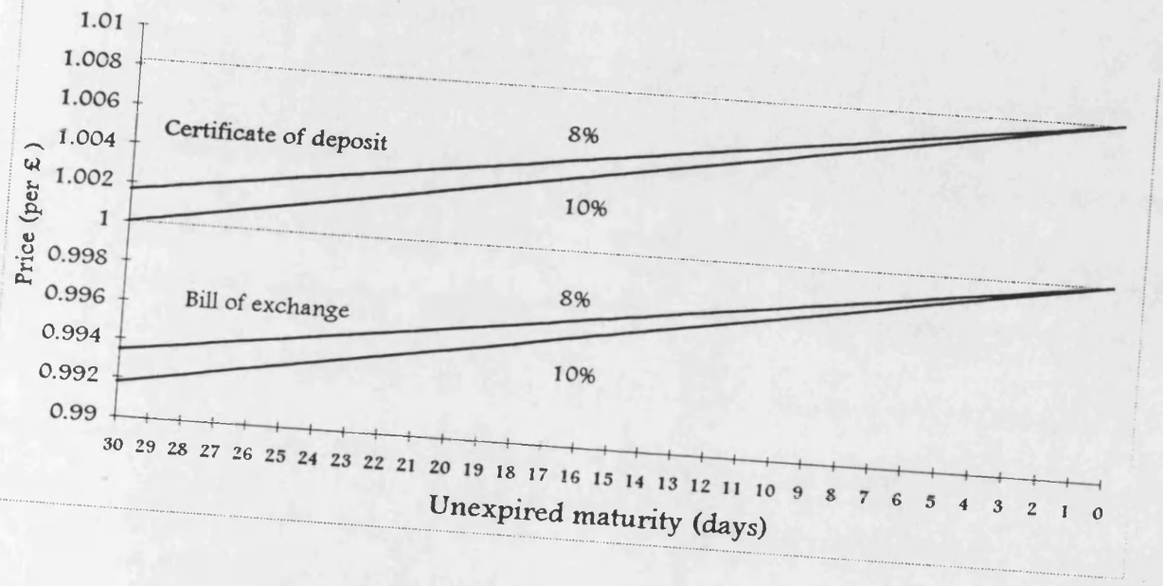


Figure 4.2
Bill yield – LIBOR (one month)
1988-1993

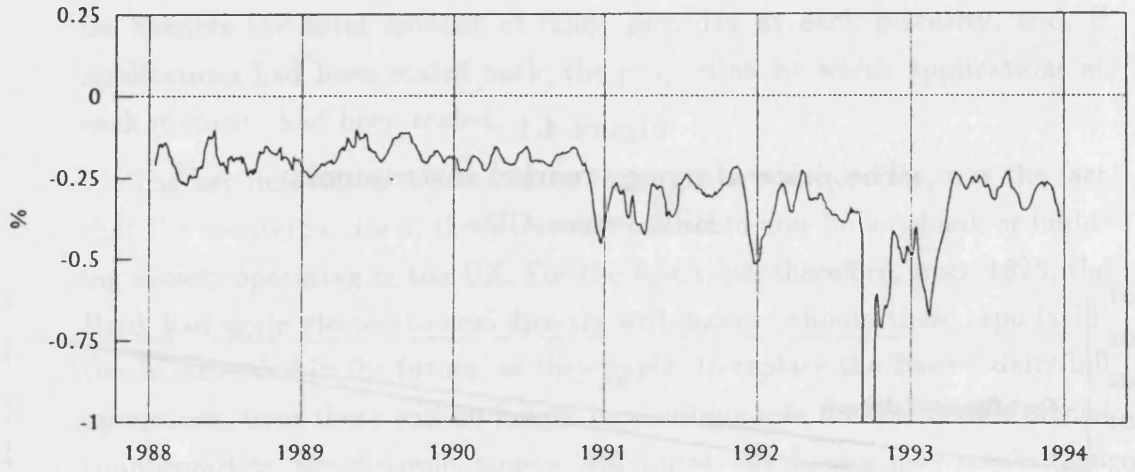


Figure 4.3
CD yield – LIBOR (one month)
1988-1993

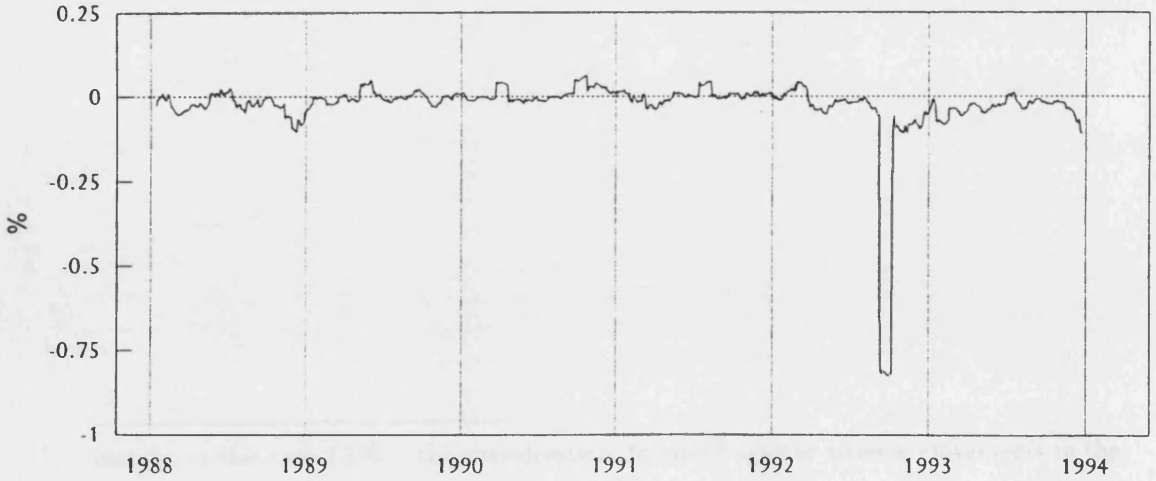


Figure 4.4
Official discount rate
1975-1993

Figure 4.5
Official rate – one month LIBOR
1975-1993

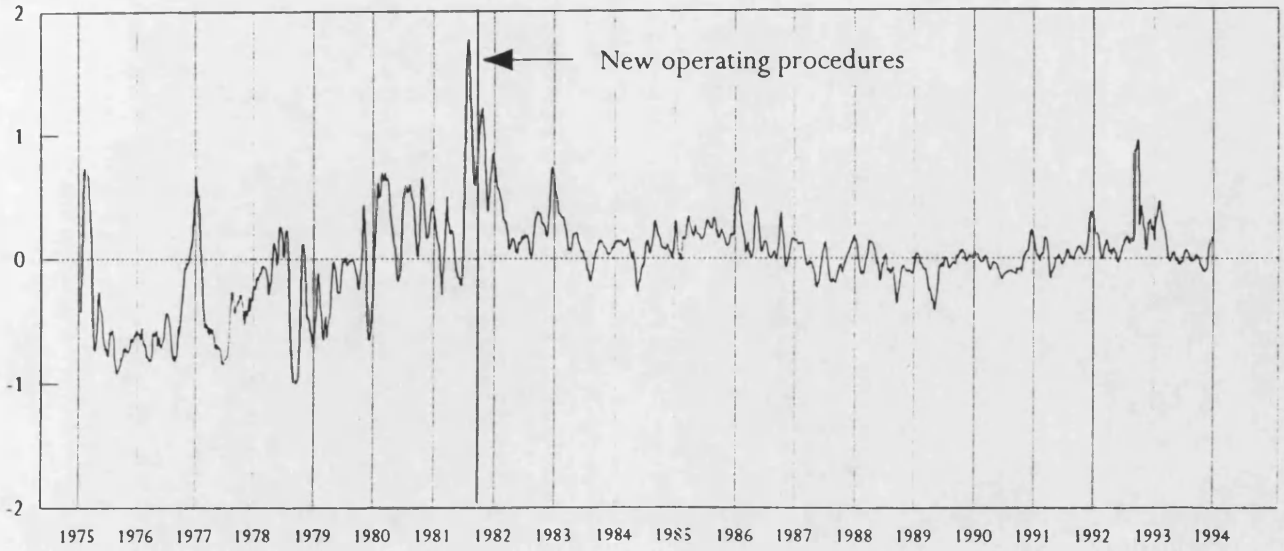


Table 4.1
Sterling money market instruments outstanding (1993)

Instrument	**Total** (£bn)	**Proportion** (%)
Treasury bills	4.85	1.7
Bank bills	20.84	7.4
Certificates of Deposit	60.15	21.4
Commercial Paper	4.00	1.4
Secured loans	7.08	2.5
Unsecured loans	183.84	65.5

Source: Bank of England Quarterly Bulletin (February 1994), and CSO Financial Statistics (April 1994).

Table 4.2
Payment statistics - average daily clearings (1993)

System	**Daily Value** (£bn)	**Daily Volume** (000's)
CHAPS	91.5	42.0
Town Clearing	4.5	0.3

Source: APACs (December 1993).

Chapter 5

The Term Structure of Money Market Interest Rates

In order to insulate money market interest rates from the fluctuations which would accompany variations in banks' aggregate reserves, the Bank of England undertakes to offset these aggregate reserve variations on a daily basis. The previous chapter has shown that the Bank does this by typically standing ready to lend reserves at its chosen stop rate, which is then transmitted to other sterling money market interest rates. The precise manner of this transmission, however, has rarely been studied, although Llewellyn (1990) and Flemming and Barr (1989) are notable exceptions. From the description in the previous section it was seen that the Bank typically lends across a range of maturities, operating as it does by inviting institutions to sell bills in a number of maturity bands. At which maturity, then, does the Bank's influence lie? And, more importantly, how does the term structure of interest rates in the money market behave when it is expected that the stop rate will change?

This chapter seeks to explore the relationship between the Bank's stop rate and money market interest rates under the specific operating procedures adopted by the Bank, taking as its point of departure the intuition that institutions are not indifferent as to the maturity of the bills which they sell to the Bank (i.e. the maturity of the funds which they obtain from the Bank).

For example, it seems reasonable that institutions would prefer to sell longer dated bills to the Bank when they are expecting an increase in the stop rate. This is because there exists a chance that the funds so obtained can be relent subsequently at a higher rate. But this means that the cost of short-term funds will not, in general, be equal to the stop rate, and will certainly not be independent of these maturity choices. What is more, short-term interest rates may, temporarily, even move in the *opposite* direction to the Bank's intended stop rate change. This may frustrate the Bank's intentions, causing it to move its stop rate sooner than it otherwise might have.

In order to address these issues, a model of money market interest rates is presented in section 2, in which expectations about an impending stop rate are summarised by a single parameter in a recursive structure[1]. This simplified framework allows an analytic expression for the term structure to be derived, from which it is possible to evaluate the impact of an expected stop rate change on money market interest rates[2]. In section 3, after considering the effects of relaxing some of the assumptions underpinning the model, daily money market interest rate data is examined for the behaviour predicted by the model. Section 4 then discusses some of the consequences of this rate behaviour and suggests how it may be eliminated. Finally, section 5 contains some concluding remarks.

[1]This chapter is a revised version of a paper written with John Whittaker, which is forthcoming in *The Manchester School.* See Schnadt and Whittaker (1994b).

[2]For a model of the term structure which investigates this for the US case see Balduzzi *et al* (1993).

5.1 A model of money market interest rates

5.1.1 Assumptions

The common situation in the sterling money market on any working day is a money market shortage, which is typically of the order of £1 billion. To offset this shortage, the Bank of England invites its designated counterparties to sell eligible bills for subsequent repurchase (i.e. via a repurchase agreement) or on an outright basis. On most working days, therefore, institutions can offer bills to the Bank under repurchase agreement, with the maturity of these deals typically chosen by the Bank to be somewhere between 14 and 33 days. Simultaneously, institutions can offer bills to the Bank for outright sale, although their unexpired maturity is subject to a maximum which is typically set by the Bank at 33 days. On those days (currently rare) on which there is a cash surplus, the Bank invites offers to buy Treasury bills to enable banks to avoid carrying excess reserve balances which earn zero interest.

Since the Bank always meets demands for reserves (or absorbs surpluses of reserves) in these ways, it is the rates of interest at which it is prepared to undertake these deals which determine sterling money market interest rates (see chapter 3). However, the relationship between official lending rates and market rates is not straightforward, due to two characteristics of the Bank's current operating procedures. First, the Bank always offers to make loans for maturities which exceed one day and, secondly, market participants can typically exercise discretion over this maturity.

In order to model the implications of this discretion, the following simplifying assumptions are made (the effects of relaxing some of these assumptions is discussed in section 2). First, the aggregate net reserve position in the money market each day is always a shortage, which is relieved by the Bank buying bills (either for subsequent resale or an outright basis): that is, the Bank always lends reserves. The rate at which it lends this cash is known as the stop rate, denoted by s_t, which is an unannounced minimum lending rate. Second, the asset portfolios of the Bank's dealing counterparties always contain sufficient bills of all relevant maturities, so that their choice of bills to

sell at any permitted maturity is not restricted by a lack of available paper. Third, participants in the money market are competitive, immune to default risk and hold the same expectations at any time; this allows the existence of a unique market rate of interest for each type and maturity of loan contract. Finally, money market participants are risk neutral, implying that the expectations hypothesis of the term structure applies exactly[3].

Under these simplifying assumptions, money market institutions collectively may be treated as a single unit (i.e. a representative bank) whose only decision variable is the maturity at which to borrow from the Bank. Profit maximisation implies that this maturity is chosen so as to minimise the interest cost of obtaining reserves each day from the Bank, given expectations of the future path of the stop rate and the assumption that optimisation will be conducted in the same way in the future. Individually, institutions price-take both the market cost of funds (represented by the money market term structure) and the cost of funds at the Bank (stop rate), and will exploit any arbitrage opportunities. This behaviour then implies that points on the term structure (or yield curve) reflect the cheapest expected cost of borrowing from the Bank, given current and expected future spot rates.

When news causes expectations of the future path of the spot rate to be revised, the term structure instantly assumes the shape which is consistent with the above optimising behaviour under the new information set. The behaviour of money market interest rates under these assumptions is now studied.

5.1.2 The general framework

In practice, money market loan contracts are always for an integral number of days, with one day being the shortest contractible maturity. It is therefore appropriate to discuss the determination of the market rate of interest for one day loans[4], which is denoted by i_t. Under the assumptions of homogenous

[3]Cuthbertson (1992) and Hurn, Moody and Muscatelli (1993) find empirical support for the expectations hypothesis in the term structure of UK money market rates.

[4]These are referred to as overnight loans, and are discussed in detail in chapter 7.

expectations and risk neutrality, the relationship between one day rates and longer rates is then given by the expectations hypothesis,

$$i_t^m = \frac{1}{m} \sum_{j=0}^{m-1} E_t (i_{t+j}) , \tag{5.1}$$

where i_t^m is the m day rate on day t (with interest understood to be compounded daily) and the E_t operator denotes the mean of the probability distribution of its argument given information available at time t.

An institution seeking to borrow one day funds from the Bank of England on day t can do so in number of ways. In particular, it can sell a bill whose unexpired maturity is $n \in \{\underline{n}, \ldots, \bar{n}\}$, where $\underline{n}$ and $\bar{n}$ are, respectively the shortest and the longest maturity which the Bank will accept, and then lending the borrowed funds in the money market from day $t+1$ to day $t+n$. Using (5.1), the expected cost of one-day funds via these alternative routes, denoted as $c_t(n)$, can be expressed as a function of the stop rate and future one day rates:

$$c_t(n) = ns_t - \sum_{j=1}^{n-1} E_t (i_{t+j}) , \; n \in \{\underline{n}, \ldots, \bar{n}\}, \tag{5.2}$$

where s_t is the Bank's official stop rate for lending (via outright purchase or repurchase of bills) at all acceptable maturities on day t. Thus, if the Bank bought bills which had only one day left to run, i.e. $\underline{n} = 1$, institutions could obtain one day funds directly, at a cost equal to the stop rate s_t. If institutions chose to obtain one day funds by selling a longer dated bill to the Bank, i.e. $n > 1$, then their cost of one day money would depend not only on the current stop rate, but also on the rate(s) at which they could relend funds for the remaining $n - 1$ days.

As discussed in chapter 3, money market shortages, together with competition between market participants, will imply that the market one day rate is driven to equality with the cost of obtaining one day funds from the Bank:

$$i_t(n) = c_t(n) , \tag{5.3}$$

which will, in this case, be a function of the maturity of the paper which is sold to the Bank and over which market participants have discretion. Profit-maximisation on each day t implies that market agents choose the maturity

of their borrowing from the Bank so as to minimise $c_t(n)$, with the result that the one day rate is determined by

$$i_t^* = \min \{ c_t(n) \mid n = \underline{n}, \ldots, \bar{n} \} . \tag{5.4}$$

The current one day rate i_t^* thus depends, in general, upon expected one day rates $E_t(i_{t+j})$, $j > 0$, which depend, in turn, upon expected future stop rates. The next section thus derives a simplified version of the manner in which expectations of future stop rates are formed in practice.

5.1.3 Expectations of stop rate changes

It is evident, from the discussion in chapter 4, that when the Bank changes its stop rate, this is always done in discrete steps, invariably of $\frac{1}{2}$ % or 1% up or down, with the stop rate being constant in between these changes. Whilst the Bank does not preannounce rate changes, it does provide signals as to the general direction of its future policy[5]. Moreover, money market participants are aware of how the Bank has responded in the past to the various economic data which it considers relevant. Based on this information, the likely form of expectations at any moment is that market participants expect, say, a $\frac{1}{2}$ % change in the stop rate but are unsure of the date of this change. Expectations of this date should, however, become firmer as the event approaches.

To simplify the problem at hand, therefore, it is assumed that on any day t market participants expect a change in the stop rate of a *known* magnitude with the rate staying constant thereafter[6]. They do not, however, know the *timing* of this change, so it is necessary to specify the probability of the change in the stop rate occurring on each day $t + j$, $j > 0$.

---

[5]Such signals may be explicit policy announcements, or more subtle messages transmitted via the Bank's dealing patterns in the market. Market participants are also aware of the general form of the Bank's *reaction function* and will thus come to expect an interest rate reaction depending on the behaviour of certain economic indicators.

[6]This is probably a poor representation of expectations when the stop rate is changing frequently. An alternative representation, which yields qualitatively the same results, is that future rate changes are expected to follow a Poisson process.

One way to proceed would be to make the problem finite, by assuming that agents believe the change will definitely occur by some date $t + j = \tau$, and then assigning probabilities $p_t(t+j)$ as viewed from day t that the change will take place on each day in the interval. Working backwards in time from τ, and assuming optimal choices of maturity n on each day, it would be possible to solve for r_t using dynamic programming techniques.

There are two drawbacks to this approach. In the first place, the general solution quickly becomes intractable as the number of days $\tau - t$ on which the rise may occur increases. In the second place, it is not plausible that agents should assume the change in the stop rate *must* occur before some fixed date τ. This would imply agents believe now that if, on the future day $\tau - 1$, the change has still not occurred, they will then believe it must occur on the following day.

To overcome these drawbacks, a model in which probabilities of a stop rate change have a recursive structure is proposed. This leaves undefined the time horizon for the occurrence of the rate change and admits a simple analytical formulation of the term structure of interest rates.

Let the current stop rate have the value s_0, and let $p_t(t + 1)$ be the probability that this will change to s_1 on the next day $t + 1$. Now assume that this probability has the same value on all the following days $t + j$ on which the change has not occurred, i.e. $p_{t+j}(t+j+1) = p_t(t+1) = \rho$, where $0 < \rho \leq 1$. In other words, if the stop rate change has *not* occurred by day $t + j$, i.e. $s_{t+j} = s_0$, then

$$s_{t+j+1} = \begin{cases} s_1 & \text{with probability } \rho \\ s_0 & \text{with probability } 1 - \rho\,. \end{cases}$$

Alternatively, if the stop rate change *has* occurred by day $t+j$, i.e. $s_{t+j} = s_1$, then by assumption $s_{t+j+k} = s_1, k = 1, \ldots, \infty$.

With this structure, ρ is the only variable which measures the probability of a stop rate change on future days. Loosely expressed, a higher value of ρ signifies an increased probability that the change will occur sooner, rather than later: if $\rho = 1$ then market participants are certain that the stop rate will change on the following day.

5.1.4 The term structure

Consider, first, the situation *after* the stop rate has changed. By assumption, market participants now know for certain that the stop rate will remain constant at s_1[7]. Given continuing cash shortages, the cost of borrowing one-day funds from the Bank on the day of the stop rate change is, from (5.2), $c_t = s_1$, and is independent of the maturity of paper sold to the Bank. This is true for all future days as well, which implies that, in the absence of any expected change in the Bank's stop rate, $i_t = s_1$ for all t. From (5.1) it is then clear that the term structure is flat, i.e. $i_t^m = s_1$, $m > 1$.

Now consider the situation before the stop rate change has occurred, i.e. $\rho > 0$. Given the above structure of probabilities the optimisation problem on each day $t+j$ is the same as on day t for as long as the stop rate change has not occurred. Hence $i_{t+j} = i_t$ if $s_{t+j} = s_0$, otherwise $i_{t+j} = s_1$. The probability on day t that the stop rate rise will not occur during the following j days is $(1-\rho)^j$, and thus

$$E_t\,(\,i_{t+j}\,) = (1-\rho)^j\, i_t + (1\,-\,(1-\rho)^j)\, s_1. \tag{5.5}$$

Substituting this into (5.2), with $s_t = s_0$, gives the expected cost of obtaining one day funds from the Bank of England, and hence the one day market rate, as

$$i_t(n) = c_t(n) = s_1 - (s_1 - s_0)\left(\frac{\rho\, n}{1-(1-\rho)^n}\right), \tag{5.6}$$

where $n \in \{\underline{n}, \ldots, \bar{n}\}$. Recall that the one day rate is then derived, from (5.4), by market participants choosing that maturity n of borrowing from the Bank which minimises (5.6).

Inspection of (5.6) indicates that $i_t(n)$ is a monotonic function of n, which is increasing when $s_1 < s_0$ and decreasing when $s_1 > s_0$. This ensures that there will be unique, optimal choice of n, denoted by n^*, which then depends

[7]Essentially, this assumes that ρ takes on a value of zero once the stop rate has changed. Although patently unrealistic, this assumption is of limited consequence to the analysis since it is the behaviour of market participants *before* an imminent change in the stop rate which is of primary interest.

only upon the *direction* of the stop rate change, i.e. on the value of s_1 relative to s_0.

Before investigating the optimal choice of maturity in more detail, note that the term structure can be expressed as a function of n^*. Substituting (5.6) into (5.5) gives the expected one day rate as

$$E_t(i_{t+j}) = s_1 - (s_1 - s_0)\left(\frac{\rho\, n^*(1-\rho)^j}{1-(1-\rho)^{n^*}}\right),$$

which, from (5.1), gives an expression for the term structure:

$$i_t^m(n^*) = s_1 - (s_1 - s_0)\frac{n^*}{m}\left(\frac{1-(1-\rho)^m}{1-(1-\rho)^{n^*}}\right). \qquad (5.7)$$

The shape of this term structure depends on the unexpired maturity of the bills which institutions will, at any time t, choose to sell to the Bank of England. The term structure equation (5.7) indicates that, in general, market rates will *not* be equal to the Bank's current stop rate. Indeed, the only market rate which will remain equal to the Bank's stop rate is the m-day rate whose maturity m corresponds to the maturity of the bills n^* which are sold to the Bank, i.e.

$$i_t^{n^*}(n^*) = s_0.$$

It will now be shown via some examples that the behaviour of all other market rates falls into one of two categories, depending on their maturity. *Market rates for which $m > n^*$ will correctly anticipate the new expected stop rate.* Thus, if the stop rate is expected to rise these rates rise above the current stop rate and if the stop is expected to fall these rates fall below the current stop. *Money market rates for which $m < n^*$ will do just the reverse, i.e. they will move in the opposite direction to the expected stop rate change.* In general, therefore, the expectation of a stop rate change causes the yield curve to "pivot" about the maturity $m = n^*$, with the direction of this pivoting being determined by the direction of the stop rate change (i.e. the sign of $s_1 - s_0$) and the extent of the pivoting being determined by the strength of expectation (i.e. the magnitude of ρ).

5.1.5 A numerical example

The behaviour of market rates in this model is now illustrated for the case in which agents expect a $\frac{1}{2}$ % change in the stop rate from $s_0 = 10\%$. Values for $\underline{n}$ and $\bar{n}$ are assumed, for the purposes of the example, to be 3 days and 14 days respectively[8]. Using these parameters, figure 5.1 plots how the cost of one-day funds, given by equation (5.6), varies with the unexpired maturity n of bills sold to the Bank, given some assumed probability ρ of the stop rate change.

Consider, first, an expected fall in the stop rate, i.e. $s_1 - s_0 = -\frac{1}{2}$. Figure 5.1 verifies that the cost of one day funds (5.6) is then monotonically increasing[9] in n. Given that institutions will always want to obtain the cheapest possible one day funds, they will obviously prefer to sell bills to the Bank which have the shortest permissible maturity, hence $n^* = 3$. For this choice of maturity, the one day rate is

$$i_t^* = 9.5 + \frac{3}{2}\left(\frac{\rho}{1-(1-\rho)^3}\right),$$

which is *above* the current stop rate for $\rho > 0$. Intuitively, this is because institutions are effectively being locked into borrowing at the existing stop rate when they expect this rate to fall in the near future. Thus they are borrowing funds at a higher rate than they expect, on average, to earn when they subsequently relend these funds[10]. From (5.7), the term structure is then

$$i_t^m = 9.5 + \frac{3}{2m}\left(\frac{1-(1-\rho)^m}{1-(1-\rho)^3}\right).$$

This is plotted in figure 5.2 (for two values of ρ). The figure indicates that the term structure for $m < 3$ lies *above* the current stop rate, and very

[8]The choice of $\underline{n} = 3$ days is in keeping with the Bank's practice of never buying bills with an unexpired maturity of less than three days. The choice of $\bar{n} = 14$ days is made since this is the cut-off maturity for band 1 bills.

[9]Higher values of ρ increase the slope of this cost function.

[10]This implicit penalty is clearly higher the longer the unexpired maturity of the bills which are sold to the Bank, which is why the cost function is increasing in n, and also why institutions will choose to sell bills with the shortest permissible maturity ($\underline{n} = 3$).

short-term interest rates thus move in the opposite direction to the expected stop rate change. Conversely, m-day rates for $m > 3$ lie below the current stop rate, approaching the new stop rate asymptotically with higher values of m. Market interest rates at all maturities $m > 3$ thus anticipate the expected fall in the stop rate in a well behaved manner. Notice that, when ρ rises, indicating a shortening of the expected interval to the stop rate fall, the term structure "pivots" in a clockwise manner, with the pivotal maturity being at $m = 3$.

In the case of a stop rate rise, i.e. $s_1 - s_0 = \frac{1}{2}$, figure 5.1 shows that the cost of one day funds, equation (5.6), is monotonically decreasing[11] in n. The optimal strategy for money market institutions, who wish to minimise their cost of one day funds, is thus to borrow from the Bank of England for the longest possible period. Hence $n^* = 14$ in this case. For this choice of unexpired maturity the one day rate is

$$i_t^* = 10.5 - 7\left(\frac{\rho}{1-(1-\rho)^{14}}\right),$$

which is *below* the current stop rate for $\rho > 0$. Now, contrary to the case of the stop rate fall, institutions are faced with an opportunity to "lock in" funds at the existing stop rate which they expect, on average, will earn a higher return due to the impending stop rate rise. Their implied cost of one day funds is thus lower than the stop rate[12]. From (5.7), the term structure is then:

$$i_t^m = 10.5 - \frac{7}{m}\left(\frac{1-(1-\rho)^m}{1-(1-\rho)^{14}}\right).$$

This is plotted in figure 5.3 (for two different values of ρ), which shows that the term structure for $m < 14$ lies *below* the current stop rate. Longer term rates, for $m > 14$, anticipate the change in the stop rate and lie above the current stop rate. Now, when ρ rises, indicating a strengthening of ex-

[11]As with the case of the stop rate rise, the slope of the cost function is steeper the higher is the probability ρ.

[12]This profitable opportunity naturally increases with the unexpired maturity of the bills sold to the Bank, which is why the cost of one day funds falls with n, and also why institutions will want to sell bills with the longest permissible unexpired maturity ($\bar{n} = 14$).

pectations about the stop rate increase, the term structure "pivots" in an anti-clockwise direction, with the pivotal maturity being at $m = 14$.

5.1.6 Asymmetric pivoting

An asymmetry in the behaviour of the term structure of money market interest rates between an expected fall and an expected rise in the stop rate is clearly evident from figures 5.2 and 5.3. In particular, the maturity about which the term structure "pivots" in each scenario is different. This comes about because institutions can exercise *discretion* over the maturity at which they borrow from the Bank of England, i.e. because $\underline{n} < \bar{n}$. This discretion is irrelevant when the stop rate is not expected to change, since the cost of one days funds from the Bank is the same regardless of the unexpired maturity of the bills sold to the Bank. When the stop rate is expected to change, however, the cost of one day funds is not the same under all choices of unexpired maturity. Institutions thus exercise their discretion, choosing to sell bills with the shortest (permitted) unexpired maturity when the expected change is a fall in the stop rate and the longest (permitted) unexpired maturity when the expected change is a rise in the stop rate.

The source of institutions' discretion over the unexpired maturity of the bills which they sell to the Bank – and hence of the asymmetric behaviour of the term structure – is the Bank's chosen policy of buying bills *with a range of unexpired maturities.* The asymmetry would disappear if the Bank of England chose, instead, to supply cash by buying bills of *a single unexpired maturity*, say 14 days. This would leave institutions no choice but to borrow at this maturity, i.e. $\underline{n} = \bar{n} = 14$ days, and would cause the asymmetry outlined in the example to disappear[13]. The term structure would thus continue to pivot[14], but in a symmetric fashion about the maturity $m = 14$ days.

[13]This would hardly be a sensible method of operating, given that the extant stock of bills with an unexpired maturity of exactly 14 days is unlikely to be very large. A much more practical alternative, which would eliminate term structure pivoting altogether, is suggested in section 3.

[14]Albeit in opposite directions, depending on the expected stop rate change.

Given that the Bank does not operate this way in practice, the asymmetry observed in the example provides a convenient means of testing the above model: short-term sterling interest rates should show *different* behaviour ahead of an expected stop rate fall than ahead of an expected stop rate rise. Specifically, the term structure of money market interest rates should reflect the asymmetric behaviour outlined above. In cases where the stop rate is expected to rise, some short term rates should be below the stop rate, whilst longer rates should anticipate this rise (as in figure 5.3). In cases where the stop rate is expected to fall, almost all money market rates should anticipate this fall (as in figure 5.2). This is, however, not the only empirical implication of the model.

The actual time path in this model of the one day rate and hence all other longer term rates, naturally depends on how ρ changes over time. Although this has not been modelled explicitly, the likely pattern is that ρ fluctuates but generally rises towards unity (i.e. expectations become firmer) as the date of a change in the stop rate is approached. From figures 5.2 and 5.3 it is clear that changes in ρ imply more dramatic changes in very short term rates than in longer term rates. Thus another implication of the model is that shorter term rates should be more volatile than longer term rates.

Again, however, the effects should be asymmetric. Specifically, instances when the stop rate is expected to rise should be associated with greater volatility in money market rates, over a wider range of maturities, than instances when the stop rate is expected to fall.

5.2 Empirical evidence

5.2.1 Relaxing the assumptions

Before confronting the data, it is necessary to mention two possible ways in which the assumptions of section 1 may be infringed in practice, thereby reducing the clear asymmetry which is being sought.

First, there are ways in which the Bank of England might deliberately forestall the cheap borrowing opportunity ahead of a rate rise. It could do this, for example, by buying securities so as to bring about money market surpluses on the days immediately preceding the rise[15]. Alternatively, the Bank could, under its current operating procedures, purposely opt to buy only band 1 (3 to 14 day) bills on these days, or offer only a repurchase agreement with an even shorter maturity, which would reduce the opportunity. In practice, however, the second of these actions would suffer from the disadvantage that they could be interpreted as *confirmation* of the Bank's intention to raise rates. This would cause ρ to rise which, according to figure 5.1, would in turn act to *increase* the spread between very short-term rates and the existing stop rate, thus potentially counteracting the effects of the Bank's preemptive action.

Secondly, and more seriously, it is often the case that the size of the money market shortage represents a significant fraction of the total outstanding value of eligible bills. The assumption that institutions always have sufficient bills of the maturity which they would prefer to sell to the Bank is therefore unlikely to be satisfied at all times, a fact which is substantiated in the following chapter, which looks more closely at the sterling bill market[16]. As shown in figure 5.1, when a stop rate fall is expected institutions will choose

[15]In the event of a surplus, the Bank *borrows* rather than lends funds at the stop rate. As shown in chapter 3, by standing ready to borrow funds when the money market is in surplus the central bank places a floor on money market rates.

[16]As evidence of this, note that 7-day LIBOR was often higher than the stop rate during the period of declining rates from October 1990 onwards (see figures 5.4 and 5.5), which would be consistent with a shortage of short dated bills.

to rediscount bills with the shortest possible unexpired maturity $\underline{n} = 3$. If they can only rediscount longer dated bills, this would mean that the cost of short term funds ahead of a stop rate fall would be higher than suggested by the figure, and the pivotal maturity would be at some longer maturity $n > 3$. In the case of rising stop rates, a shortage of bills with long unexpired maturity would work in this same direction and would tend to reduce the amount by which money market rates fell below the stop rate.

In summary, these violations of the assumptions, especially the second, imply that our predictions should be qualified; they imply that money market rates may in practice rise by somewhat more than suggested by the model before a fall in the stop rate, and fall by somewhat less than predicted before an expected rise in the stop rate.

With the above qualifications in mind, the remainder of this section seeks to verify the presence in sterling money market interest rates of the predicted behaviour. Given the difficulty of observing the magnitude of expectational variables, no attempt is made to relate theory to fact quantitatively, although it is still necessary to identify points in time at which the stop rate was expected to change. Inspection of the time path of the stop rate (figure 5.4 and 5.5) shows obvious autocorrelation in the movements of this rate. Over relatively long periods, therefore, the stop rate was either rising, constant, or falling, and these periods have been identified from figures 5.4 and 5.5 as:

Stop Rate	Period	Days
Increasing	6 Jun 88 - 4 Oct 89	348
Constant	5 Oct 89 - 3 Oct 90	260
Decreasing	4 Oct 90 -4 Sept 91	240

It is assumed in what follows that during these periods of rising (falling) stop rates, there was a general expectation that the next stop rate change would be in the same direction. For these periods, then, three key statistics were examined, namely the level of money market rates relative to the stop rate, the volatility of money market rates, and market rate behaviour in the

period immediately surrounding a stop rate change. Although the shortest maturity interest rate in the sterling money market is one day (i.e. overnight), there are several other important features of the overnight market which affect rates there (see chapter 7), making them a particularly poor variable for purposes of investigating the term structure behaviour outlined above. For this reason, the next section utilises daily data for seven day, one month and three month interest rates on unsecured sterling deposits, all of which are heavily traded each day[17].

5.2.2 Market rates relative to the stop rate

The model which has been presented above proposes that market rates up to maturity $m < \bar{n}$ should lie below the stop rate ahead of an expected rise in the stop rate, whilst such pivoting behaviour of the term structure should be virtually undetectable in the case of a stop rate fall. The mean differences between various interbank rates and the stop rate $(i_t^m - s_t)$ were thus calculated[18] for each of the three periods identified above and are provided in table 5.1.

For the rising period, the mean differences for both the 7-day rate and the 1 month rate are significantly less than zero, showing that, on average, i_t^m is less than s_t despite the qualifications discussed above. For the falling period, the positive differences for these maturities can be taken as a measure of the effect of the shortage of eligible bills discussed in the previous subsection. The fact that these signs are reversed for the 3 month rate indicates that the "normal" effect of expectations (i.e. market rates anticipating the direction of the stop rate change) is dominant at this maturity, which is consistent with the notion of "pivoting" presented in the model.

[17]These data were obtained from Richard Pattinson (Barclays), and are LIBOR rates taken off Reuters at approximately 11h00 daily.

[18]Note that LIBOR rates are yields, whereas the stop rate is a discount rate. All data were thus converted to continuous compound yields to facilitate valid comparison.

5.2.3 Rate volatility

In section 1 it was shown that variations in the probability of a future change in stop rate (a shift in expectations) are associated with greater variations in market rates when that change is a rise than when it is a fall. On the assumption that the pattern of changing expectations whilst rates are rising is similar to that during falling rates, the implication is that market rates for shorter maturities should exhibit more volatility when rates are rising than when they are falling. In the latter case, however, market rates at very short maturities should nonetheless be somewhat more volatile than those at longer maturities. To examine this behaviour, the standard deviation of day-to-day first differences of these and longer term rates was computed over each of the three designated periods and are provided in table 5.2.

The table indicates that volatility is similar across these different maturities when the stop rate is constant and it is greater in all cases when the stop rate is changing. Comparing the rising and falling periods, the volatility of shorter term rates (7 day and 1 month) is higher when the stop rate is rising than when it is falling, but this difference is not apparent in the case of the longer term rates (3 month). These findings are thus in accordance with the predictions. Note that in the case of the 1 month rate, this difference in volatility between the rising and falling periods is significant at the 95% confidence level.

5.2.4 Rate behaviour around a change in stop rate

The above tests have compared averages over the periods of rising and falling rates respectively. Whilst their results are conclusive, these tests are crude in that no attention is given to the behaviour of market rates just ahead of the rise (fall).

It has been asserted above that knowledge of the timing of an imminent change in the stop rate may become more precise as the date of the change approaches. This implies, in the case of a rise in the stop rate, that money market rates should first fall and then rise abruptly when the stop rate rises.

Sharp increases should therefore be visible when comparing market rates on the day of the rise with those a few days preceding this. The symmetrically opposite effect preceding a rate fall should, however, be virtually absent, since the firming of expectations in this case rather implies a smooth transition of nearly all market rates to the new stop rate.

Following Dale (1992), the changes in the 1 month rate in the 3 days prior to changes in the stop rate (i.e. from day $\tau - 3$ to day τ) were computed. These changes, expressed as a percentage of the change in the stop rate and averaged over increases and decreases respectively in the stop rate between June 1988 and September 1991, are shown in table 5.3.

As it turns out, the mean change for the rate increases is greater than the mean change at the decreases at the 99% confidence level, which again convincingly confirms the presence of the effect which is being sought. The 1 month rate rises, on average, by nearly the amount of the stop rate during the 3 days before a rise in the stop rate, whilst the symmetrical effect at a rate fall is not significant.

5.3 Assessment

5.3.1 Problems

By standing ready to lend reserves on a daily basis, the Bank of England clearly reduces the fluctuations in money market interest which would otherwise occur in response to movements in aggregate reserves. Using the terminology of chapter 2, it is evident that the Bank of England engages in "smoothing" of money market interest rates. It has been shown above, however, that by choosing to operate in the manner that it does, the Bank does not smooth these rates perfectly. Indeed, the Bank's operations may, through market expectations of changes in its official rates, actually *impart* fluctuations to money market interest rates. Clearly these rate fluctuations – which may be considerable – are confined to very short-term interest rates, and are thus of limited macroeconomic significance. This does not mean, however,

that they do not carry any negative consequences.

Volatile money market interest rates, even if these are very short-term rates, pose problems for financial institutions. Recall from chapters 3 and 4 that markets in very short-term loans, such as overnight loans, are precisely those markets which are favoured by banks and other financial institutions for purposes of reserve (cash) management. Given the repetitive but random nature of day-to-day fluctuations in the reserve needs of financial institutions, they will rationally choose to manage these temporary fluctuations with instruments of a very short maturity. Borrowing or lending long-term funds to deal with temporary reserve needs would expose financial institutions to greater, and unnecessary, interest rate (price) risks. it would also typically require a larger – and more costly given the existence of a spread – volume of borrowing or lending for a given sequence of temporary reserve disturbances, since these disturbances tend to be distributed around zero. Of course, when very short-term interest rates are volatile and less predictable, reserve management using these short-term instrument is likely to become more difficult and more costly. Banks may then be induced to pursue more conservative reserve management strategies to avoid paying what they perceive as penal rates of interest, when these are temporarily high. They will also, in turn, pass on some of these costs to their customers. In sum, therefore, increased volatility in very short-term interest rates is likely to be consistent with a less efficient money market, at least as far as its role in redistributing reserves is concerned.

Interest rate volatility is not, however, the most problematic aspect of the Bank's operations. More troubling is the fact that they may cause some money market interest rates to move in the *opposite* direction to an impending change in official rates. Clearly this change in the monetary instrument is only considered necessary by the Bank because a policy target, for example the price level or an exchange rate, is, or is expected to, diverge from its desired range and a higher or lower level of short-term interest rates is seen as the required step to correct this divergence. Of course, expectations on the part of market participants will, through adjustments in the term structure, initiate this process, thereby reinforcing the Bank's intentions. Hawtrey (1938), in

his study of Bank rate, recognised this importance of expectations in what is now referred to as the transmission mechanism:

> "... when the use of Bank rate to restrict credit became an established practice, traders, being aware of the intentions of the Bank, were inclined to anticipate them. When Bank rate went up ... a trader would reason that this was intended to have a restrictive effect on markets, and that, if the effect was not brought about, the rate would simply go higher and higher until it was .... [These] psychological reactions are in reality no more than a tendency which in any case exists. Were they absent, that would only mean that Bank rate would have to be raised higher" (pp.249-250).

The model in this chapter has shown that, by operating in the manner that it does, the Bank of England provides less scope for these anticipatory forces to work since, until the Bank actually does change its official lending rate, some short-term rates will behave in a manner which clearly frustrates the Bank's intentions. While this may be of limited relevance when the Bank is targeting the price level, which depends to a significant extent on longer-term nominal interest rates, it is of considerable relevance when the Bank is targeting an exchange rate, which is far more sensitive to the level of short-term interest rates. Until sterling was withdrawn from the Exchange Rate Mechanism (ERM) in September 1992, the exchange rate was an important policy target. And, whilst the Bank's operating procedures can on no account be held responsible for the inability of sterling to be kept within its designated ERM band, they consistently made it cheaper, rather than more expensive, to speculate against the currency. Repeated announcements by the monetary authorities and the government that they were prepared to raise official interest rates led, on several occasions during 1992, to short-term interest rates falling below the Bank's official rates (see figure 5.6). Consequently, adopting a short position in sterling by borrowing the currency on a short-term basis to purchase another currency (e.g. Deutschemarks) was therefore made cheaper at precisely the time that this behaviour should have been penalised, making the exchange rate target harder to maintain.

Whilst the significance of this errant interest rate behaviour should not be overemphasised, it remains true that it has on occasion – by the Bank's own admission – led to official interest rates being moved sooner than intended:

> " This development [ie. a cut in the US discount rate] immediately led to expectations of further reductions in interest rates in the United Kingdom .... The Bank moved its dealing rates down by 0.5% in the early part of that week. *This move was made more quickly than in earlier episodes because it had proved difficult at times in the past to secure sufficient offers of bills to meet the cash shortages during periods when the Bank's dealing rates had entered a phase of gradual reduction and when there were very confident expectations that they had further to fall*" (*Bank of England Quarterly Bulletin*, December 1982, p.484; emphasis added).

This illustrates a recurring problem in the Bank' operations. When the Bank is expected to lower its stop rate, institutions become reluctant to offer longer-dated bills for sale to the Bank, which means that the money market shortage is not removed during the Bank's open market operations. Instead, institutions rely increasingly upon the Bank's discount facility for this purpose, where secured funds are available, usually at a short maturity (overnight). This tends to cause very short-term (i.e. overnight) money market rates to "spike" upwards, since the wider market remains short of reserves until fairly late in the day and the funds which are borrowed from the Bank may attract a higher interest rate than normal.

5.3.2 A simple reform

The rate behaviour described above arises from what appears to be a simple practice on the part of the Bank of England, namely the provision of reserves in the money market via the purchase of eligible bills in a range of maturities. This practice has a considerable history (see chapter 1), dating back to when bills were virtually the only short term security that was available in ample supply, thus making them an obvious candidate for this purpose (see Bank of England 1982a). Of course, many other avenues for providing reserves now exist, in particular *repurchase agreements* (or repo), which is now the preferred technique of most central banks (see Kneeshaw and Van den Bergh 1989). Although the use of repurchase agreements was actually pioneered by the Bank of England during the 1870s, these were seen primarily as a means of *withdrawing* reserves rather than supplying them, and in any event was

not actively employed for either purpose after world war one. More recently, of course, the Bank has begun to use repo more – both against bills and, recently, against gilts – but outright bill purchases still form the cornerstone of its open market operations, comprising about half of the Bank's total daily transactions nearly 75% of the time. But what is the advantage of using repo, when the source of the above problems lies in the *maturity* of the Bank's operations?

By restricting its counterparties to offer securities for sale under a repo, instead of an outright, basis, the Bank would gain significantly more flexibility in the maturity dimension of its operations. The Bank cannot, at present, reduce the unexpired maturity of the bills which it is prepared to accept, since this would reduce the available supply of bills which were eligible for sale to the Bank to the point where these were insufficient to cover the large daily money market shortages. Under a repo, however, the maturity of the underlying securities is irrelevant[19] (as long as they exceed the maturity of the repo), so the Bank could then easily reduce the maturity of its daily lending to as little as one day (i.e. overnight). This maturity would pose no additional operational problems, given that the Bank already operates in the money markets at least once a day. The advantage of lending over a shorter maturity is clear from figure 5.1, which indicates that, under the assumptions made at the outset of this chapter, the cost of one day funds is then driven, in the limiting case of one day, to equality with the Bank's stop rate *regardless of interest rate expectations.* Whilst the term structure of money market interest rates would still exhibit "pivoting" behaviour, the pivotal maturity would then always be one day, eliminating the tendency for short-term market rates to move in the opposite direction to any expected interest rate changes, as well as reducing the volatility in short-term rates.

Experience in other countries provides some evidence for this claim. The Fed, for example, conducts very few outright open market purchases, preferring instead to use repo transactions whose maturity varies but is usually overnight (see Meulendyke 1992 and Feinman 1993a). Comparative evidence

[19]Except, of course, to the cash margin, which reflects the price risk of the securities offered as collateral to the lender.

on the average day-to-day variability of overnight interest rates presented in Kasman (1992, p.21) shows that sterling overnight rates were about twice as volatile as the federal funds rate in the US. Similarly, the operations of the Banque de France comprise biweekly repos whose maturity is 7 days. In addition, the Banque employs so-called "fine tuning" operations where this maturity is reduced to one day (see Banque de France 1990), which implies that French overnight interest rates show considerably less variation than sterling overnight rates.

As the next two chapters will illustrate, the maturity of the Bank's money market operations are not the only source of volatility in short-term sterling interest rates. It will be shown, for example, that the Bank's reliance upon bills as the primary intervention asset has been a further contributing factor. By moving to the exclusive use of repo as opposed to outright transactions, the Bank would also gain more flexibility over the assets which it routinely purchased. The entire stock of (fully paid-up) gilt-edged securities, for instance, would become available for this purpose, relieving eligible bills of their exclusive status in the Bank's operations. The recent adoption of the bimonthly gilt repo facility on a permanent basis by the Bank (see chapter 4) is encouraging in this regard and suggests that this evolution may already be under way.

5.4 Concluding remarks

It is widely believed that the major influence of interest rates on economic behaviour is through long-term rather than short-term interest rates. Perhaps the most famous proponent of this view was Maynard Keynes in the *General Theory* (1932), who was thus led to argue that:

> "... [p]erhaps a complex offer by the central bank to buy and sell at stated prices gilt-edged bonds of all maturities, in place of the single bank rate for short-term bills, is the most important practical improvement which can be made in the technique of monetary management" (pp.205-6).

And whilst central banks in many industrial countries did seek to influence both short-term and long-term rates in precisely this manner during the interwar period, they now conduct their operations – without exception – so as to influence short-term interest rates. Thus, although short-term interest rates do have a direct effect on economic behaviour, central banks rely to a considerable extent upon market expectations to determine long-term interest rates. If they believe that long-term rates are a more potent channel for influencing economic activity, why then do they not, as Keynes has suggested, set long term interest rates directly, for example by operating via 20-year gilt-edged securities rather than one month bills?

The analysis in this chapter provides one reason why this may not be feasible in practice. Quite simply, the interest rate instrument of central banks is such that they cannot avoid influencing *all* interest rates, since short-term and long-term interest rates are not independently determined. To the contrary, the possibility of intertemporal arbitrage in financial markets imposes restrictions on the behaviour of short-term interest rates relative to long-term interest rates. These were summarised, under conditions of risk neutrality, by equation (5.1) at the outset of this chapter, which is referred to as the expectations hypothesis of the term structure. Thus, if a central bank chooses to influence the level of 30-day interest rates over time, this not only has implications for all long-term rates, *but for all short-term rates as well.* Of course, this remains equally true whether the central bank decides to influence 30-month or even 30-year interest rates. If the central bank did indeed try to operate its monetary policy through determination of the 30-year rate, however, it would then have to accept the resultant behaviour of all shorter-term rates which, by a simple extension of the analysis in this chapter has demonstrated, can be expected to be neither stable nor consistent with the achievement of the central bank's policy targets. From this perspective, then, it is entirely appropriate that central banks restrict their influence over interest rates to shorter rather than longer maturities.

5.5 Figures and tables

Figure 5.1
The cost of one day funds:
equation (5.6)

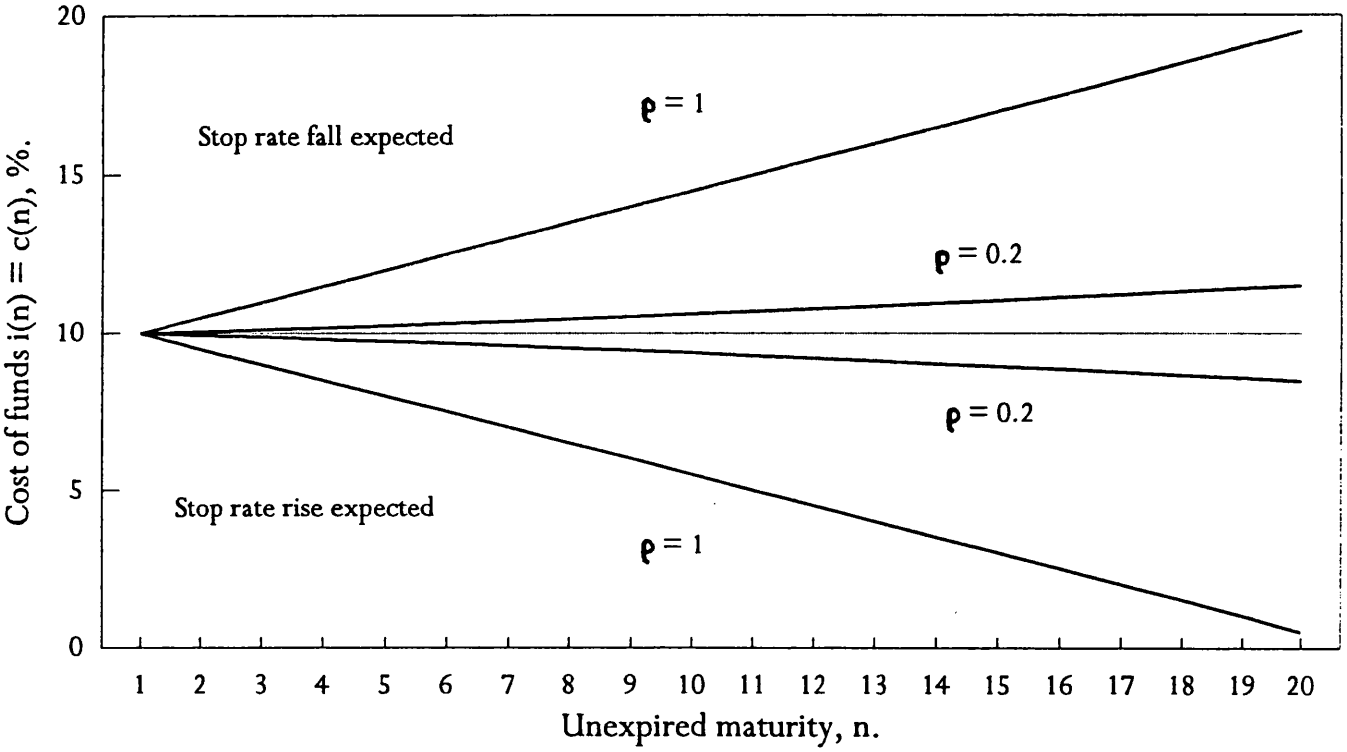


Figure 5.2
The money market term structure
before an expected fall in the stop rate

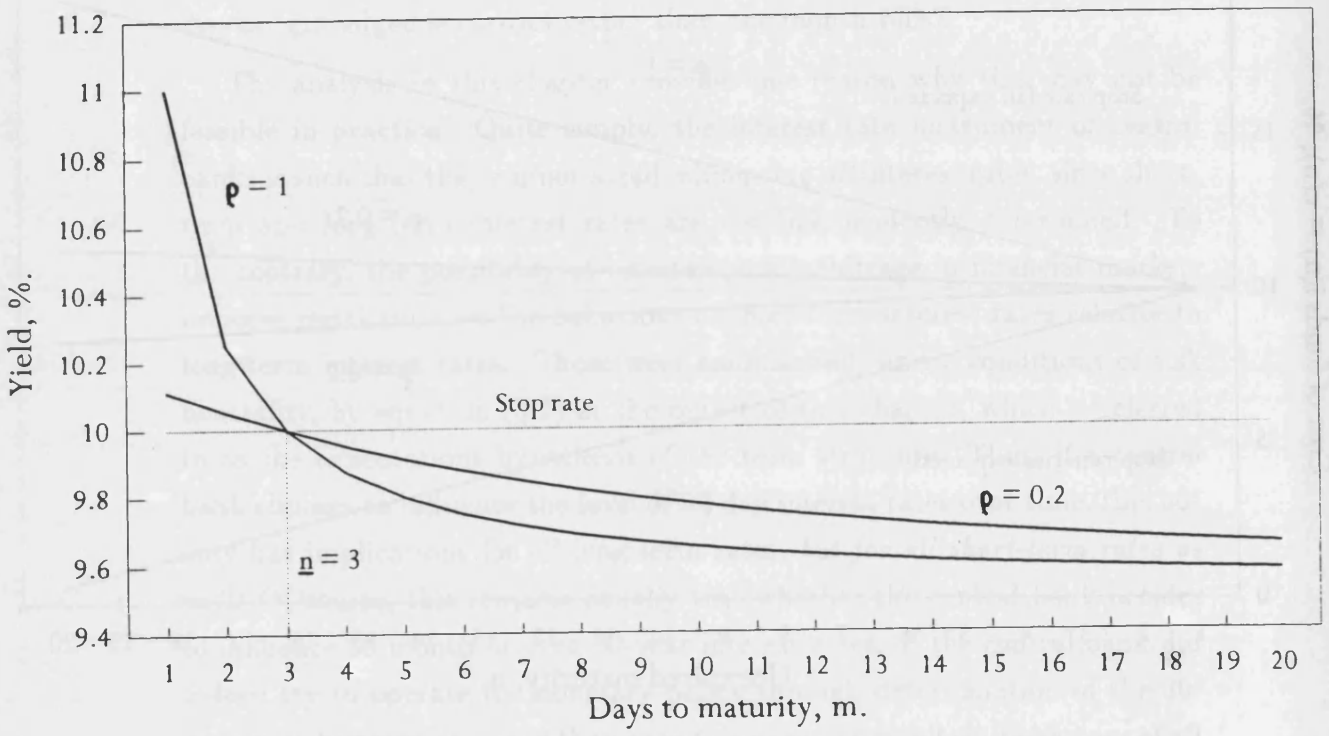


Figure 5.3
The money market term structure
before an expected rise in the stop rate

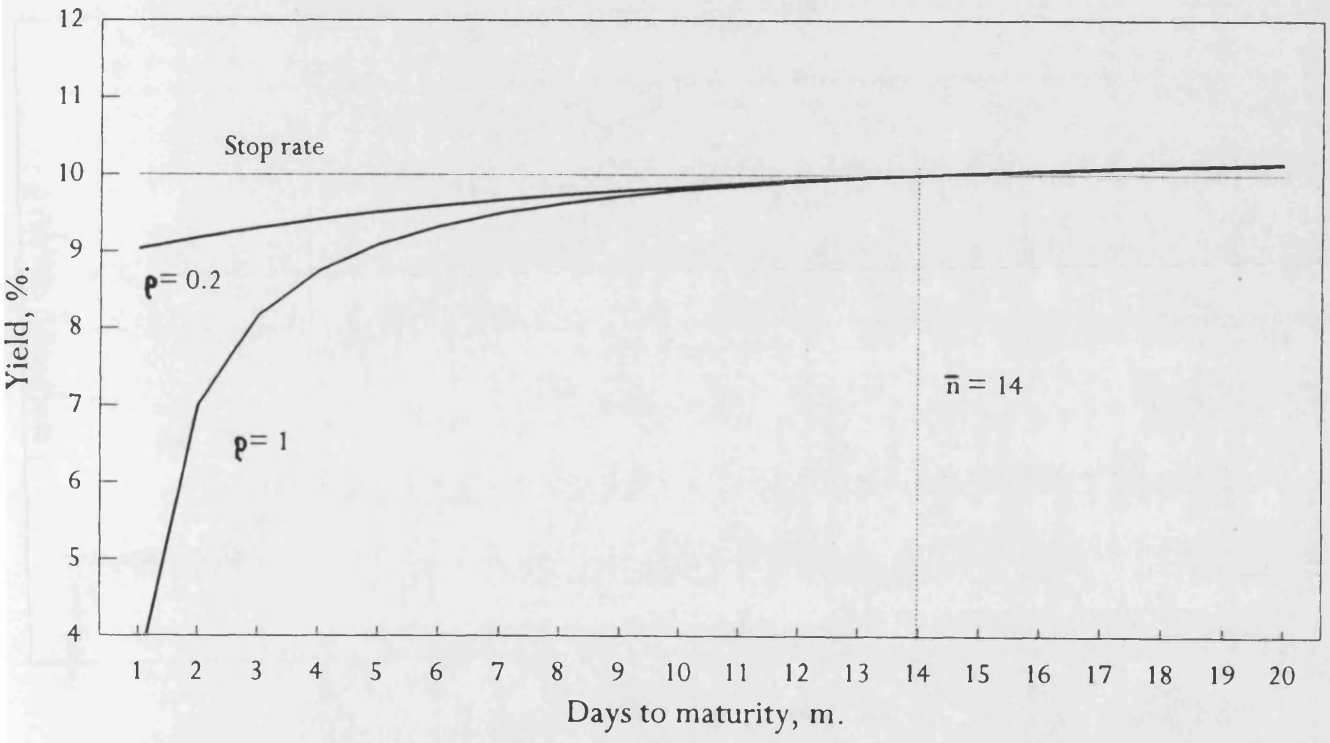


Figure 5.4
The stop rate and 7-day LIBOR
1988-1992

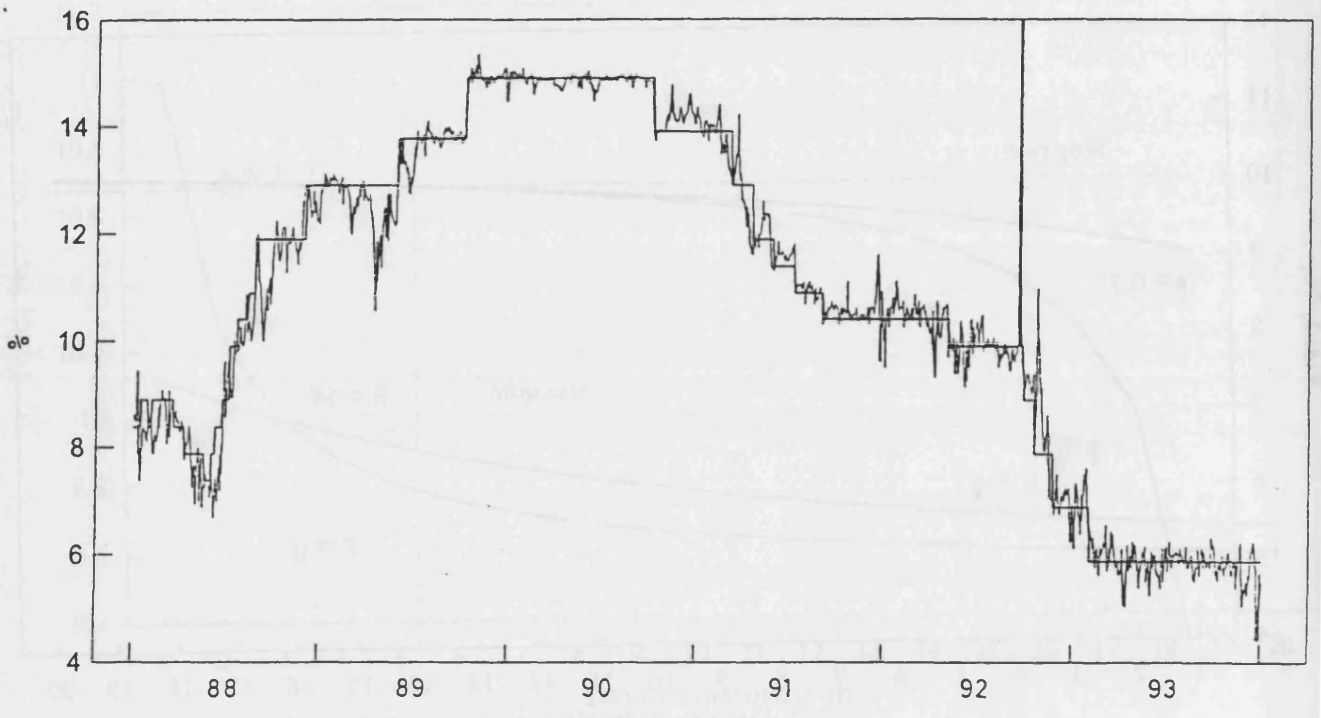


Figure 5.5
The spread between the stop rate and 7-day LIBOR
1988-1992

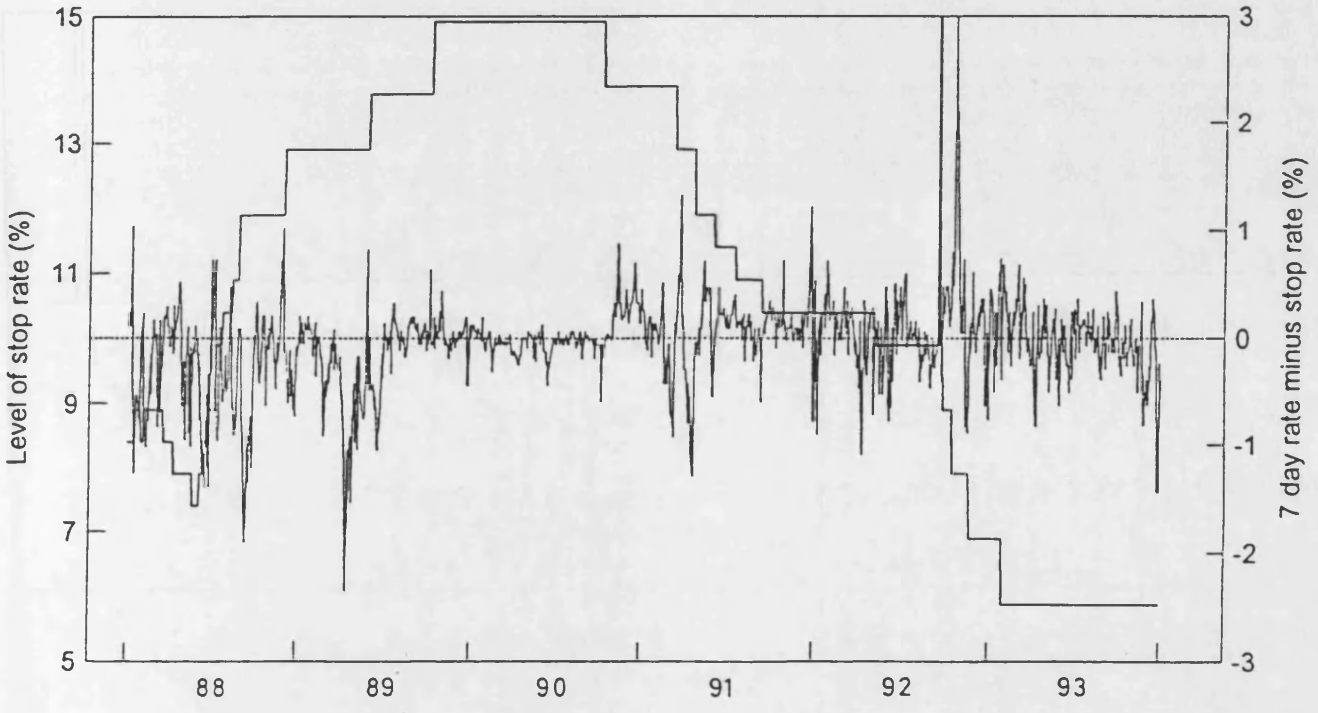


Figure 5.6
The stop rate and 7-day LIBOR
January – September 1992

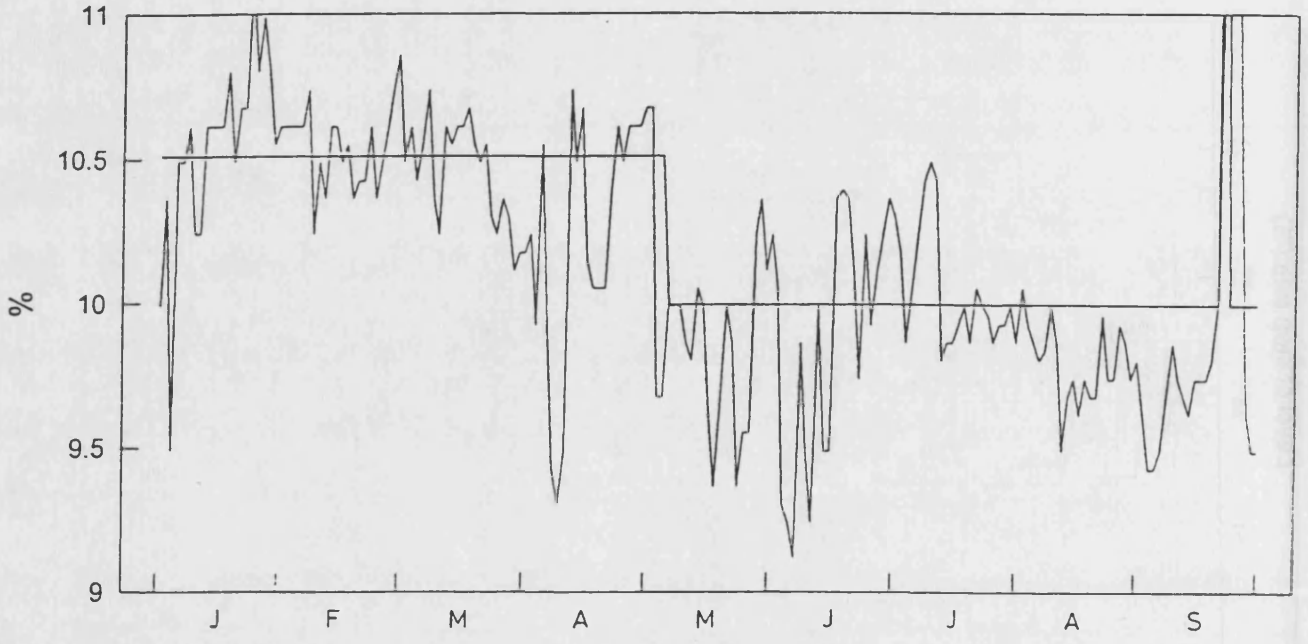


Table 5.1
Mean difference (%) between market rate and stop rate

Period	Maturity		
	7 days	1 month	3 month
Rising	-.348	-.350	.054
	(.025)	(.010)	(.012)
Constant	.041	.073	.033
	(.007)	(.006)	(.009)
Falling	.043	.133	-.230
	(.016)	(.0.12)	(.014)

Standard errors in brackets.

Table 5.2
Volatility of money market rates

Period	Maturity		
	7 days	1 month	3 month
Rising	.039	.022	.011
	(.005)	(.005)	(.002)
Constant	.005	.004	.004
	(.001)	(.002)	(.009)
Falling	.030	.011	.009
	(.008)	(.004)	(.005)

Standard errors in brackets.

Table 5.3
Mean change in the one month rate around a stop rate change

Stop Rate	Number of events	Mean change in 1 month rate
Increases	12	82.5 (21.8)
Decreases	11	4.9 (11.0)

Standard errors in brackets.

Chapter 6

The Excess Liquidity Premium on Sterling Eligible Bills

The yield on sterling eligible bills of exchange has, for many years, been below the yield on equivalent money market assets. This differential undoubtedly reflects the special role which these instruments play within the sterling money market. For two reasons, eligible bills are considered by UK banks to be so-called "prime liquidity". First, bills are routinely purchased by the Bank of England in its money market operations. This means that the holder of an eligible bill is guaranteed, under most market conditions, to exchange this asset for cash. Second, and because of this feature, bills are classified by the Bank of England as reserve assets, which institutions may hold in order to satisfy their statutory liquid asset requirements. There is an obvious circularity here: eligible bills are considered as reserve assets by the Bank *because* they are liquid; but bills are only considered as liquid *because* they are eligible for sale to the Bank. The superior liquidity of bills, therefore, is derived solely from their eligibility and implies that eligible bills will be voluntarily held even when their yield is below that of equivalent money market assets. Eligible bills, in other words, carry a liquidity premium.

Unfortunately, this liquidity premium ($\phi_t > 0$) is not directly observable. What *is* observable, is the *spread* ($S_t > 0$) between the yield on eligible bills

(i_t^b) and the yield on other sterling money market instruments (i_t). However, this yield spread will be equivalent to the liquidity premium in the absence of any constraints on the demand for, and supply of, eligible bills, i.e.

$$S_t = i_t - i_t^b = \phi_t \,.$$

To see why, consider what would happen if, in the absence of any such constraints, the yield spread on eligible bills became larger than the (unobservable) liquidity premium. In practical terms, this would mean that the yield on eligible bills was below the yield on some equivalent asset, such as an unsecured sterling loan for example, i.e. LIBOR, by *more* than was warranted by the extra liquidity of bills. Faced with this yield differential, bank treasurers would clearly rather lend via the unsecured deposit market rather than via the bill market. Conversely, corporate treasurers would rather borrow by drawing a bill. Wherever possible, both parties would engage in arbitrage, selling (drawing) bills and lending in the unsecured deposit market. These forces would quickly tend to narrow the yield spread between bills and LIBOR, to the point where their *liquidity-adjusted yields* became equal. Thus any spread would reflect only the liquidity premium, which would tend to stay more or less constant over time.

As this chapter indicates, the reality is less straightforward. The annual average spread between one month LIBOR and one month eligible bills has varied considerably in recent years, being as low as 10 basis points in 1987 and as high as 50 basis points in 1992 (see table 6.1 and figures 6.3 and 6.4). Indeed, the trend over the past five years has been for this yield spread to *widen* rather than to narrow. This implies that there has in fact been an *excess liquidity premium* on eligible bills, i.e.

$$S_t > \phi_t \,,$$

and indicates the likelihood of constraints which have tended to counter the equilibrating forces suggested above. These constraints, moreover, are likely to have emanated from the Bank of England's own regulation of the bill market, particularly its imposition of accepting limits on eligible acceptors. Thus the Bank's practices are not only responsible for the appearance of a

liquidity premium on eligible bills, but may also be responsible for an excess liquidity premium on these instruments in recent years. But why should an excess liquidity premium matter?

Unlike many other sterling money market instruments, like certificates of deposit (CDs) or commercial paper (CP), bills have been, and still are, the fulcrum of the Bank's money market operations, through which the Bank seeks to influence the level of sterling money market interest rates generally. The yield on bills, therefore, is relevant to the behaviour of those institutions which routinely buy, hold and sell these bills to the Bank (such as the discount houses and, latterly, the clearing banks) in a manner that the yield on CDs or CP is not. Discount houses, for example, are obliged by the Bank of England

> " ... to make to banks and to non-bank members of either CHAPS or the Town Clearing, on demand and in any trading conditions, ... continuous and effective bid prices ... at which they stand committed to deal in eligible bills for same day settlement and in marketable amounts" (Bank of England 1988, para 9).

Clearly if eligible bills earn less – on a liquidity adjusted basis – than equivalent money market assets, yet have to be bought and held by some institutions as a necessary part of the normal operation of the UK banking system and monetary policy, then these institutions will face incentives to recoup their costs by generating cheaper funding opportunities (or more profitable lending opportunities). Indeed, only by being able, occasionally, to borrow very short-term funds at a lower than average rate (or being able to lend funds at a higher than average rate) will the holders of bills be able to fund these instruments profitably. This is precisely the behaviour which does characterise very short-term sterling interest rates (i.e. overnight rates), which are widely known to be more volatile than those elsewhere. This suggests that the cost of the excess liquidity premium is not being borne so much by the holders of bills, but by the users of the market for short-dated sterling funds, which includes domestic as well as foreign banks, and implies that this market is less efficient than it could be.

Section 1 presents a brief history of the sterling bill market followed, in section 2, by an empirical picture of the behaviour of eligible bill yields (in

particular one month yields) which confirms the presence, in recent years, of a considerable widening in the spread between LIBOR and the yield on eligible bills. Section 3 then provides a profile of the supply side of the bill market, where eligible bills are created as credit instruments. With this profile in place, section 4 goes on to examine several hypotheses as to the supply constraint which is most likely to have driven the "scarcity" of eligible bills. Having identified the likely causes of the excess liquidity premium on eligible bills, section 5 traces some of the effects of this spread. Some concluding remarks are made in section 6.

6.1 The evolution of the bill market

6.1.1 Emergence

Inland bills first became legal in England in 1697, three years after the formation of the Bank of England, although they had been extensively used before this to finance trade throughout Western Europe. During the eighteenth century the Bank routinely rediscounted bills, usually at the maximum legal discount rate[1], and thereafter at Bank rate (see chapter one, figure 1.1). In order to be considered as eligible for rediscount at the Bank, however, bills had to bear two good London names and have a tenor not exceeding 65 days. The fact that bills could be rediscounted with the Bank increased their attraction as high-quality, liquid assets and reinforced their role as the primary short-term financial instrument.

In addition to serving as a source of trade finance, bills served as an early mechanism whereby funds were redistributed within the nascent banking system in the absence of a branch network (see chapters 1 and 3). Bill brokers developed to facilitate this trade in bills, and by 1822 some twenty five bill brokers were in business, although this number fluctuated considerably (Scammel 1968). In 1825 a banking crisis occurred which was precipitated, in

[1]The Usury Law of 1714 stipulated a maximum legal interest rate of 5% until 1833, when it was repealed. See Scammel (1968).

part, by the Bank's refusal to discount bills. This prompted a reorganisation of the financial system which not only heralded the emergence of joint-stock banking (in 1826), but also led to the bill brokers being transformed into a unique set of financial institutions known as discount houses. Consequently, commercial bills retained their place at the centre of the sterling money market, a position they were to hold for nearly eighty five years.

In 1873 the overall growth of commercial bills reached a turning point, whilst inland bills began to suffer a gradual decline relative to foreign bills as an instrument of finance. One explanation for this decline was the development of branch banking, which had reduced the role of bills in redistributing cash within the banking system. King (1936) argues that it became a matter of indifference to the banks whether they financed their customers by discounting their bills or by granting loans and advances, and to the customers the flexibility of the overdraft system had definite advantages. Nishimura (1971) disputes this explanation, claiming that branch banking only reached maturity in the 1890s. He argues, instead, that it was the development of communication and transport which reduced the general level of inventories (and so, too, the need for bill finance) for internal trade, whilst at the same time stimulating foreign trade and the demand for foreign bills.

6.1.2 War and the dominance of Treasury bills

The outbreak of war in Europe in 1914, and the attendant disruptions of trade and increases in government borrowing, brought about a sharp decline in the issue of commercial (bank) bills and a rapid increase in the issue of Treasury bills. Although Treasury bills had first been issued in 1877, they had thus far played a minor role in the finance of government. This changed dramatically with the advent of the war, when Treasury bills rapidly outstripped commercial bills as the dominant money market instrument. Indeed, they remained dominant until the 1960s, at first due to the depression in the 1930s which further disrupted trade, and then due to the second world war[2]. This development meant that the Bank of England, over this period,

[2]For a brief history of the sterling Treasury bill see Midland Bank (1961).

routinely made advances against Treasury bills instead of commercial bills in its lending operations with the discount houses.

It is noteworthy that, during the 1930s, the business of the discount houses came under severe threat due to the government's "cheap money" policy, under which the Bank of England dropped its Bank Rate to a level of 2%[3]. Low demand for loans and advances due to the depression led banks to invest in Treasury bills, which pushed the discount rate on Treasury bills even lower, to below 1%. This caused the houses to incur losses on their bill portfolios since the clearing banks, under a cartel agreement, refused to lend cash on call to the houses at less than 1% under Bank Rate. After several discount houses failed, it was eventually decided to resuscitate the market via a series of gentlemen's agreements in 1934. The clearing banks agreed not to participate in the weekly Treasury bill tender, not to buy Treasury bills which had run for less than seven days and, lastly, to buy bills at a rate which was at least as high as their minimum call lending rate to the houses. Without these agreements, which essentially guaranteed that the houses made a small profit, it is almost certain that these institutions would not have survived the 1930s.

The bill market which emerged from the second world war thus had a small number of discount houses (eleven) and an outstanding issue of Treasury bills which, in 1958, exceeded the issue of commercial bills almost sixfold (see Bank of England 1961 and 1967). Bank bills, which were commercial bills accepted by a British bank or one of the eighteen "accepting houses", were still considered eligible for discount at the Bank of England, but formed a small part of the Bank's lending operations, which mainly comprised advances to the houses against security of Treasury bills.

6.1.3 The resurgence of the bank bill

By 1965, the amount of commercial bills had risen to the point where their issue equalled that of Treasury bills. The reasons for this post-war resurgence,

[3]Where it remained for nineteen years; see Scammel (1968).

according to a commentator at the time, included the abolition of *ad valorem* stamp duty on bills in 1961, a reduction in acceptance fees, a desire on the part of the Treasury to keep the Treasury bill issue low, and "a shift in the attitude of the banks towards the commercial bill, and a suspected shift in that of the Bank of England" (Law, 1965, p.341). At that time, the Bank's policy regarding commercial bills was to maintain the standards of quality associated with the prime bank bill, and to allow reasonable development of bills of all classes, given their usefulness as financial instruments. To this end, the Bank had regularly purchased small amounts of bills to monitor their quality.

The corset

Renewed growth in eligible bills during the 1970s was prompted by the scheme of supplementary special deposits[4], otherwise known as the "corset", which was introduced by the monetary authorities in 1973 to curb excessive credit growth. Under this scheme, banks had to place non-interest bearing reserves with the Bank of England whenever their eligible liabilities grew beyond a specified rate. Acceptances were, however, excluded from the definition of eligible liabilities. Predictably, therefore, bank bills quickly became a dominant channel for short-term lending, with the result that the bill issue rapidly expanded. From a level of about £350 million at the beginning of the first corset period (1973), the amount of bills eventually peaked at £2700 million just before the scheme was abolished (1980). The corset had hardly been abandoned, however, when the monetary authorities, again as part of their attempts to meet monetary targets, embarked on a policy of "overfunding" the Public Sector Borrowing Requirement (PSBR). This policy, even more than the corset, once again placed the eligible bill at the centre of the sterling money market.

Overfunding

The PSBR represents the difference between public sector revenue and expenditure and thus measures the borrowing (or funding) need of the public

[4]For details and definitions see Bank of England (1982b).

sector[5]. Overfunding, as its name suggests, occurs whenever the public sector borrows *more* funds than are needed to meet the PSBR[6]. Since most public sector accounts are held with the Bank of England, a natural consequence of overfunding is the accumulation of public sector balances at the Bank and a corresponding reduction in clearing bank balances. Overfunding thus results in an increase in the *money market shortage*[7], which requires increased lending operations by the Bank of England in order to replenish aggregate reserves.

An initial response by the Bank to the introduction of overfunding was to reduce the amount of Treasury bills issued at the weekly tender in order to lessen the impact of these bill sales on the money market shortage. However, shortages remained so large that this strategy soon depleted the outstanding stock of Treasury bills to the point where discount houses (and banks) had insufficient Treasury bills to cover the shortage. The Bank thus had to rely increasingly upon purchases of bank bills to remove the shortage. This increased the demand for bank bills *as eligible liquid assets*, which tended to depress their yields relative to comparable assets leading, in turn, to bills becoming more attractive as a source of short-term finance.

New operating procedures

In 1980 the Bank announced that its money market operations would in future rely less upon direct lending to the discount houses and more upon open market purchases of eligible bills. These purchases would, however, still be conducted via the discount market. In order to ensure an adequate supply of bills, held by the discount market, two further steps were taken. First, the list of eligible accepting banks was extended to include those foreign banks which had a substantial and broadly based sterling acceptance business in

[5]For a good account of funding in the post-war period up to 1984, see Bank of England (1984).

[6]This policy was motivated by an attempt to reduce broad money growth without resorting to higher short-term interest rates; see Goodhart (1989b).

[7]The relationship between overfunding and the money market shortage is not one-for-one, however, due to the presence of other factors which influence the shortage. See chapter 4, section 3.

the United Kingdom and whose bills could command the finest rates in the market. Second, eligible banks were required, as part of their liquid asset requirements, to lend a minimum amount of funds on call to the discount houses (see chapter 4).

These changes clearly rejuvenated the role of the bank bill as a prime liquid asset to banks and as a source of short-term finance to borrowers. Between 1978 and 1983, for example, UK banks' holdings of bank bills relative to Treasury bills rose from about 50% to over 90%, whilst over this period the volume of sterling acceptances rose by nearly 300%. The changes also reinforced the position of discount houses as intermediaries in the sterling money markets, a position which had been under increasing pressure from the large, active interbank markets in unsecured sterling funds which had developed throughout the 1970s. The development of these markets meant that banks could now borrow funds directly from other banks, financial institutions or corporations, without recourse to the discount market as an intermediary.

During the early 1980s the policy of overfunding, due to the large money market shortages it created, led to a dramatic and sustained increase in the portfolio of eligible bank bills held by the Bank of England. As shown in figure 6.1, this "bill mountain" had reached a peak of £12 billion in 1985 (against a Treasury bill issue of about £1 billion) and the Bank's bill purchases had become increasingly large relative to the stock of outstanding eligible bank bills. In the 1985-86 fiscal year it was decided to discontinue the overfunding policy[8] and to adopt the following *full fund rule*[9]:

> "The authorities will seek to fund the net total of maturing debt, the PSBR and any underlying increase in the foreign exchange reserves by sales of debt *outside the banking sector*" (emphasis added).

Although it was called a full fund rule, this was a misleading name, as the

[8]This was due, in part, to the reduced emphasis on broad money targets and, in part, due to "distortions" introduced by overfunding, in particular the possibilities for bill arbitrage (see section 4).

[9]For a summary of funding policy see Annex A to chapter 2 of the Chancellor's *Financial Statement and Budget Report 1990-91*, from which the definition below is taken, or the Treasury Bulletin (1992).

rule would continue to imply either overfunding or underfunding depending on the behaviour of the non-funding sector (i.e. banks). Net purchases of gilt-edged securities (gilts) by banks, for example, would result in overfunding, since the government would then have to issue gilts in excess of the funding requirement and vice versa.

Initially, the new full fund rule resulted in a small degree of underfunding, reversing the previous trend of large money market shortages and bill purchases by the Bank. In 1988-89, however, overfunding of some £2 billion again occurred, even with a PSBR which was in surplus. Two subsequent amendments to the full fund rule also increased its *future* potential to cause overfunding. In 1988-89 net purchases of government debt by the building society sector were no longer considered as funding, whilst in 1989-90 all sales of Treasury bills were excluded from the definition of funding. The upshot was that overfunding, and hence large money market shortages, continued to require substantial amounts of bank bills to be sold to the Bank of England on a daily basis. As figure 6.2 indicates, whilst gross annual assistance by the Bank of England (i.e. total bill purchases and lending by the Bank) peaked in 1985 and declined for some years thereafter, it subsequently increased beyond its 1985 level, totalling over £250 billion in 1992.

6.2 The yield on eligible bills

The yield on eligible bills over the period 1975-1993 is presented in figures 6.3, 6.4 and 6.5, which depict, for a maturity of one month, the spread between LIBOR and the yield on Treasury bills, the spread between LIBOR and the yield on bank bills, and the spread between the yield on bank bills and Treasury bills, respectively[10].

The adoption of new money market operating procedures by the Bank of England in September 1981 indicate a clear break in the time pattern of these spreads. As noted in the previous section, the period leading up to

[10]The data is daily (weekdays), but has been smoothed using a 21-day centred moving average to facilitate presentation.

this date was characterised by rising money market shortages, which saw the Bank buying substantial amounts of Treasury bills and a growing number of bank bills, and reducing the Treasury bill tender. Not surprisingly, therefore, Treasury bill yields responded to this increased demand and decreased supply by falling substantially below LIBOR. The yield on bank bills – instruments which had played a more limited role in the Bank's money market operations – had, until this time, been primarily *above* LIBOR, but also began to fall in response to these developments, especially in 1980, which was a year of unusual and persistent stringency in the money market. It is also noteworthy that, before the change in the Bank's operations, the yield on bank bills fluctuated somewhat independently of the yield on Treasury bills.

After 1981, when bank bills and Treasury bills effectively became perfect substitutes (as assets), their yields were driven, as would be expected, to (approximate) equality (see figure 6.5). Similarly, the yield on both of these assets was now consistently below LIBOR and showed less fluctuation as the sterling money market adjusted to the Bank's new operating procedures. From 1982 onwards, therefore, it is no longer meaningful to distinguish between the yield on bank bills and Treasury bills. Consequently, the remainder of this chapter will simply refer to the yield on eligible sterling bills, where this refers to bank bills *and* Treasury bills and will be concerned primarily with the difference, or spread, between this yield and LIBOR.

6.2.1 Properties of the spread between 1982 and 1993

The yield spread, S_t, between LIBOR and sterling eligible bills between 1982 and 1993 is, as one should expect, a stationary process[11] which, from the sample autocorrelation and partial autocorrelation functions, appears to be an autoregressive process of order 1, i.e. AR(1). Consequently, fitting the model $S_t = \alpha_0 + \alpha_1 S_{t-1} + \epsilon_t$, where $\mid \alpha_1 \mid < 1$ and $\epsilon_t \sim N(0, \sigma^2)$, to the data over this period provides a relatively good description of the evolution of the

[11] A test for a unit root yielded a Dickey-Fuller t-statistic of -21.8, conclusively rejecting the presence of a unit root. A similarly strong result obtains over the entire sample period from 1975 and 1993.

spread over time, as shown in the table below:

Coefficient	Value	Standard Error	t-Statistic
α_0	0.12	0.006	20.25
α_1	0.53	0.016	33.81

$R^2 = 0.285$, $F_{1,3130} = 1142$.

It is encouraging that the (long-run) expected spread implied by the above equation is

$$E(S_t) = \frac{\alpha_0}{1 - \alpha_1} = 0.25 \ ,$$

or 25 basis points, which accords with most market participants' belief about the "typical" spread between LIBOR and bill yields in recent years[12]. The *actual sample mean* over the period 1982-1993, however, was higher than this, being equal to 32 basis points. A look at either figure 6.3 or 6.4, moreover, indicates that the spread varied considerably about its (long-run) expected value, increasing sharply in 1984-1985, then decreasing briefly and then increasing again over a sustained period.

In order to examine these deviations, the annual mean, standard deviation and volatility[13] of the spread was calculated for each year between 1980 and 1993. These values are provided in table 6.1. The table indicates that both the standard deviation and the volatility of the spread declined steadily over the 1980s and into the 1990s, with the exception of 1992, when the variance of the spread increased temporarily [14].

The mean spreads, however, tell a very different story. The last column of the table provides the outcomes of a (one-sided) hypothesis test of the equality between means in successive years and, where the hypothesis of

[12]Furthermore, *bill arbitrage* (see below) is possible for certain borrowers once the spread becomes wider than 25 basis points.

[13]Calculated as the standard deviation of daily differences in the spread.

[14]This is plainly visible in figures 6.3 and 6.4, and was clearly only due to the extraordinary events surrounding the withdrawal of sterling from the ERM in September 1992, which are discussed below.

equality between means could be rejected, includes the level of significance in brackets. In the period of interest, namely 1982-1993, the annual mean spreads are the same in only one instance, namely 1983-1984. In all other cases, these tests verify the pattern identified above: successive mean spreads were significantly lower in 1986-1987, and significantly higher in 1984-1985 and in each year after 1987 until 1992.

It is the increase in the spread in this latter period which will be of primary interest in subsequent sections of this chapter. From its low of about 10 basis points in 1987, the mean annual spread increased to nearly 50 basis points in 1992, before coming down somewhat in 1993. An OLS regression of the spread on a time dummy variable over this period confirmed this secular increase (t-statistics in brackets):

$$\begin{array}{rcccc} S_t & = & 0.093 & + & 0.000255\,(t)\,. \\ & & (10.2) & & (14.5) \end{array}$$

The coefficient on the time variable is highly significant, and corresponds to an annual (i.e. 261 days) increase in the spread of about 7 basis points.

6.2.2 Why were eligible bills scarce?

The observed widening of the spread between bills and equivalent assets strongly suggests the presence of forces which tended to make eligible bills more scarce over time creating, in effect, an *excess* liquidity premium on these instruments. In other words, the demand for eligible bills rose, or their supply fell, or both.

The primary source of demand for eligible bills arises from the substantial daily purchases of these instruments by the Bank of England in its money market operations. Such purchases are, in turn, driven by the size of money market shortages. The previous section has indicated that these shortages, and hence the Bank's purchases and holdings of eligible bills, rose substantially between 1982 and 1985 due to the policy of overfunding. This is corroborated by figures 6.1 and 6.2. The abolition of this policy in the 1985-86

fiscal year and the subsequent adoption of a so-called full fund rule temporarily eased these money market shortages. As the banking sector's purchases of gilts rose in 1990, however, this full fund rule again resulted in overfunding, increasing money market shortages and prompting rising bill purchases by the Bank.

This broad pattern in the demand for bills by the Bank matches reasonably closely the pattern in the annual average yield spreads. This co-movement can be seen even more clearly from figure 6.6, which plots the annual changes in these spreads against the annual changes in the Bank's holding and purchases of bills. The question, then, is why the *supply* of eligible bills did not adjust, in the manner suggested at the outset of this chapter, in response to the cheaper cost of borrowing via bills, especially over the period 1990-92. Before examining several possible answers to this question, the next section briefly presents a recent profile of the supply side of the bill market since 1987, describing the creation of eligible bills and looking at the factors which influence the volume of acceptances.

6.3 Accepting: bills as credit instruments

6.3.1 The creation of an eligible bill

Corporate customers who borrow funds via bill finance usually do so through an acceptance credit facility with one of a large number of eligible accepting banks. These 150 or so institutions are listed by the Bank of England as eligible acceptors and comprise the UK clearing banks, domestic and foreign merchant banks, and other domestic and foreign banks (a full list of eligible acceptors is attached in the appendix). Having been approached by a customer for an acceptance facility, an accepting bank will first assess the customer's credit needs which, if the bills drawn under the facility are to be claused as eligible, should be of a short-term, self-liquidating nature[15]. The

[15]In practice, however, clausing may be so broad (e.g. "receivables") as to make it impossible to discern the purpose for which funds were borrowed.

accepting bank will also assess the credit standing of the customer, so as to establish a borrowing limit under the facility and decide an acceptance fee in return for the credit risk which it will assume by accepting the customer's bills.

Unlike many other forms of credit, which may be extended entirely via book entries, the issue of bills under an acceptance facility is governed by the Bills of Exchange Act which defines the bill as an unconditional order *in writing*. Thus bills cannot be dematerialised like other negotiable securities, such as sterling Treasury bills or sterling certificates of deposit. In order to reduce the time and transport costs, as well as the risks, of drawing and signing individual bills as they are needed, one of two alternatives is usually adopted. The customer may sign a number of undated bills for the amounts and periods required and return these to the accepting bank, which will then date and discount the bills whenever the customer requires cash. Or, alternatively, the customer may simply give the accepting bank Power of Attorney to sign bills on the customer's behalf.

Once the accepting bank has dated and signed the bill it assumes a contingent liability: that is, although the accepting bank becomes liable to pay the face value of the bill at maturity, an actual liability is incurred only if the customer who drew the bill fails to make funds available at maturity. Having signed and dated the bill, the acceptor will also usually purchase, or discount, the bill, thereby acquiring an asset in the form of a negotiable instrument[16]. The customer's account is then credited with an amount less than the face value of the bill (calculated as in equation (4.1) in chapter 4). At maturity, the customer's account is debited by the amount of the face value of the bill.

Having acquired an asset in the form of a bill and paid out the consideration to the drawer of the bill, the accepting bank is then faced with a choice. It can either retain the bill on its balance sheet, in which case it must "fund" the asset (i.e. borrow cash). A profit will be realised if the bill can be funded at a cost below its yield, or if the bill is subsequently resold at a

[16]This is a comparatively recent practice in the long history of the bill, and became common in the 1970s. Before this, bills were almost always initially discounted in the market (i.e. bought by a discount house or another bank).

higher price. Alternatively, the accepting bank can remove the bill from its balance sheet by selling (i.e. rediscounting) it in the secondary market. In this case the accepting bank simply earns the acceptance fee. In practice, the choice between these alternatives will depend upon the expected path of future short-term interest rates (see section 5 below).

Regardless of whether the bill is initially retained or sold, the accepting bank will immediately deliver the physical bill to the Bank of England, where the instrument is then logged onto CMO, thus permitting any subsequent changes in its ownership due to secondary market trading to be achieved via computerised book-entry.

6.3.2 Sterling acceptances

Size, maturity, fees

Although sterling acceptances now account for over 96% of all acceptances drawn in the UK, this share has risen to such a high level only recently (during 1992). At the end of 1991, for example, sterling acceptances represented only about 80% of the total drawn. Virtually all sterling acceptances drawn are eligible, with non-eligible acceptances representing only 0.35% of bills drawn.

Total sterling acceptances outstanding at the end of 1993 stood at just over £20 billion, having exhibited considerable variation since 1987: figure 6.7 indicates that from their lowest level of just over £15 billion in June 1987, acceptances rose to a peak of almost £25 billion in October 1992 before falling again to their current level. Figure 6.8 shows that the annual growth of acceptances became negative between mid-1990 and mid-1992, thus coinciding with the recession in the UK over this period. Acceptance growth surged strongly after September 1992 (when sterling left the ERM), for reasons which will be discussed below.

The use of acceptances as compared with overall sterling bank lending is relatively small and acceptances only represented 6% of total bank credit to the UK private sector at the end of 1993. Moreover, and more importantly, figure 6.9 indicates that this proportion has fallen steadily since 1987, when

it stood at over 14%.

Drawers

Acceptances are typically drawn to finance short-term, trade-related transactions. Table 6.2 bears this out, indicating that the sectors which currently rely most heavily on acceptance finance (relative to total bank lending) are food, motor and other manufacturing, the chemical and energy industry, electrical engineering and wholesale distribution. Of these, the sectors with the largest volume of acceptance credit are the food sector (£2.7 billion), followed by wholesale distribution (£2.3 billion) and other manufacturing (£1.4 billion).

The table also indicates that the largest overall drawers of acceptances are firms in the financial sector, which suggests that finance bills now constitute a significant proportion of the sterling market. Unfortunately, the CSO's non-specific definition of "other financial firms" makes it impossible to identify the underlying activity which is generating these bills. Undoubtedly, however, this figure is likely to include bills that are issued by the (notionally) separate financing arms which many large industrial or commercial companies employ to manage their financial requirements. Since these companies will issue bills for a host of reasons, such as to raise working capital or to finance inventories (e.g. cars sold on a hire-purchase basis), the simple clausing which is required on a bill makes it impossible in practice to establish whether such bills are commercial bills or finance bills.

Acceptors

The UK clearing banks are the largest group of acceptors, with about 22% of all accepting business. Of this group, the two largest clearing banks, Barclays and National Westminister, account for over half of the acceptances, making them the largest individual acceptors. However, the overall accepting share of the clearing banks, although it is large, has fallen significantly over recent years (see figure 6.10). British accepting houses (merchant banks) and Japanese banks are the next largest group of acceptors, with about 15% and 12% of the market share respectively. Whilst the market share of the accepting houses has been roughly constant during recent years, Japanese

banks have assumed a marginally increasing share of the acceptance market. This is in sharp contrast to American banks, whose share of the accepting market has fallen from 8% in 1987 to less than 3% at present. Overall, foreign banks account for over 50% of all acceptances drawn in London (sterling and non-sterling).

6.3.3 The volume of acceptances

Like many other forms of credit, variation in the volume of acceptances will be influenced by the overall level of activity in the macroeconomy. In addition, factors which are specific to those sectors which traditionally rely on acceptance credit, such as the demand for industrial and manufacturing output, retail sales and the level of imports and exports, which are themselves not independent of overall activity, will also play an important role.

Another factor driving the volume of sterling acceptances is clearly the cost of acceptance credit relative to other sources of short-term finance. Under an acceptance facility the cost of one month funds (say) is given by the prevailing market discount rate on one month bills, plus an acceptance fee. Corporate borrowers will compare this cost with alternative sources of one month funds, which will include unsecured wholesale loans over a range of currencies[17] and maturities (e.g. overnight, 7 days and one month). These loans are usually obtained at a spread above LIBOR, where this spread will depend upon the creditworthiness of the borrower and will also include an associated cost[18].

---

[17]For example, borrowers requiring one month *dollar* funds have sometimes been able to raise these more cheaply by issuing sterling bills and swapping the proceeds from sterling into dollars.

[18]According to practitioners, this cost (which is usually $\frac{1}{16}\%$) arises due to the cash ratio requirement which is levied as a fraction of banks' retail liabilities. Funding a loan is therefore more expensive, since the bank not only has to borrow the full amount of the loan, but also the non-interest bearing reserves which then need to be held against this liability. In the case of bill, the bank has a choice between funding the bill, or selling in the secondary market, and hence an associated cost is not levied in this case.

For borrowers who require sterling funds, therefore, the relative cost of acceptance finance will depend mainly on the level of the discount rate relative to LIBOR (plus the associated costs), since the additional credit-risk premium is charged by the lender irrespective of whether the credit is extended via acceptance or unsecured loan[19]. From the discussion in the previous section (and figure 4.5) it is clear that acceptance credit was almost always cheaper than the alternative, often by a significant margin. However, this spread does vary considerably over time, which is likely to cause the use of acceptances to vary on a month-to-month basis.

Finally, the payment of taxes is likely to affect the volume of acceptances. Inland revenue statistics[20] show clearly that corporation tax payments in the UK are almost all made in the months of January, April, July and October. These place additional seasonal cash flow demands on firms, which they are likely to meet via short-term borrowing such as acceptance credit.

In order to assess the significance of these factors in the monthly variation in sterling acceptances, the following OLS regression was estimated on monthly data over the period from January 1988 to December 1993:

$$\text{GROWTH}_t = \beta_0 + \beta_1 (\text{OUTPUT}_t) + \beta_2 (\text{SPREAD}_t) + \beta_3 (\text{TAX}_t) + \epsilon_t ,$$

where $\epsilon_t \sim N(0, \sigma^2)$. The dependent variable, GROWTH, is the monthly percentage change in total sterling acceptances. Three independent variables were included to measure the contribution of macroeconomic, tax and relative cost factors to the monthly variation in acceptances. OUTPUT is the month-on-month percentage change in the CSO index of total industrial production. SPREAD is the average monthly spread (in basis points) between LIBOR and the yield on one month eligible bills. Finally, TAX is a dummy variable to measure the significance of corporation tax payments, taking the value of unity in the tax-paying months of January, April, July and October, and the

[19]For borrowers who require non-sterling funds, the relevant comparison is between one month LIBOR in this currency and the all-in cost of swapping their bill funding (which is in sterling) into this currency.

[20]See for example chart 1.2 of the Inland Revenue Statistics (1993, HMSO:London).

value of zero in all other months. The results of this regression are presented in the table below:

Variable	Coefficient	Standard Error	t-Statistic
CONSTANT	-3.67	1.05	-3.5
OUTPUT	-0.14	0.63	-0.23
SPREAD	8.92	4.11	2.17
TAX	8.06	1.43	5.65

$R^2 = 0.35$, $Durbin - Watson = 2.14$, $F_{2,70} = 11.99$

The coefficient on monthly changes in industrial production is insignificant and of the wrong sign[21]. The coefficient on the yield spread variable is significant at the 5% level, suggesting that acceptances do respond to changes in their relative price. The magnitude of the coefficient implies that a 10 basis point increase in the spread is consistent with an increase of just under 1% in the growth of acceptances. The tax dummy coefficient is also significant (at the 1% level) and its value implies that acceptances are about 8% higher in tax-paying months than in non tax-paying months. The overall significance of the regression indicates that the joint hypothesis that all the coefficients are zero can be rejected at the 1% level.

6.4 The excess liquidity premium: causes

The profile of the bill market developed in the previous section has indicated that eligible acceptances comprise a *diminishing* share of sterling bank lending to the private sector, currently representing only 5% of bank lending. Moreover, clearing banks, which are the largest potential acceptors of eligible bills, have reduced their share of the acceptance market from 40% at the end

[21]This is not altogether surprising, and one might expect *annual*, rather than monthly, changes in national output to be a more important explanatory factor. Indeed, regressing annual changes in the volume of acceptance on annual changes in industrial production yielded a significant, positive coefficient.

of the 1980s, to less than 25%. At the same time, persistent daily money market shortages have caused the Bank to continue to buy large quantities of eligible bills on a daily basis. Thus the demand for bills by the Bank has remained high, increasing from 1989 to 1992. On the other hand, empirical evidence was found which indicates that the volume of acceptances *does* respond to a change in the relative cost of bill finance, increasing as this gets cheaper. Yet the supply of bills clearly did not respond sufficiently to meet the increased demands by the Bank of England, resulting in a "scarcity" of eligible bills and reflected in the excess liquidity premium on bills. This section now examines why the supply of bills has failed to respond to this relative price incentive between bills and other sources of short-term funds.

6.4.1 Was supply limited by borrowers?

The effects of the recession

Although the previous section has indicated that the monthly volume of acceptance credit *is* responsive to the relative price of acceptance credit, it is nonetheless true that the overall volume of acceptances is ultimately driven by the needs of companies to finance inventories and domestic and international trade. Hence the annual growth in acceptances also fluctuates with the overall growth of production and trade. Most indices of production and trade show that a severe recession occurred in the UK between 1990 and 1992, which will in turn have reduced the demand for acceptance credit. Indeed, annual growth (measured monthly) of sterling acceptances became negative in mid-1990, only becoming positive again in mid-1992, thus roughly coinciding with the recession. It might be argued, therefore, that the effects of the recession placed an inherent limit on the volume of acceptances, which could account for the failure in the supply of acceptances to respond to the clear yield advantages of borrowing via this channel.

During this period, 1990-1993, average monthly acceptances fluctuated between £20-23 billion. Over the same period, the value of imports averaged about £110 billion per annum, whilst total bank lending to the UK private sector averaged about £360 billion. This would suggest that acceptances

still represented a relatively small part of companies' overall borrowing needs when compared either with the overall level of trade or bank lending, and could have increased substantially in volume before approaching some "natural" limit. It is unlikely, then, that depressed demand for acceptance credit due to the recession could have constrained the overall volume of acceptances. To the contrary, depressed conditions are likely to have increased considerations such as the cost of credit, making bill finance more, rather than less, attractive.

Did borrowers use substitutes?

The experience of the US dollar banker's acceptance market, which contracted considerably during the mid-1980s, may offer another explanation. A study of this market by Jensen and Parkinson (1986) indicates that a significant factor in its decline was the increasing use by borrowers of non-bank sources of finance such as commercial paper[22]. This was prompted by the deterioration in the credit-rating of many US financial institutions throughout the 1980s, which made alternatives such as commercial paper a cheaper source of funds than acceptances. Is it possible that a similar shift towards non-bank finance occurred in the UK, leading to a decline in the issue of eligible bills?

The sterling commercial paper (SCP) market was launched in 1986 by the Bank of England due to increased demands by borrowers for this instrument. As noted in chapter 4, SCP is the liability of the issuer only (one name paper) and, consequently, issuers do not have to pay an acceptance fee. If their credit-standing is well known, this means they could (potentially) borrow more cheaply via SCP than via bills. By the end of 1987, one and half years after its inception, SCP outstanding had reached £2 billion. Subsequent growth in the SCP market, however, was much slower, with the total issue only reaching £5 billion at its highest and averaging only about £3 billion.

There is little evidence, therefore, that the introduction of commercial paper has been accompanied by any significant substitution away from sterling

[22]For additional discussion of the US banker's acceptance market see Helfrich (1976), Hill (1986) and LaRoche (1993).

bills. Indeed, the opposite conclusion is warranted, namely that *the presence of cheap bill finance has been the single most important reason why the use of sterling CP has remained so limited since its introduction.* This conclusion is corroborated by the strong growth in other commercial paper markets, such as those on the continent (see Alworth and Borio 1993), whose introduction was even more recent than that of sterling CP. Total commercial paper outstanding in France and Germany, for example, already exceeds sterling CP outstanding. Until bill yields fall relative to other money market rates, therefore, this state of affairs may be expected to continue.

Bill arbitrage

Perhaps the strongest reason why constraints on eligible bill issue are unlikely to have emanated from the borrowing side is *bill arbitrage.* As its name suggests, bill arbitrage consists of borrowing by issuing eligible bills and relending these funds in the sterling deposit market (see Bank of England 1982b). Of course, this is only profitable if the cost of raising bill finance, which comprises the discount rate plus an acceptance fee, is below the rate at which funds can be lent in the money market[23] (i.e. LIBID). Thus bill rates have to be sufficiently low to compensate the borrower's acceptance fee and also the prevailing bid-ask spreads in the money market[24]. Prime borrowers, for whom the acceptance fee is the lowest, thus have the greatest potential to engage in bill arbitrage. Figure 6.12 indicates that for prime borrowers (who typically face an acceptance fee of $\frac{1}{8}\%$), bill arbitrage opportunities presented themselves during 1988 and 1991, and especially in 1992-1993.

Borrowers are not unconstrained in the amount of bill arbitrage which they can and will undertake. Acceptance facilities, like all other bank loans, are subject to bank-imposed credit limits. Furthermore, acceptances facilities which are used to their limits for arbitrage purposes will leave the company with less flexibility in the event of a genuine cash need or financial con-

[23]Note that this lending may be in *another* currency, such as dollars. There have been instances where profitable arbitrage has been possible by raising sterling funds in the bill market, swapping these funds into dollars, and then lending these dollar funds.

[24]Bid-ask spreads in the one-month sterling unsecured deposit market are typically in the region of $\frac{1}{8}\%$.

tingency. Finally, administrative costs will be incurred, which imply that a minimum spread may be necessary before arbitrage is seen as worthwhile. Spreads of 12.5 basis points were not uncommon after 1991, however, and are consistent with a riskless profit of £6250 for each £5 million of bills drawn and relent in the money market. This will have provided a sufficient incentive for such arbitrage during this period, and borrowers would have engaged in such arbitrage to the fullest extent possible, suggesting again that the demand for acceptance credit is unlikely to have been a significant explanatory factor in the observed scarcity of bills over this period.

6.4.2 Was supply limited by the Bank of England?

Regulation of the bill market

Like any bank, a central bank will be concerned about credit risks, and will monitor and limit its exposure to particular debtors in order to manage these risks. The credit risks assumed by central banks are clearly limited by the fact that they have long ceased to operate as commercial banks and thus hardly ever engage in commercial lending[25]. In addition, the assets held by most of the major central banks often comprise government securities, which are seen as (virtually) riskless. From the earlier discussion of the evolution of the Bank of England's daily operations it is clear that this is less true of the Bank. A large proportion of the Bank's assets consists of eligible bank bills, which are *not* riskless assets, and which involve credit exposures to eligible accepting banks (and the drawers of bills). For this reason, the Bank has regulated the issue of eligible bills.

Originally, bills eligible for rediscount at the Bank of England had to bear two good London (and, later, British) names, which automatically excluded any bills accepted by foreign banks. Similarly, the Bank showed a preference for bills drawn upon self-liquidating (i.e. trade) transactions. During the interwar period, when the Bank did not actively purchase commercial bills

[25]Although it is generally understood that they may *have* to undertake such risky loans - dubbed "last resort" loans - in the event of a crisis which threatens the stability of the financial system.

and their outstanding issue remained small relative to Treasury bills, these regulations were of little consequence[26]. Nonetheless, throughout this period, and particularly after the second world war, the Bank's policy was "to maintain the standards of quality long associated with the London prime bank bill" (Bank of England 1961, p.28). To this end, the Bank imposed *accepting limits* on the major acceptors, which were then the British merchant banks, although these limits were *informally* established and policed. According to the Bank, "... the prime status accorded by the Bank to the acceptances of the accepting houses depends upon the readiness of the latter to make their balance sheets available to the Bank and to answer any enquiry that the Bank might wish to make as a result of their sampling of the bills coming onto the market" (p.29). In addition, the Bank discouraged accepting houses from issuing finance bills by "prior consultation" with the houses and by "[taking] only a limited proportion of finance bills [in its sampling purchases, which] tends to restrict the market's willingness to take up finance bills" (p.29)[27].

Given that the Bank itself did not purchase large quantities of eligible acceptances after the second world war, the issue of eligibility and accepting limits was clearly of secondary importance within the Bank and within the overall functioning of the sterling money market at the time. The policy of overfunding the PSBR, however, described in section 1, dramatically changed the status of eligible bills and placed them at the centre of the Bank's money market activities. As the Bank was forced to purchase ever larger quantities of eligible bills under its obligation to relieve the growing money market shortages, it must have become apparent that the existing eligibility criteria were too narrow to provide sufficient eligible bills. To some extent, there-

[26]Although some foreign banks which had operated for many years in the sterling market probably found it irksome, and unfair, that the acceptances of their British rivals traded at a premium above their own bills.

[27]Evidence as to the impact of these restrictions is contained in the Radcliffe Committee Report (1958), which investigated Britain's post-war monetary system. According to the Report, "the evidence is that [the accepting houses] are never accepting as much as they could" (para. 190) but goes on to claim that variation in total eligible acceptances "coincides very largely with the movements of the total value of UK [external] trade" (para. 193).

fore, the advent of overfunding created a dilemma: *increased money market shortages meant that the Bank had to increase its asset acquisition, but in a manner which would not compromise its overall credit risk*[28].

The only true way out of this dilemma (of how to increase asset purchases without also increasing credit exposure) would have been for the Bank to increase its purchases of government securities (via repurchase agreements so as not to distort their prices). This has become the preferred money market intervention technique of most central banks. Instead, the route which was chosen by the Bank was to include, for the first time, foreign banks on the list of eligible acceptors[29]. As before, issue limits were again placed upon all eligible acceptors:

> "The Bank does not propose to impose any direct limit for monetary control purposes on the volume of acceptance business written by eligible banks. It does, however, intend to set limits on the holdings within its own portfolio of individual acceptors' paper; all limits would be increased in equal proportion should the Bank's total portfolio, for reasons of market management, expand beyond the notional level to which the basic limits are related. These [basic] limits ... will be related to the size of the acceptor's capital and to the proportion of the bank's global business that is conducted in sterling. If the Bank's holdings of an individual acceptor's bill approached the limit, the Bank would discuss the situation with the acceptor concerned" (*Aide Memoire Giving the Bank's Guidance on its Practices when Purchasing Eligible Bills from the Discount Market*, Bank of England, 4 August, 1981).

Accepting limits

The calculation of these accepting limits was published by the Bank in November 1986 in a public notice concerning eligibility. As stated above,

[28]That the increased purchases for bills had caused existing eligibility criteria to become binding is evidenced by the fact that the Bank, on several occasions, rediscounted *finance bills*, even though these were explicitly defined as ineligible. Such occurrences have continued through the 1980s, and market participants readily acknowledge that, in practice, the Bank regularly rediscounts bills which are not claused on the basis of a self-liquidating transaction.

[29]These banks were chosen on three criteria: their overall credit-standing, the quality of their existing sterling acceptance business, and the ability of British banks to have similar access in the home-country of the bank under consideration (i.e. a "reciprocity principle").

accepting limits are based on two criteria, namely capital and the extent of the acceptor's sterling business. Thus an acceptor's *basic limit* (l_i) is given by

$$l_i = \text{Capital}_i \times \left(\frac{\text{Sterling Liabilities}_i}{\text{Total Liabilities}_i} \right) . \tag{6.1}$$

This places an upper limit on the aggregate stock of eligible bills equal to $L = \sum_{i=1}^{n} l_i$, where n is the number of eligible acceptors. If all acceptors issued up to their limit, therefore, an individual bank's maximum, or limit, share (s_i) of bills in issue can be calculated as

$$s_i = \frac{l_i}{L} , \tag{6.2}$$

which clearly varies inversely with the basic limits of other acceptors. The limit share s_i, however, bears no specific relationship to the size of the Bank's bill purchases. In practice, the Bank only holds a bill portfolio which is some fraction π, $0 < \pi < 1$, of the aggregate basic limit[30]. Thus an acceptor's *actual limit share* (s_i'), which determines the maximum proportion of the Bank's portfolio which may be accounted for by that acceptor's paper, is calculated as

$$s_i' = \frac{s_i}{\pi} . \tag{6.3}$$

Note that $s_i' > s_i$, i.e. s_i' is a scaled-up measure of s_i. Since the Bank's actual holdings of bills is not published, however, acceptors are required instead to apply their limit share to the most recently published measure of the total stock of bills outstanding (B). Thus an individual acceptor's actual limit (l_i') is given by

$$l_i' = s_i' B \tag{6.4}$$

$$= l_i \left(\frac{B}{\pi L} \right) , \tag{6.5}$$

which depends positively on the acceptor's basic limit l_i and the outstanding stock of bills B, and negatively on the basic limits of other acceptors L and the Bank's purchases of bills π[31].

[30]In 1986 the Bank's portfolio was estimated as equal to one third of the aggregate basic limit (implying that $\pi = \frac{1}{3}$), but this may well have changed since then.

[31]Notice that $l_i' > l_i$, i.e. acceptors' actual limits exceed the basic limits implied by their capital and share of sterling business.

Accepting limits, since they depend on an acceptor's capital and sterling business, are not equally binding across eligible acceptors. Foreign acceptors, due to their more limited sterling business, will face tighter limits than domestic acceptors, whilst smaller domestic acceptors (e.g. merchant banks), due to their smaller (absolute) capital base, will face tighter limits than larger domestic acceptors (e.g. the clearing banks). Indeed, the size of the clearing banks is such that they are unlikely to be constrained by accepting limits, regardless of how these limits are actually calculated. This is borne out by the fact that, in practice, the clearing banks are typically unaware of their accepting limit, whilst most other accepting banks are not only aware of their limits, but frequently reach these limits.

This is significant, for it means that *the degree to which accepting limits constrain the overall level of acceptances will depend crucially on the accepting behaviour of the clearing banks.* A reduction in the level of acceptance credit supplied by the clearing banks, therefore, prompted (for instance) by capital considerations, will require a greater proportion of the existing demand for acceptance credit to be supplied by non-clearing acceptors. These acceptors, however, are only able to respond to this demand as far as their limits will allow, which then constrains the supply of acceptance credit. Indeed, from (6.5) is it apparent that a reduction in the supply of acceptance credit by the clearing banks, by reducing the outstanding stock of bills B, actually makes the accepting limits of non-clearing banks *tighter.*

This impact of accepting limits may well explain the observed scarcity of bills identified in section 2. There *is* evidence that the clearing banks have, in recent years, reduced their share of acceptances (figure 6.10). This may be due to the fact that capital considerations have played an increasing role in the lending decisions of financial institutions. Consequently, the profitability of loans made under acceptance facilities, where fees have become increasingly competitive, will have been compared with the profitability of loans under alternative lending arrangements, and may have been found to be less attractive[32]. The considerable loan loss provisions made by some of the clear-

[32]Note the important distinction between accepting a bill and purchasing a bill accepted by another bank. The latter entails virtually no credit risk, since a bill is two-name paper.

ing banks during the recession of 1990-1992, especially due to their exposure to the property sector, will undoubtedly have sharpened these comparisons. Whatever the reason for the observed reduction in acceptances by the clearing banks, however, this reduction will have caused demands for acceptance credit to shift towards other acceptors. Accepting limits on these banks may have been such that they were unable to accommodate these demands. For obvious reasons banks will not wish to divulge their limits, nor will they readily admit that they are up against these limits. Nonetheless, casual evidence obtained from these banks suggests that many acceptors were frequently at their accepting limits[33].

Another shred of evidence in favour of the above hypothesis lies in the different time pattern of the (annual) growth of acceptances by the clearing banks as compared with those of all other accepting banks. As figure 6.15 indicates, the growth rates of acceptances for these two groups is highly correlated, but the growth rate in acceptances for the clearing banks is much more *variable* than that of the other acceptors[34]. Thus, while the acceptance growth of the clearing banks became progressively negative through 1991, that of the other accepting banks remained much closer to zero, suggesting that there may have been the type of spillover effect described above. Then, in September 1992, acceptance growth for the clearing banks became sharply positive and in October rose to nearly 50%. In contrast, the acceptances of the other accepting banks only grew by about 25%, suggesting that their capacity to expand acceptances was constrained relative to the clearing banks.

This sharp reversal in acceptance growth – which saw the annual growth

---

Accepting a bill, on the other hand, entails a direct credit exposure to the drawer of the bill, which is remunerated by the acceptance fee. It is this return which a bank will weigh up against the return on its other lending business.

[33]Customer relationships are valuable, and thus acceptors may continue to accept bills even when they have reached their limit. They must, however, hold and fund these bills for some time before selling them into the wider market, which usually incurs a cost given the low yield on bills.

[34]This is partly the result of the fact that the clearing banks have a smaller share of acceptances, which means that any given change in acceptances will necessarily reflect a greater percentage change.

rate in acceptances surge above 25% in October 1992 – is likely to have been triggered by the huge money market shortage resulting from the large purchases of sterling by the monetary authorities just before the currency left the ERM. Even though the Bank offered temporary facilities whereby it could relieve most of this shortage, its purchases of bills rose significantly. The market discount rate on bills thus fell sharply (see figure 6.4), stimulating the observed increase in the demand for acceptance finance. This is also the time when bill arbitrage opportunities were at their greatest (see figure 6.12).

Although the experience of September 1992, together with some degree of overfunding of the PSBR, kept shortages large for most of 1993, conditions in the money market have been less stringent of late, for two reasons. First, the Bank has made permanent – and more frequent – the temporary facility invoked in 1992 (see chapter 4). And, secondly, the government has now modified its definition of funding to include government securities purchased by banks and building societies, thereby reducing the potential for its previous funding rule to result in overfunding. These changes have reduced the Bank's demands for bills, and so have probably eased the impact of accepting limits on bill yields. Furthermore, they have in all likelihood reduced the chances that these limits will again become binding in the future.

Having identified a plausible candidate for the causes of the observed scarcity of bills, the following section now traces some of its effects.

6.5 The excess liquidity premium: effects

6.5.1 Funding and rediscounting costs

The fact that the Bank of England buys bills each day should result in the market discount rate being (approximately) equal to the Bank's stop rate. To see why, note that any difference between the market discount rate and the Bank's stop rate would imply a difference between the market price for a bill and the Bank's price for a bill. Using equation (4.1), this price difference may be expressed as

$$P_{Bank} - P_{market} = \frac{m\,(d-s)}{365}\,, \tag{6.6}$$

where d is the market discount rate, s is the Bank's stop rate, and m is the unexpired maturity of the bill. When $d \neq s$ there would thus exist either a sure capital gain or capital loss on bills which were sold to the Bank of England. Under normal circumstances, therefore, the market price of bills would adjust to eliminate such a sure gain or loss.

This has not, however, happened in practice. Figure 6.11 indicates that market discount rates on one month bills have typically been *below* the official stop rate (the same is true for the discount rate on three month bills). The explanation for this differential is clearly the same as that profferred, in the previous section, for the yield spread between bills and other money market instruments.

Given the price differential between market bill prices and the Bank's price, any institution which accepted a bill and then immediately sought to rediscount this bill with the Bank of England, would incur a capital loss. In practice, therefore, acceptors have had an incentive *not* to rediscount their bills with the Bank, selling them into the market instead. In any event, as noted in chapter 4 the Bank does not purchase bills until seven days have elapsed since the day of their acceptance, which means that all eligible bills must be funded for at least seven days before they can be sold to the Bank. However, the existence of an excess liquidity premium on eligible bills – namely the large yield spread between eligible bills and other money market assets between 1990-1992 – implies that there has not been an incentive to hold bills either, given that they could not be profitably funded. Indeed, aside from the fact that bill holdings contribute towards satisfying banks' liquidity requirements, there has been little incentive for acceptors or any other financial institution, such as a discount house, to hold any bills at all in recent years.

6.5.2 The ownership of eligible bills

Who, then, holds the outstanding stock of eligible bills? Figure 6.13 indicates that the Bank of England, through its regular bill purchases, has owned an increasing share of the overall stock of bills and at the end of 1993 owned over 50% of all sterling eligible bills outstanding. Discount houses, traditionally the largest non-official holders of bills, have held a declining share of the stock of bills, and now own less than 10% of non-official holdings. In contrast to the houses, and as shown in figure 6.14, the clearing banks have become the largest non-official holders of bills, now regularly owning over 90% of non-official holdings. Within this group, the "big four" clearing banks are the major holders and account for over 80% of the groups' holdings. Note that this is a fairly recent development, with the major increase in clearing bank holdings occurring after the middle of 1990.

This development in the ownership distribution of the stock of non-official bill holdings has had several, interrelated consequences. First, the increase in the bill holdings of the larger clearing banks has undermined the role of the discount houses as the official counterparties to the Bank of England. Instead of acting as principals, the discount houses now operate largely as passive agents in the operations of the Bank, "passing through" the bill offers of the clearing banks to the Bank on most occasions. Clearly, therefore, the timing and magnitude of the daily bill sales to the Bank now depend to a significant extent on the portfolio decisions of a few institutions. In itself, this is not a remarkable situation, since these decisions have been the domain of a relatively small number of discount houses for many years. Coupled with the fact that these large bill portfolios incur an opportunity cost in foregone revenue, however, this situation implies that the clearing banks face incentives *to generate cheap funding through reserve management strategies which utilise the timing and magnitude of their bill operations with the Bank.*

6.5.3 Volatility

For example, a clearing bank may be in a position to sell enough bills to the Bank in its 9h45 operations to "remove" the entire money market shortage, thereby alleviating the aggregate reserve "pressure" that would otherwise result from the shortage. If the clearing bank itself remains short of cash (i.e. has a borrowing need in excess of the money market shortage) some other financial institutions will necessarily be long of cash (i.e. has excess reserves)[35] and will be seeking to lend, usually for a short period such as overnight. By delaying its demands for cash in the market, the clearing bank may succeed in lowering the cost at which it eventually borrows, as lenders will gradually lower the rate at which they are prepared to lend as the end of the trading day approaches. In this way, a clearing bank is able to fund its large bill portfolio profitably because it can deploy this portfolio to generate cheap short-dated funding sufficiently often to offset the lower yields on its bills[36].

Alternatively, a single clearing bank may be in the position where, by delaying its sales of bills to the Bank, a substantial amount of the money market shortage remains outstanding throughout the day. If the bank is also long of cash, other banks will necessarily be short of cash and will be seeking to borrow in the overnight market. This excess demand for funds will tend to drive up the price of overnight funds, as borrowers become increasingly willing to pay a higher overnight rate as the end of the day approaches.

Obviously these sorts of funding/timing strategies are not without risks, particularly since the aggregate cash position of the banking system is not known with certainty. Hence an expected intraday movement in short-term

[35]This follows because the aggregate reserves of the banking system are close to zero after the money market shortage has been removed. See Schnadt (1994a).

[36]This strategic behaviour on the part of clearing banks is discussed in greater detail in chapter 7. Note, however, that this behaviour is *not* collusive, as clearing banks compete vigorously for funds in the money market. Clearing banks are, however, usually large net takers of short-term wholesale funds. Strategic behaviour on the part of one clearing bank which results in a lower overnight rate, therefore, may indirectly lower the funding costs of another clearing bank.

rates does not always materialise. Similarly, clearing banks are, to some extent, constrained by their existing bill portfolio and their cash position and thus they cannot simply pursue these strategies at will. Nevertheless, the breakdown of the timing of the Bank's operations (in chapter 4) – which indicates that large shortages are sometimes removed entirely in the morning or, alternatively, that they sometimes remain until the money market closes – suggest that these strategies are deployed. A more obvious indication, perhaps, is the volatility in short-dated sterling interest rates, which is examined more closely in the following chapter. Whilst the funding strategies of the clearing banks are not the only cause of volatility in short sterling rates[37], they are certainly a prime suspect. These strategies, moreover, have had a tendency to create periodic cycles of "feast or famine" in the bill market.

6.5.4 The secondary bill market

These cycles are promulgated as follows. Although the clearing banks are the largest individual acceptors in the UK, they only accept one-fifth of all eligible sterling bills. Since they routinely hold over three-quarters of all eligible bills, the clearing banks have become large net buyers of bills. Most other accepting banks, being unable to fund their bills profitably, are thus usually net sellers of bills into the market. The periodic attempts by clearing banks to generate lower short-dated interest rates, as described above, have two effects on this scenario. Accepting banks (and discount houses) temporarily become able to fund their bills at a profit (since short-term rates are temporarily low) and thus sell fewer bills into the market. Secondly, short-term borrowers temporarily switch from issuing bills to borrowing unsecured funds in the money market. This reduces the overall flow of bills to the clearing banks, whose portfolios become depleted due to their regular bill sales to the Bank. As bills become increasingly "scarce", bill rates are bid down whilst short-term money market rates begin to tighten, often "spiking" to high levels as

[37]Reserve requirements elsewhere are generally higher than those of the UK, and reserve maintenance periods are longer, which will tend to insulate short-term rates from movements in aggregate reserves.

large money market shortages are removed with increasing difficulty. This in turn induces a greater flow of bills into the market, as holders of bills can no longer fund them profitably and borrowers again switch to bill finance.

These market dynamics are a far cry from a "conventional" picture of secondary market activity, such as may be found in the sterling CD market (or, say, in the US Treasury bill market). Eligible bills are *not* held by a wide group of institutions, and are *not* employed as liquid assets in the conventional[38] sense, i.e. as an asset which is purchased when an institution has excess reserves, or sold when an institution must cover a reserve deficit. Instead, secondary market activity is essentially *unidirectional*, with (non-clearing bank) acceptors selling bills to the clearing banks (and, to a limited extent, to the discount houses). Even this, unidirectional, activity is limited, since the clearing banks themselves will not usually sell their acceptances into the market, but will "swap" acceptances instead, since this avoids the bid-ask spread.

Thus, whilst the primary bill market is clearly attractive to borrowers, since they can obtain funds at below market rates, the secondary bill market is essentially redundant. The classic function of this secondary market, namely reserve redistribution, has long been carried out in the short-term, unsecured deposit market, which is much deeper and more liquid. Unfortunately, this is also where those banks who face the funding and rediscounting costs associated with eligible bills must, and do, attempt to recover these costs. Consequently, that market in which reserve redistribution does take place, is also rendered less efficient.

[38]The conventional or classical definition of a liquid asset is an asset which is *actually* employed to manage cash on a day-to-day basis; it is hard, otherwise, to think why liquidity would be such a valuable attribute. Of course, financial institutions also consider as liquid those assets which *could* be readily exchanged for cash in the event of some unforeseen contingency.

6.6 Concluding remarks

In order to provide reserves to the banking system on a regular basis, the Bank of England has, virtually since it opened its doors, elected to purchase private sector assets, namely bills of exchange. Such bills were, originally, a natural choice for this purpose, as they were the only short-term financial assets which existed in any quantity. Having chosen to purchase these assets, however, the Bank has also had to regulate their issue for prudential reasons, to contain its credit exposure to individual banks or firms.

One way in which the Bank has regulated the issue of eligible bills is via accepting limits. And, whilst the notional constraints placed by these limits on the overall supply of bills has never been reached, let alone approached, this chapter has outlined how, in practice, the supply of eligible bills has recently been constrained by these limits – due to circumstance rather than to design. It is difficult, otherwise, to explain the yield on eligible bills, which has fallen, for prolonged periods, below the level suggested by the liquidity premium due to their eligibility status.

The Bank did not always purchase bank bills. During and between the two world wars, the Bank routinely bought Treasury bills, whose supply had far eclipsed the stock of bank bills. Obviously, this made regulations in the form of accepting limits unnecessary for prudential purposes. However, the pursuit of overfunding by the government dramatically changed this situation:

> "[During the 1970s] increased attention was given to the control of the stock of money, but a buoyant trend developed in bank lending to the private sector. At least in the short-run, unduly rapid monetary growth could most easily be moderated by selling government debt to the non-bank private sector on a larger scale than hitherto in relation to the Government's borrowing requirement [i.e. overfunding]. This combination of circumstances created persistent underlying shortages in the money market. The Bank responded in the first instance by reducing the quantity of Treasury bills on offer at the weekly tender. Even so, the Bank still found it necessary to buy Treasury bills - often more than were being issued. Inevitably, the proportion of these in the banking system's portfolio declined and the Bank was then *obliged* to operate increasingly by purchasing bank bills... " (Bank of England 1982, p.88; emphasis added).

The Bank's resumption of bank bill purchases, therefore, was a reaction to the circumstances created by overfunding. The advantage of overfunding was that it could – mechanically – lower bank deposits and so too the measured money stock, thereby assisting the monetary authorities in meeting their explicit broad money targets (introduced in 1976). This mechanistic reduction was possible because most public sector accounts were – and still are – held with the Bank of England and overfunding simply decreased private bank deposits by increasing public sector deposits at the Bank of England. This flow of deposits (reserves) from the private to the public sector, however, had the disadvantage of generating larger money market shortages, which then had to be offset by increased purchases of *private sector* assets (i.e. bank bills) by the Bank. The same conditions which made overfunding possible, therefore, similarly rendered it impossible that the shortages so created could be met through an increase in the issue and purchase of Treasury bills by the Bank.

Even though overfunding as a policy objective was abandoned in 1985/86 (see section 1) the phenomenon of overfunding has continued to accompany government funding in recent years. This was due to the funding policy pursued by the government, which excluded gilts sold to banks and building societies from the definition of funding. Consequently, the Bank has continued to purchase large quantities of bank bills. Although this definition was scrapped in 1992/93, a modification to the funding rule in 1990/91 had excluded Treasury bills from the definition of funding. According to the Treasury,

> "... because the underlying intention is that the public sector does not finance itself by creating money, the authorities need to ensure that funding is not concentrated in very short-term debt, where the public liabilities have the attributes of money and liquidity .... [Thus] Treasury bills, which have such characteristics, no longer count as funding" (Financial Statement and Budget Report 1990-91, HMSO, p.25).

This position is inconsistent for at least two reasons. First, eligible bank bills and Treasury bills are, in terms of liquidity, perfect substitutes: both may be sold to the Bank and both satisfy banks' liquidity requirements.

Secondly, all gilt-edged securities, regardless of their *initial* maturity, will eventually have an unexpired maturity which makes them indistinguishable from a Treasury bill (except, perhaps, for a coupon). They, too, count towards banks' liquidity requirements. If the funding policy of the Treasury were consistently applied, therefore, it should similarly exclude bank bills and gilt-edged securities with a short unexpired maturity from the definition of funding, which it clearly does not[39].

Inconsistent or not, the current funding policy of the government is that Treasury bills are excluded from the definition of funding. Until this definition is changed, it is unlikely that their issue will increase enough for them to play a significant role within the sterling money market or within the Bank's money market operations. Consequently, the Bank is likely to rely increasingly upon repurchase agreements involving gilt-edged securities as a means of supplying reserves to the banking system. Not only are such repos more flexible in terms of their maturity – a feature whose advantages were outlined in the previous chapter – but they also involve no credit exposure to banks or firms and thus do not require additional regulation or monitoring by the Bank.

It is not inconceivable, therefore, that the Bank may eventually curtail altogether its purchases of bank bills. These instruments will then cease to be considered as primary liquid assets and will cease to bear a liquidity pre-

[39]This is exemplified by the following exchange during the proceedings of the Treasury and Civil Service Committee during 1993:

> 303. Why do you not allow sales of Treasury bills to count as funding?
>
> (*Mr Lamont*) As I have said, I have the instinctive apprehension about monetary financing.
>
> 304. Unless I am mistaken, if the banks buy short-term government debt in the form of Treasury bills this does not count as funding. But if they buy gilts within six months or less to maturity this presumably does count as funding. What is the reason for this distinction?
>
> (*Mr Lamont*) I am sure we should not go in the direction of funding more and more of our debt on a short-term basis. You can argue about where we eventually draw the line (*Fourth Report from the Treasury and Civil Service Committee*, 1992-93, HC578, HMSO, p.45).

mium (other than that which is due to their being tradable on a secondary market). This will affect both the firms which rely upon bill finance as a cheap source of short-term funds, as well as the financial institutions which hold bills to satisfy their liquidity requirements. Borrowers will switch to the cheapest alternative which, in the case of prime names, may be commercial paper, which avoids the acceptance fee on a bill. Sterling commercial paper programmes are thus likely to become more actively used. Other borrowers, however, may well find that bills remain an attractive means of raising finance. Financial institutions, on the other hand, are likely to hold larger amounts of gilt-edged securities for liquidity purposes. Indeed, this demand for additional liquid assets may, finally, provide the impetus for increasing the Treasury bill issue. Lastly, the cessation of eligible bank bill purchases by the Bank may see the restoration of more orderly conditions in the overnight market, and very short-term sterling interest rates, the current behaviour of which is examined in the following chapter.

6.7 Tables and figures

Figure 6.1
Eligible bank bills held by the Bank of England (£mn)

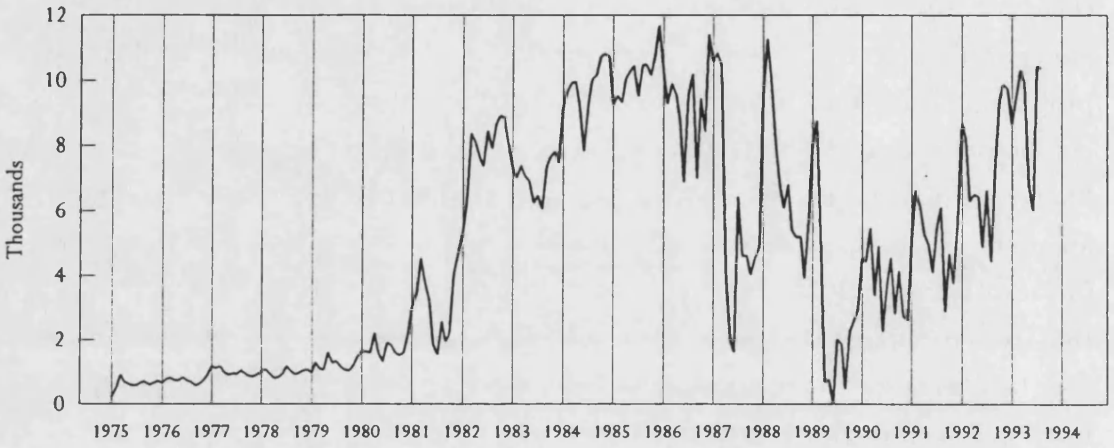


Figure 6.2
Gross annual assistance by the Bank of England (£bn)

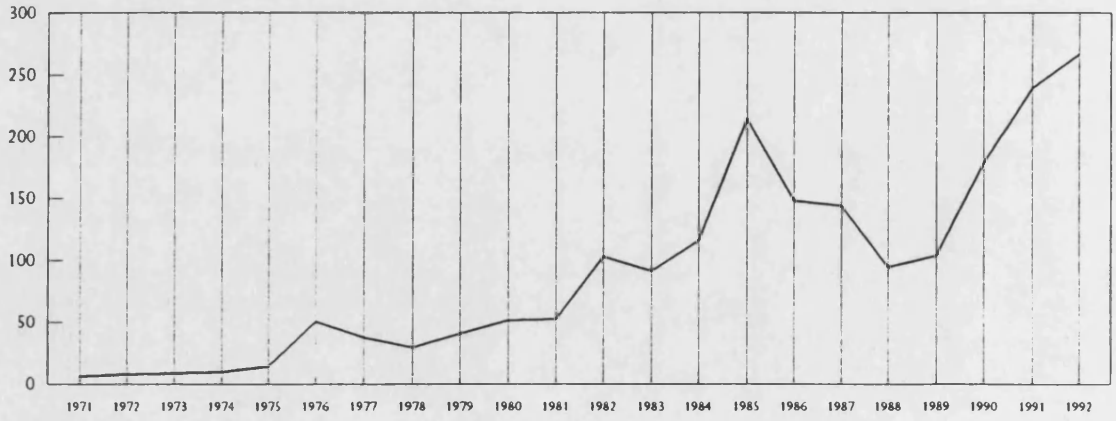


Figure 6.3
LIBOR/Treasury bill yield spread (one month)

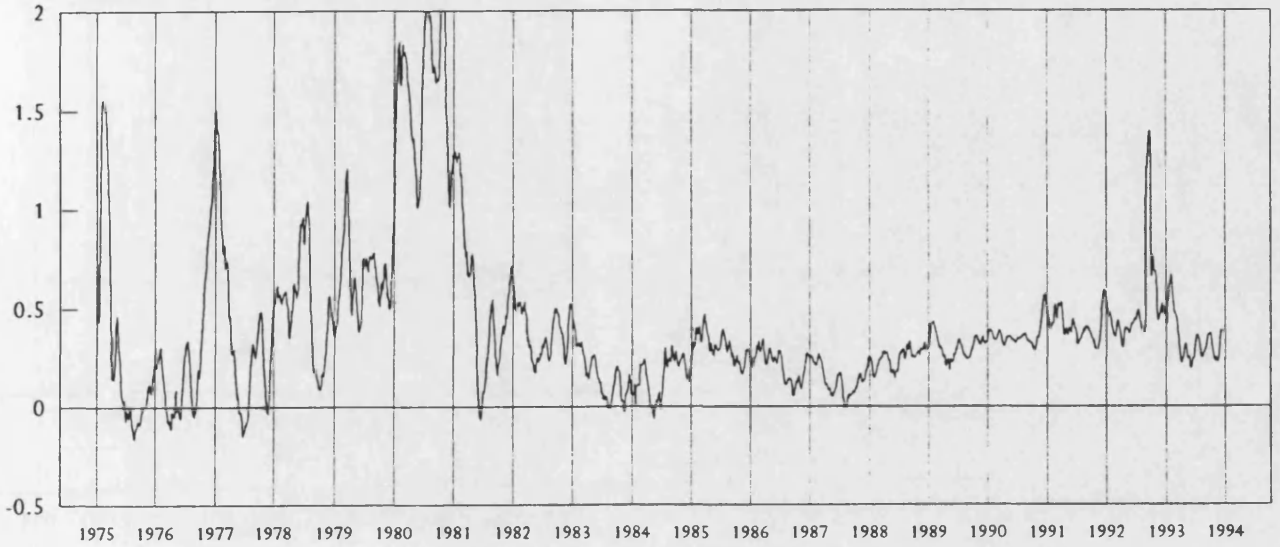


Figure 6.4
LIBOR/bank bill yield spread (one month)

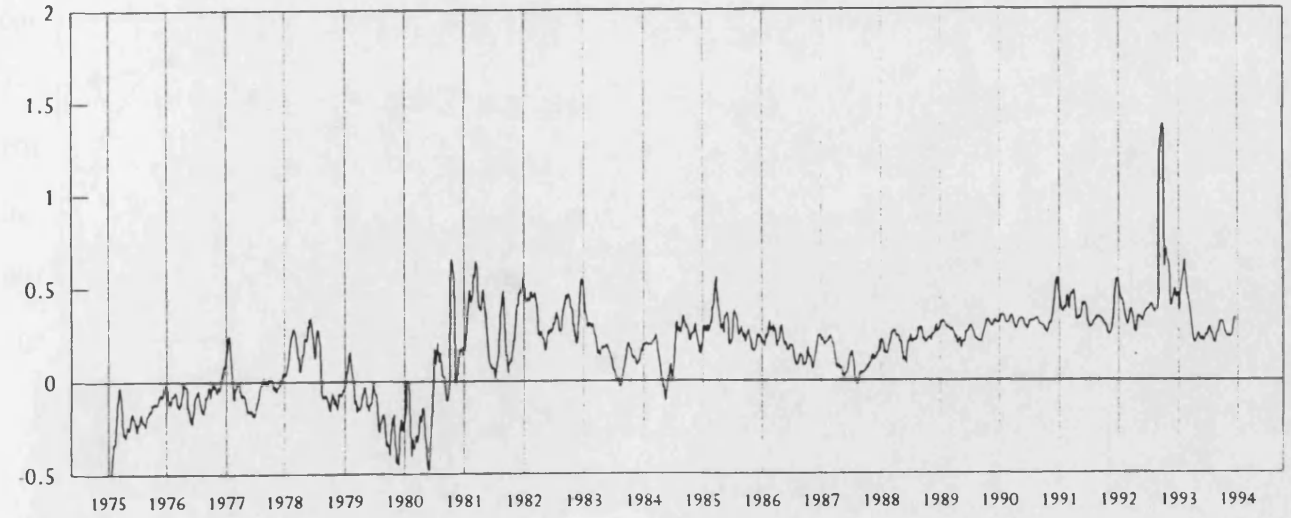


Figure 6.5
Treasury bill/bank bill yield spread (one month)

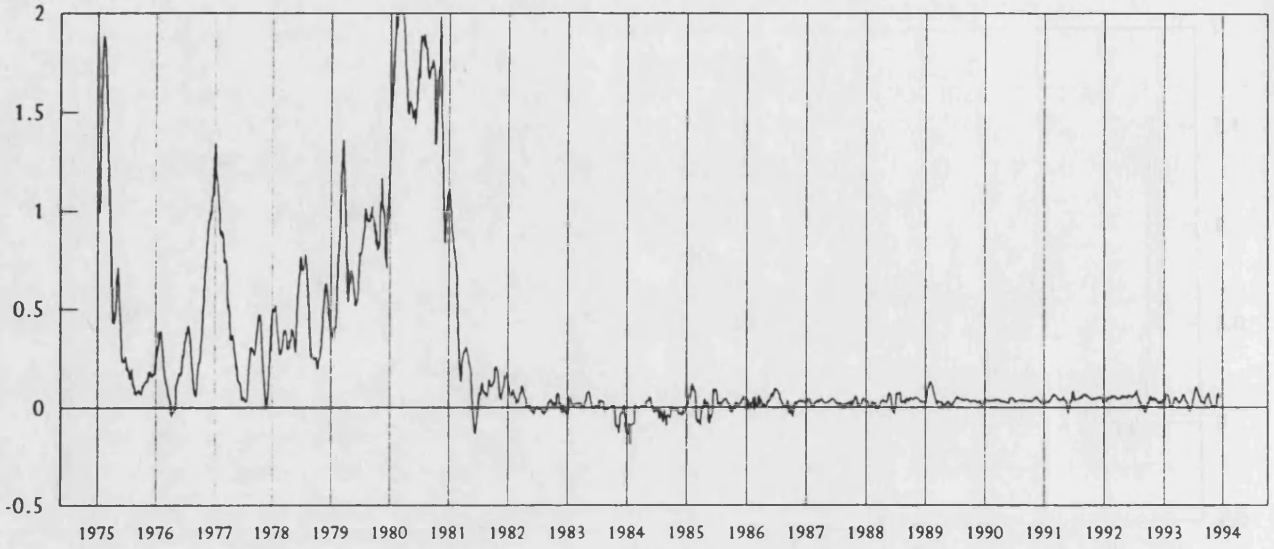


Figure 6.6
Average annual LIBOR/bank bill yield spread and gross annual assistance

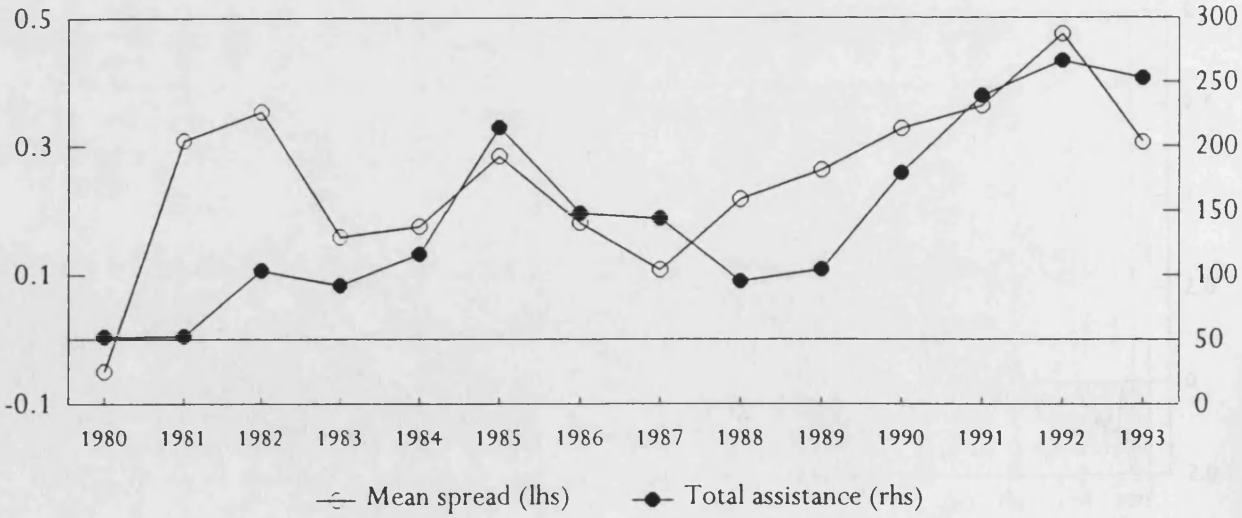


Figure 6.7
Sterling acceptances:
total outstanding (£mn)

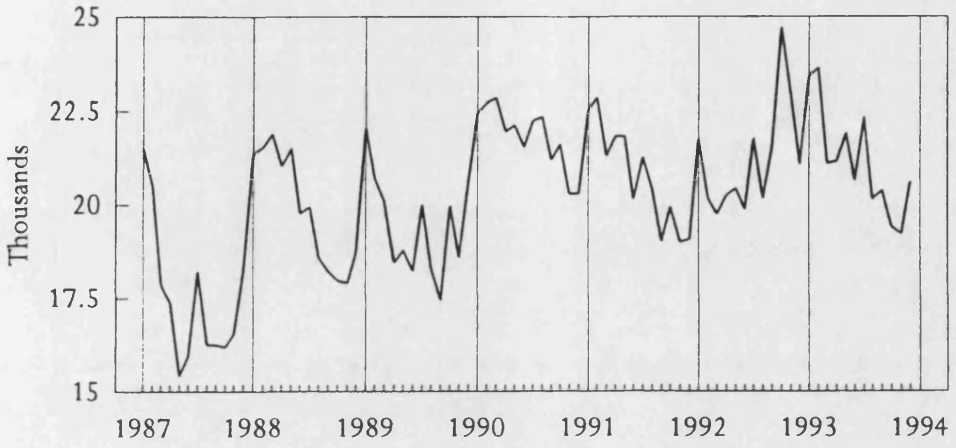


Figure 6.8
Sterling acceptances and industrial production:
annual growth rates (%)

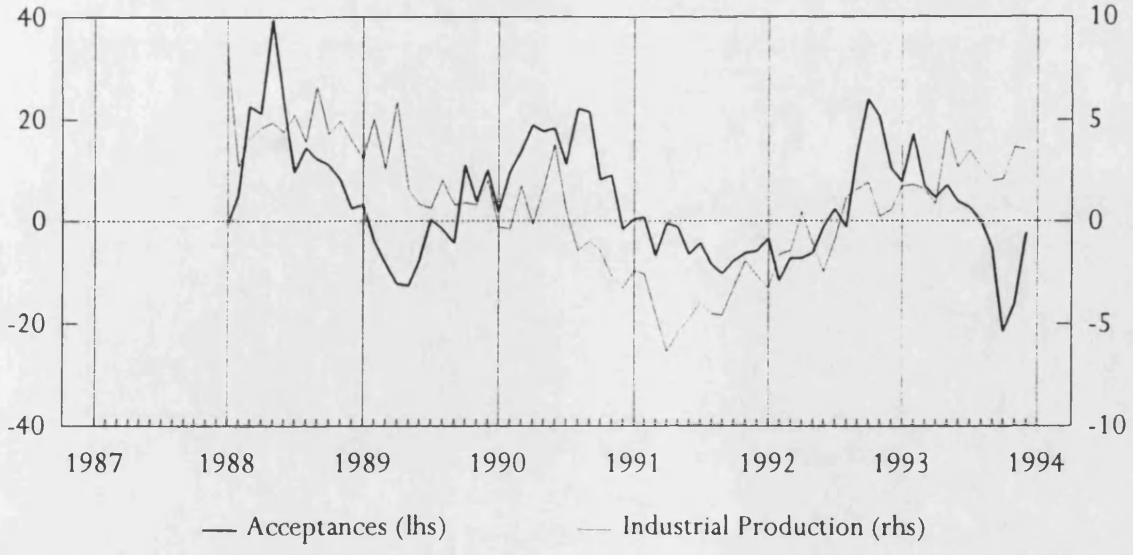


Figure 6.9
Sterling acceptances:
share of total bank lending to the UK private sector (%)

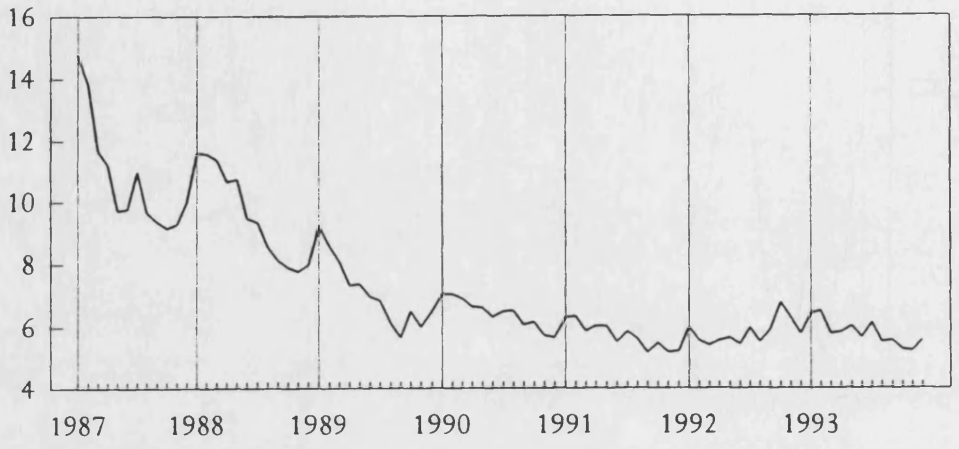


Figure 6.10
Sterling acceptances:
share accepted by clearing banks (%)

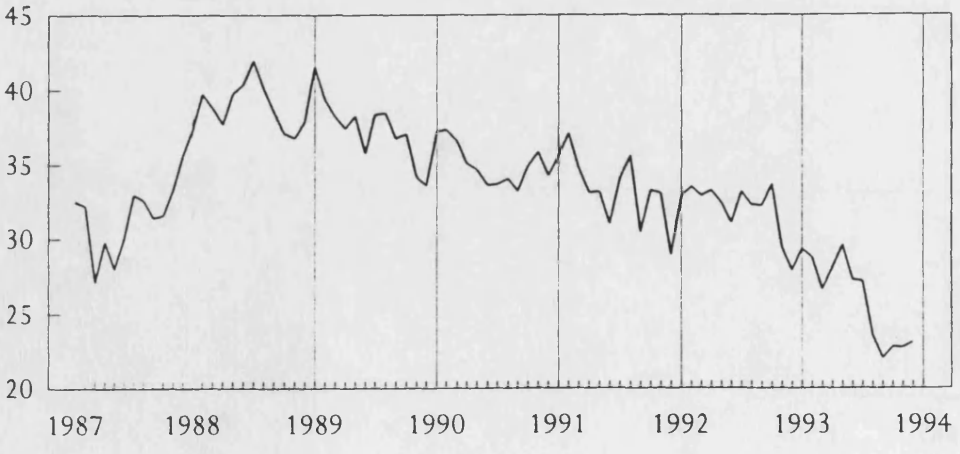


Figure 6.11
Market discount rate – band one stop rate (one month)

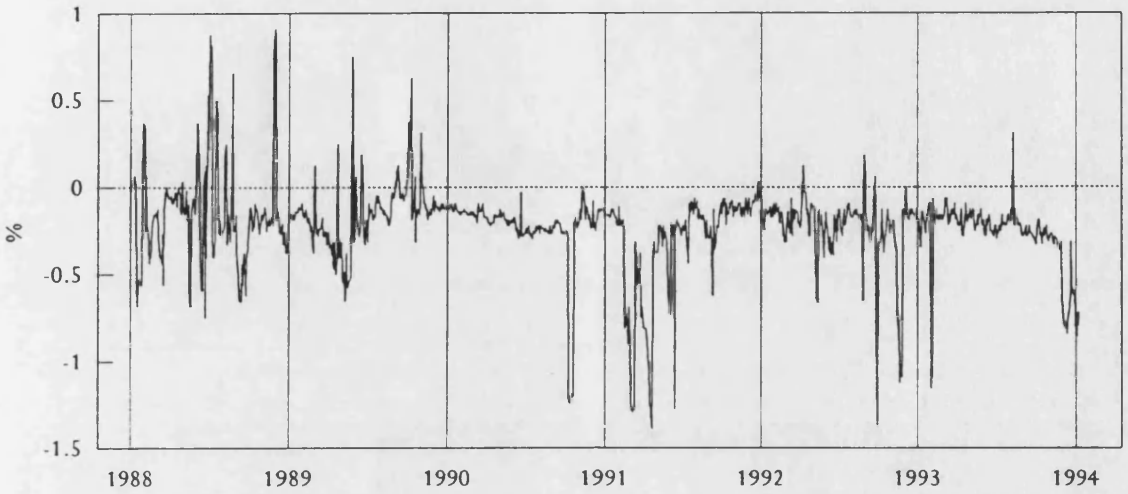


Figure 6.12
Bill yield (+fees) – LIBID (one month)

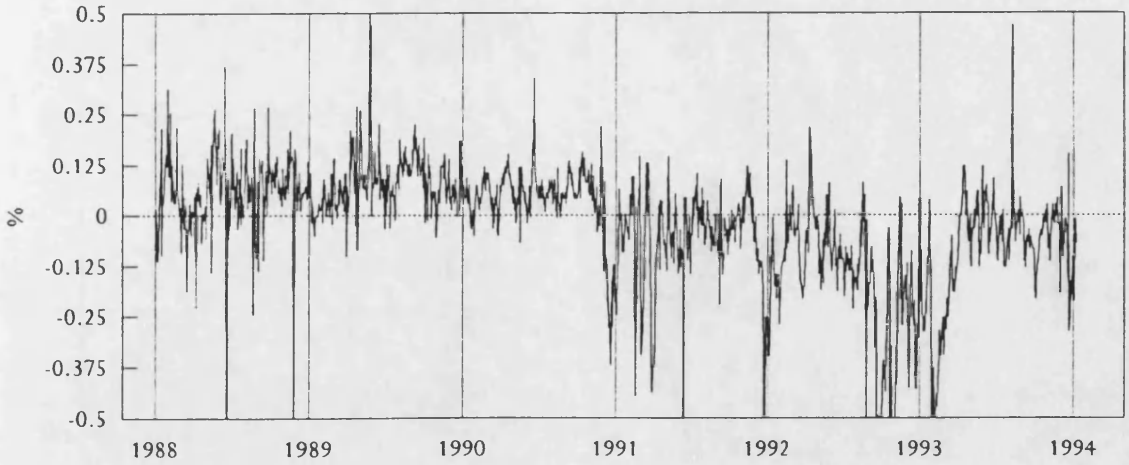


Figure 6.13
Bank bills:
total holdings (£mn)

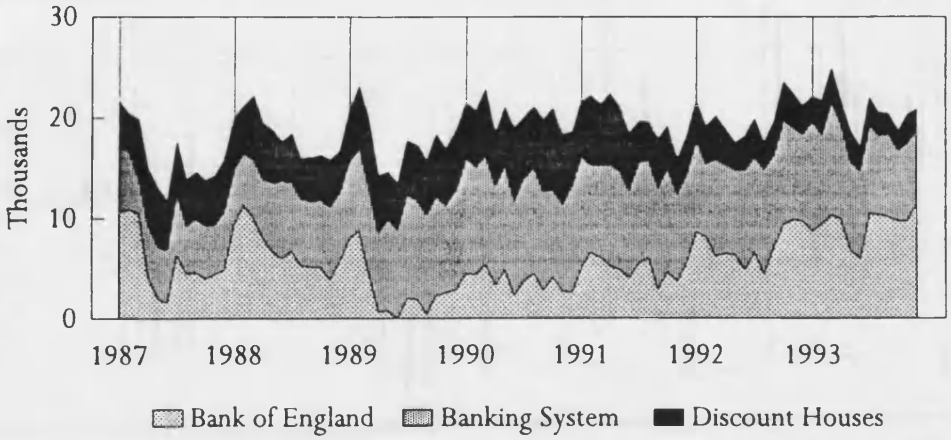


Figure 6.14
Bank bills:
non-official holdings by clearing banks (%)

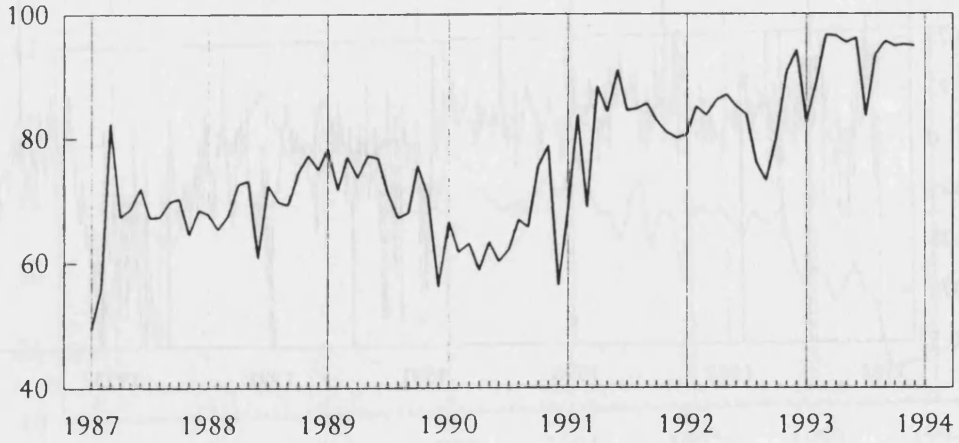


Figure 6.15
Annual acceptance growth (%):
clearing banks versus other eligible acceptors

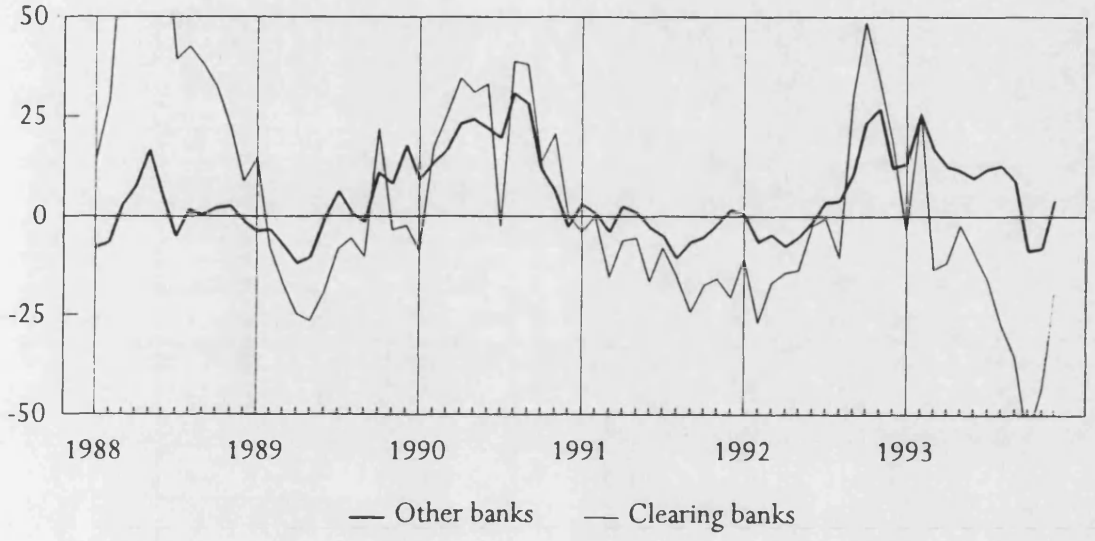


Table 6.1

The LIBOR/eligible bill yield spread 1980-1993

Year	Mean	Variance	Volatility	$H_0 : \mu_T = \mu_{T-1}$?
1980	-0.05	0.51	76.06	n/a
1981	0.31	0.27	0.33	n/a
1982	0.35	0.20	0.15	yes
1983	0.16	0.13	0.11	no (1%)
1984	0.18	0.17	0.17	yes
1985	0.29	0.21	0.17	no (1%)
1986	0.18	0.14	0.13	no (1%)
1987	0.11	0.12	0.12	no (1%)
1988	0.22	0.12	0.09	no (1%)
1989	0.26	0.07	0.06	no (1%)
1990	0.33	0.08	0.03	no (1%)
1991	0.36	0.14	0.09	no (1%)
1992	0.48	0.73	0.06	no (5%)
1993	0.31	0.14	0.08	no (1%)

Source: Datastream

Table 6.2

Largest Users of Acceptances

Sector (CSO classification)	Acceptances	
	£mn Outstanding (% of Total)	% of Total Bank Lending to Sector
Other finance	4907 (27)	10
Food, drink, tobacco	2737 (15)	29.7
Wholesale distribution	2326 (13)	15.9
Other manufacturing	1363 (7)	12.2
Leasing companies	902 (5)	3.5
Other retail distribution	807 (4)	7.3
Electrical engineering	785 (4)	15.9
Business services	778 (4)	2.5
Oil and natural gas	602 (3)	8.5
Other energy industries	558 (3)	16.7
Motor manufacturing	523 (3)	25.6
Chemical industry	520 (3)	16.7

Source: CSO Financial Statistics.

Chapter 7

The Behaviour of the Sterling Overnight Interest Rate

This chapter investigates the behaviour and the determinants of the shortest maturity interest rate in the sterling money market – the overnight rate. In the US the overnight rate is known as the federal funds rate and is generally considered to be an important indicator of the stance of monetary policy. Consequently, it is one of the most highly scrutinised variables in the financial markets and in the wider economy. The federal funds market has also been the subject of numerous academic studies, which have examined the institutional structure of the market and the banking system more generally, and have traced the effects of this structure on the behaviour of the fed funds rate[1].

The sterling overnight rate, in contrast, is not "targeted" by the Bank of England in the same manner that the federal funds rate is targeted by the Fed. The overnight rate is thus not considered as an important information variable as regards monetary policy in the UK and is not closely monitored by anyone aside from those institutions which are active in the money market. Indeed, apart from two published papers by the author (Schnadt 1994a, 1994b) and a recent working paper by Ayuso, Haldane and Restoy (1994),

[1]See Minsky (1957), Poole (1968), Ho and Saunders (1985), Spindt and Hoffmeister (1988) and Brunner and Lown (1993).

no analytical studies of the sterling overnight funds market and of the determinants of the overnight rate have been conducted.

Why might such issues be important? One reason has been put forward by Poole (1990):

> "In debating how best to stabilise the economy, a tendency exists to dismiss day-to-day operating procedures as irrelevant .... But ignoring the very short run issues is a mistake. The analysis of monetary policy should proceed in the same way as the analysis of possible schemes for successful speculation in the securities markets. A common research design in finance is to specify what hypothetical speculators would do and when they would do it. Investors cannot buy shares at the monthly average price without buying every day. Thus, in finance research, our hypothetical speculators buy and sell at actual market prices and pay actual transactions costs. The analysis of monetary policy should proceed in the same way. (p.40)

Even though the Bank of England does not attempt to influence the overnight interest rate directly, it certainly influences this rate indirectly. This was clearly demonstrated in chapter 5, where the term structure of sterling interest rates was derived, and the behaviour of very short-term interest rates was seen to depend crucially on the interest-rate expectations of money market participants. The level of overnight rates, therefore, is fundamentally linked to the expected path of official interest rates and may contain information about these expectations, at least to money market participants.

A more important reason for studying the overnight funds market, however, pertains to *efficiency*. It has been argued in previous chapters that the money market, especially the overnight funds market, is vital to banks for liquidity management purposes. Activity in the sterling overnight market is substantial – easily £15 *billion* daily – and is many times larger than the volume of trading in other sterling financial markets such as the London Stock Exchange[2]. If the overnight interest rate is extremely volatile as compared with other interest rates – which this chapter will confirm – then this may be detrimental to the efficiency of the overnight funds market in

[2]Daily turnover on the stock exchange rarely exceeds £2 billion.

facilitating reserve management[3]. Money market participants, for example, will have to devote additional resources to predicting rates more accurately, or to seeking out alternative means of managing their liquidity. Studying the microstructure of the overnight market, therefore, and the determination of prices there, is essential to an understanding of its allocative efficiency, and may also provide insights into the possible design of alternative – more efficient – institutional arrangements.

It is not the intention of this chapter to develop a specific theoretical model of the sterling overnight market. Instead, it aims to provide a descriptive framework and, more importantly, to characterise empirically some of the factors which influence both the level and the volatility of the overnight rate. Whilst the level (or the mean) of the overnight rate and its determinants may be examined using a simple regression model, the volatility of this rate is more usefully studied by using recent developments in the empirical modelling of volatility.

Most economic models, when subjected to empirical estimation, are formulated as follows:

$$y_t = \underline{\theta}' \cdot \underline{x}_t + \varepsilon_t \,, \quad t = 1, \ldots, T \,, \tag{7.1}$$

where y_t is the variable whose behaviour is to be explained, $\underline{x}_t$ is a vector of explanatory variables, $\underline{\theta}$ is a vector of coefficients, and ε_t is a random component, or error. Since the explanatory variables are given, and are thus non-stochastic (non-random), statements about the relationship between these and the dependent variable rely crucially on the statistical properties of the error term, which forms the stochastic part of the model. By far the most common assumptions, and the basis of classical (OLS) regression, are that $\varepsilon_t \sim N(0,\, \sigma^2)$.

These assumptions, particularly the latter assumption concerning constant variance, essentially ensure that the distribution of the disturbances is

[3]The dramatic increase in the volatility of federal funds rate between 1979 and 1982, for example, led to a spate of surveys and studies of the effects of this volatility; see Evans (1981), Kasriel (1981) and Rasche (1985).

unchanging across observations (i.e. over time if the data are time series), and that the dependent variable y_t has the conditional distribution

$$y_t \mid \underline{x}_t \sim N(\underline{\theta}' \cdot \underline{x}_t, \sigma^2).$$

Unsurprisingly, the statements of economic interest are then typically about the conditional mean of the dependent variable rather than its variance. Of course, the assumption of constant variance (of the error) may not be valid in practice, in which case conventional regression techniques may give misleading results[4]. This problem – heteroscedasticity – is known to pervade cross-section data, but is also prevalent in the distribution of financial prices, of which short-term interest rates are an example. A substantial literature has emerged recently which examines whether the variance of the random component in (7.1) exhibits (a particular form of) time-variation, a phenomenon which has been called autoregressive conditional heteroscedasticity or, more simply, ARCH (see Engle 1982). And, whilst conventional estimation techniques are then more or less maintained, the statements of economic interest are usually about the variance, or volatility, of the dependent variable and the nature of its time-variation.

These developments in the literature suggest a convenient way of organising the empirical examination in this chapter. Section 1 presents a description of the overnight market and of the recent behaviour of the overnight rate, whilst section 2 discusses the factors which may influence this behaviour. Section 3 then examines the determinants of the *level* of the overnight rate; that is, a model of conditional mean is estimated using conventional OLS regression. Thereafter, section 4 examines the *volatility* of the overnight rate; that is, a model of the conditional variance is estimated using ARCH regression. Finally, some concluding remarks are made in section 5.

[4]In particular, the estimators $\hat{\underline{\theta}}$ become inefficient; see Greene (1993).

7.1 The overnight funds market

7.1.1 What are overnight sterling funds?

The market for overnight funds is by far the most active component of the sterling money market. Here, banks, other financial institutions and corporations borrow (lend) unsecured sterling funds – for same day settlement – with a maturity of one day. This "day", however, is not defined in terms of real time, e.g. 12 or 24 hours, but refers to the next *trading* day. The end of one trading day is marked by the cut-off time at which it is no longer possible to trade funds for same day value. In the UK, this is governed by the rules of the large value payment systems discussed in chapter 4, and is normally 15h10 for non-clearing banks and 15h30 for clearing banks.

This has two noteworthy consequences. First, although banks may continue to trade overnight funds after 15h30 on the same calendar day, these are already considered as having taken place on the following trading day. In principle, therefore, there is no difference between an overnight trade conducted after 15h30 on one calendar day and before 15h30 on the next calendar day. Normally, however, participants in this market will trade most actively on the morning and early afternoon of a trading day. Secondly, overnight funds traded before 15h30 on a Friday have a maturity of *three* days rather than one day, given that the markets are closed on Saturdays and Sundays. Similarly, overnight funds traded before bank holidays also carry a longer maturity than one day. If, for instance, both Friday and Monday are bank holidays – this is commonly the case at Easter – then an overnight loan made on the Thursday becomes a five day loan since it will only be repaid on the following Tuesday. Effectively, therefore, the shortest maturity at which it is possible to borrow or lend funds is not always one day, an institutional feature which is discussed further below.

The average size of trades in the overnight market naturally differs according to the needs of the borrowing or lending institution, but is usually around £5 million. Obviously smaller banks and corporations will be able to lend and borrow smaller amounts, whilst the larger clearing banks may make

single trades as large as £1 billion. Quoted spreads are usually $\frac{1}{8}\%$, although these are usually only firm for interbank trades. Corporate participants typically face a slightly wider spread of about $\frac{1}{4}\%$.

7.1.2 UK reserve requirements

There is an important institutional difference in the UK between clearing (settlement) banks and all other money market participants. It is that the latter all utilise accounts at clearing banks for purposes of trading in the money market – indeed, for any trading, financial and non-financial – whilst clearing banks themselves use operational (or reserve) accounts at the Bank of England. The reserve accounts of the UK banking system may thus be thought of as a "two-tier" structure, in which only the clearing banks face an official reserve requirement. As noted in chapter 4, the clearing banks are obliged, by the Bank of England, to achieve at least a zero *closing* (or overnight) balance in these operational accounts[5]. Non-clearing banks, on the other hand, do not face any official reserve requirements, although their reserve management problem is nonetheless similar to that of the clearing banks. The interest return on funds left overnight in transaction (current) accounts will be at a substantial discount to the market overnight rate, whilst the cost of funds when these current accounts are overdrawn overnight will be at considerable premium above the market overnight rate. Thus customers of a clearing bank will not want to have either a positive or a negative closing balance in their current accounts and will use the overnight funds market to lend any surplus balances, or to borrow deficit balances[6].

Of course, the activity of the customers of a clearing bank as they withdraw funds from, or deposit funds into, their current accounts naturally alters the cash (reserve) position of that clearing bank. Since this customer activity is not perfectly predictable, clearing banks remain inherently uncertain over

[5]Recall that clearing banks aim to end the day with a small, positive *target* balance each day.

[6]On occasion, however, the overnight rate may rise sufficiently for customers to utilise their overdraft facilities *voluntarily*.

the likely evolution of their intraday cash position, and hence of their closing reserve position. For instance, a clearing bank may have bought or sold funds in the overnight market so as to "square" (i.e. balance) its own reserve position, only to receive further customer deposits (withdrawals) which then leave its operational balance positive (negative). These customer transactions thus necessitate further trading in the overnight market by the clearing bank, as it must sell (buy) funds to square its position. It is unlikely, therefore, that a clearing bank will be able to end the trading day (i.e. 15h30) without either a positive or a negative reserve position, which may be quite substantial (e.g. £250 million). It is for this reason that clearing banks – and not other banks – are able to engage in a "final" round of lending as described in chapter 4.

Given that the Bank of England imposes a daily reserve requirement on the clearing banks, it must then ensure that the reserves are made available to satisfy this requirement. Earlier chapters have argued that this reserve provision is the primary purpose of the Bank's daily money market operations. Indeed, if the Bank did not make reserves available *during* the day, it would have no choice but to make them available at the *end* of the day, probably in the form of an overnight overdraft to the clearing bank (or banks) whose operational account was in deficit. The relationship between the Bank and the clearing banks is, in this sense, directly analogous to the relationship between a clearing bank and its customer. When a customer overdraws their transaction account by more than the negotiated limit, the clearing bank more-or-less *automatically* lends funds to the customer. Naturally, bank customers will try to avoid this situation, as they will be charged a penalty rate by the clearing bank and may be granted less (or more costly) credit in the future. Clearing banks, however, cannot avoid incurring a deficit in their operational balances if *aggregate* reserves are insufficient to meet their reserve requirements. This will be the case whenever there is a money market shortage – which was defined in chapter 4 as a reserve shortfall of the clearing banks below their reserve targets. The shortage thus represents the level of reserves which must be provided by the Bank of England during the course of the day to ensure that the clearing banks meet their reserve targets.

7.1.3 The impact of reserve accounting and central bank operations on the overnight rate

The reserve maintenance period in the UK is extremely short – one day. Unlike their overseas counterparts in the US or on the Continent, therefore, who typically face positive reserve requirements with reserve maintenance periods of as long as two weeks or one month, UK clearing banks do not have any scope for reserve *averaging*. It is a well known that such averaging tends to reduce the day-to-day variability of overnight interest rates, since banks can choose to defer their overnight borrowing (lending) if they believe the current cost of overnight funds is high (low) relative to its expected level on subsequent days (Poole 1968). Consequently, one would expect the day-to-day variability in overnight rates to be higher in the UK than elsewhere. This is certainly borne out in practice. For example, Kasman (1992) and, more recently, Ayuso, Haldane and Restoy (1993) find that sterling overnight rates are significantly more volatile[7] on a day-to-day basis than US or Continental overnight rates.

It is also well known that, under the reserve accounting employed elsewhere (e.g. the US or Europe), overnight interest rates become more volatile as the end of the maintenance period approaches, essentially because banks then have less flexibility in accommodating any deviations from their average reserve targets. Spindt and Hoffmeister (1988) have studied this phenomenon in the fed funds market, noting that "because of the discrete character of reserve maintenance measurements and settlement, the variance of the [overnight] funds rate is higher towards the end of each business day and is highest near the end of settlement days ... [Furthermore] funds rate variance may be higher on days prior to holidays and on Fridays ..." (p.412). They present evidence – corroborating the results of an earlier study by Dyl and

[7]Volatility in these comparisons may be defined in a number of ways. Kasman, for instance, chooses to use the deviation of daily rates from a centered moving average, whilst Ayuso *et al* use the conditional variance of the error from an ARCH regression. A third commonly used diagnostic, which is employed below, is the standard deviation of day-to-day changes in the overnight rate.

Hoffmeister (1985) – that the volatility of the federal funds rate increases over the maintenance period and that it is significantly higher before weekends and on the final day of the maintenance period. Similar effects in continental overnight rates, particularly those relating to greater volatility near the end of reserve maintenance periods, are reported by Ayuso *et al* (1993). These results suggest that, whilst the volatility of the sterling overnight rates on a day-to-day basis is clearly higher than that observed elsewhere, a more appropriate comparison might be with the behaviour of overnight rates on the *final day* of the maintenance period, as banks are then in a similar position to those in the UK.

Of course it is not *necessarily* the case that a particular structure of reserve requirements (i.e. low level, short maintenance period) *must* be associated with greater volatility in overnight rates. If, for example, the central bank stood ready to borrow or lend overnight funds at some stated rate at all times, the market overnight rate would be kept equal to the central bank's rate irrespective of the structure of reserve requirements. An important determinant of the behaviour of overnight rates under any set of reserve requirements is thus the *money market operations of the central bank*, which usually only occur at several discrete points over the maintenance period and which influence current and, more importantly, expected reserve availability. If reserve availability is expected to become greater (tighter), the current overnight funds rate will fall (rise), as those institutions which are expecting an average reserve surplus (deficit) attempt to lend (borrow) reserves.

This expectation of future reserve availability may itself be based upon the characteristics of the central banks' operations, such as their type, their timing and their magnitude. Feinman (1993a), for instance, undertakes a detailed analysis of the magnitude and the type of open market operation employed by the Federal Reserve and finds that these are systematically related to certain variables such as the aggregate expected reserve need, the aggregate cumulative reserve position and the deviation of the federal funds rate from its expected average. This in turn implies that market participants will associate particular actions by the Fed as *signals* of its underlying policy stance, prompting an immediate reaction in the fed funds rate when the Fed

deviates from its "normal" pattern of operations.

7.1.4 Intraday variation of the overnight rate

Although the reserve maintenance period in the UK is only one day, some of the above results regarding overnight rate behaviour should still be observable on an *intraday* basis. Specifically, the sterling overnight should exhibit intraday variation, becoming more volatile as the end of the day approaches.

Unfortunately, detailed intraday data on the overnight rate is not readily available. However, intraday data comprising four daily observations – 8h00, 11h00, 14h00 and 15h30 (close) – were obtained[8] for a period spanning the first five months of 1994 (i.e. over 100 observations, four times a day). Its properties are reported in figure 7.1 and tables 7.1 and 7.2. Figure 7.1 depicts four histograms of the spread between the overnight rate and the stop rate, one for each time of day, and shows a dramatic increase in variance later in the trading day. Table 7.1 confirms this result, indicating that the variance of the overnight rate was up to five times as high at the close of the day as it was during early trading. Table 7.2 qualifies this result further, and indicates that movements in the overnight rate were larger (in absolute terms) and had a larger variance, later in the day.

Whilst these results are encouraging in that they confirm a tendency which characterises overnight rates elsewhere, they also point to the difficulty of using ordinary daily data on sterling overnight rates for purposes of analysing the *day-to-day* behaviour of this rate. Such data, which is readily obtainable from the financial press (e.g. the *Financial Times* reports the closing rate) and other financial data sources (e.g. Datastream reports a midday rate), usually consists of a single daily observation of the overnight rate. The significant intraday variation in the overnight rate, however, suggests that the properties of these alternative time series will be considerably different, and may lead to qualitative differences in the inferences which are drawn about the determinants of both the level and the volatility of the overnight rate.

[8] I am indebted to Richard Pattinson (Barclays) for these data.

What is more, the level of the overnight rate does not convey any information about the quantity of funds which was traded at that rate. Thus some intra-day movements in the overnight rate may only be applicable to a relatively small volume of trading, leading again to the danger that incorrect inferences are drawn.

There are two approaches which, to some degree, circumvent this data problem. One approach is to examine time series of the daily highs and lows of the overnight rate. Parkinson (1980), for example, argues that a good approximation of the variance[9] of a continuously evolving random variable is given by a measure which employs the extremes of this variable over some time interval t (e.g. a trading day):

$$D_T = \frac{.361}{T} \sum_{t=1}^{T} (High_t - Low_t)^2 \tag{7.2}$$

where $High_t$ and Low_t are the daily high and low respectively and T is the period of interest. These extremes may be employed further to examine the moments of the overnight rate, without requiring observations of this rate itself[10]. This is because these extremes are known to fall into a class of asymptotic, or limiting, distributions, known as *extreme value distributions*, whose statistical properties have been well-documented (see, for instance, Gumbel 1958 or Johnson and Kotz 1970). Consequently, the actual distributions of the extremes may be estimated – using maximum likelihood methods – and these results used to make inferences about the behaviour of the distribution of the underlying variable, i.e. the overnight rate. Indeed, this approach, which is taken in a recent working paper by Dale, Haldane, Murphy and Schnadt (1994), may be used to assess the significance of various institutional variables on the parameters of the distributions of the extremes, permitting efficient inferences to be drawn. Unfortunately, the technical requirements of this approach are considerable and the results more difficult

[9]Or, more precisely, of the *diffusion constant*, which determines the variance of the displacement of the random variable from its starting position after a unit interval.

[10]Such an exercise is conducted by Hols and De Vries (1991) on extremal exchange rate returns.

to interpret in terms of their economic content. Another drawback is that the volume of overnight funds traded at the highest or the lowest rate is ignored, thus potentially overstating the effect of the explanatory variables on the overnight rate at which the bulk of funds is actually traded.

An alternative approach – which is pursued in this chapter – is to utilise overnight interest rate data which takes explicit account of the *volume* of trading which occurred at different intraday rates. Rate movements which are weighted by the amount of funds traded clearly provide a better reflection of the average, or effective, daily overnight rate faced by market participants. Whilst such a trade-weighted rate is officially calculated on a daily basis in the US by the Federal Reserve – which labels this rate the *effective* federal funds rate – no such official calculation is undertaken in the UK. An unofficial series, however, is calculated by Barclays, a large clearing bank which is an active participant in the sterling overnight market each day. This series, which comprises what may be referred to as the daily effective sterling overnight rate, was obtained[11] for the period January 1991 to May 1994 (938 observations) and forms the basis of all the empirical results reported in subsequent sections. Obviously, since it is not officially calculated, this data is only an approximation of the true effective rate and is thus subject to some measurement error. In the absence of better data, this is the most attractive alternative.

Even though the effective overnight rate is calculated as a trade-weighted rate, it is visibly more volatile – on a day-to-day basis – than other short-term sterling interest rates. Figure 7.2 illustrates clearly how day-to-day variability declines at somewhat longer maturities, in this case seven days and one month (LIBOR). Note, however, that the dependent variable throughout this chapter will not be the *level* of the effective overnight rate but the *spread* between the effective overnight rate and the Bank of England's stop rate:

$$\text{SPREAD}_t \equiv (\text{effective overnight rate})_t - (\text{stop rate})_t \, .$$

This variable, which will henceforth be referred to as "the spread" will obvi-

[11] I am grateful to Richard Pattinson for this data and also for additional data on the high and the low of the overnight rate which is employed below.

ously be less than zero whenever the effective overnight rate was below the stop rate (converted to a yield), and vice versa.

It is the spread which is of interest since, from the discussion in chapter 3, the primary determinant of the overnight rate is the official interest rate itself. Indeed, in the highly stylised environment of the model presented in that chapter the overnight rate was seen to be equal to the official discount rate in the event of a money market shortage. Figure 7.3, which plots the spread, confirms that the overnight rate does not deviate systematically from the stop rate, but fluctuates considerably around the stop rate. This variation, which is indicative of institutional and other forces on the overnight rate, is the focus of the remainder of this chapter.

7.2 Determinants of the spread

Table 7.3 shows that the spread averaged only 5 basis points over the period 1991 - 1994, although the mean spread in any particular year was usually somewhat higher or lower than this. Similarly, the (sample) variance and, more significantly, the volatility[12] of the spread varied substantially in different years, being much higher in 1992[13] and 1993.

These results were confirmed using information about the daily high and low of the overnight rate over the same period (table 7.4). It is evident that the average difference between the high (low) and the stop rate is fairly wide, approaching 200 basis points. The right hand column, which reports the Parkinson estimate of volatility – equation (7.2) –, is broadly consistent with the measures of volatility contained in the previous table and indicates that the volatility of the spread increased considerably in 1992 and 1993.

A further breakdown of the behaviour of the spread on a day-of-the-week

[12]Volatility is calculated here as the (sample) standard deviation of day-to-day changes in the spread.

[13]The dramatic events surrounding the withdrawal of sterling from the ERM in September 1992, which saw short-dated money rates rise to virtually unprecedented levels, accounts for the particularly high variance observed in 1992.

basis (table 7.5) revealed two patterns. First, the average spread between the overnight rate and the stop rate is not constant over the week and exhibits *positive autocorrelation.* The pattern in the average spread is typically that the spread tends to be lower at the start of the week, higher during the mid-week, and lower again at the end of the week. Second, the variance of the spread is usually somewhat higher at the *end of the week*[14]. These two patterns are again confirmed by the behaviour of difference between the high (low) of the overnight rate and the stop rate (table 7.6). The mean difference between the high (low) and the stop rate also show positive autocorrelation, while the Parkinson estimate indicates a clear increase in volatility at the end of the week. These patterns in the mean and the volatility of the spread are presented diagramatically in figure 7.4.

This behaviour of the spread is, undoubtedly, the result of a complex *interaction* of factors, whose individual influences may be hard to distil empirically. Nevertheless, it is possible to identify three classes to which these factors are likely to belong, and also how these factors might be expected to affect either the level or the volatility of the spread.

7.2.1 Expectations

The impact of an expected change in the official stop rate on very short-term interest rates (including overnight rates) was derived in chapter 5, which presented a model of the term structure of sterling interest rates. Evidence was found to support the model, which predicts that very short-term rates will move in the opposite direction to an anticipated stop rate change and, further, that short-term rates may become more volatile ahead of such an anticipated official rate change. A major problem, however, was the *identification* of interest rate expectations, which was overcome in chapter 5 by splitting the data into three distinct periods, each of which was assumed to characterise a particular belief about future stop rates. This technique is not

[14]The tendency for Thursdays to have a higher variance than Fridays may reflect the fact that Fridays are, on occasion, bank holidays. The importance of this is taken up in greater detail in the subsequent section.

particularly useful in the current context, where the intention is to formulate a specific regression model with which to explain day-to-day movements in the overnight rate. Given the difficulty of obtaining a meaningful proxy for market expectations of a stop rate change, the omission of this factor was considered to be the best strategy, with the result that interest rate expectations are not explicitly incorporated into the subsequent analysis. Whilst this omission is unfortunate, the fact that overnight rates in 1993 were extremely volatile *despite* only a single change in the stop rate suggests that this may not be too serious.

7.2.2 Structural factors

It was noted earlier that studies of overnight markets elsewhere (e.g. Feinman 1993a) sometimes found that several institutional features – for example the magnitude, the type and the timing of central bank operations – were systematically related to the overnight rate. These factors may thus be relevant determinants of the sterling overnight rate, for structural reasons and/or because they convey information about the policy stance of the Bank of England (i.e. they act as a policy signal).

Large money market shortages, for example, will mean that large quantities of eligible bills are sold to the Bank. This may sometimes reduce market participants' bill portfolios to the point where their ability to offer suitable bills to the bank is compromised[15]. These considerations may lead, in turn, to difficulties in "removing" the shortage, causing conditions in the overnight funds market to remain tighter than normal and driving a positive spread between the overnight rate and the stop rate.

By inviting its counterparties to offer eligible bills for repurchase (repo) instead of for outright sale, however, the Bank is in a position to alleviate these difficulties to some extent. By offering a repo, the Bank allows its counterparties to sell longer dated bills, which may be in greater supply than

[15]This might be due to the supply constraints on eligible bills identified in the previous chapter, which reduce the supply flow of fresh bills, and also because banks will wish to hold a minimum level of eligible bills for prudential liquidity purposes.

short-dated bills[16]. Of course, the Bank may, for whatever reason, choose *not* to offer a repo. This may then be construed by market participants as a deliberate signal by the Bank that it is considering a tightening of its monetary policy. The Bank, on the other hand, may have declined to offer a repo for some other, perhaps purely technical, reason.

It is not possible, in general, to identify particular actions by the Bank as clear signals of its policy intentions. Indeed, there is a good reason why the Bank may always prefer to give "noisy" policy signals as opposed to signals which make its intentions totally transparent: transparent signals reduce the Bank's ability to react freely to the arrival of new information. Such new information may cause the Bank to change its policy stance, thus making it vulnerable to the criticism that its original signals were misleading, possibly bringing its competence into question. Consequently, there are no actions by the Bank which are unanimously accepted by all market participants as a clear signal about future interest rates[17].

In the light of this, the **type** of operation (i.e. outright sale versus repo) offered by the Bank of England is not explicitly incorporated into the subsequent analysis, even though there may be examples elsewhere (e.g. the US) where the type of operation is considered as a relevant signalling variable.

The **size** of the daily money market shortage, on the other hand, is a good candidate for a structural variable which may affect the overnight rate, since it is not a choice variable of the Bank and is announced each morning by the Bank (making it part of the information set of market participants).

---

[16]Recall that the repurchase date, and not the maturity of the underlying instrument, is the relevant determinant of the maturity of a repo.

[17]This does not, of course, prevent market participants from interpreting the Bank's actions as signals, usually at their own peril. A recent market comment by a prominent investment bank claimed: "... the Bank of England moved quickly last week to dispel any doubts about what had been decided when the Chancellor met with the Governor of the Bank of England for their regular monthly meeting ... By setting a repo which expires partly on the day of their next scheduled meeting ... the Bank sent the clearest signal possible that the authorities had decided to keep interest rates on hold ..." (*Goldman Sachs UK Weekly Comment*, Issue Number 162, 9 Sept 1994). *The Bank raised its official lending rate by 0.5% on the next trading day.*

As regards the impact of the size of the shortage, larger shortages may be expected to result (on average) in a somewhat higher overnight rate (and hence spread). More indirectly, and for reasons explained in the previous chapter, larger shortages may also result in greater volatility in the spread, which is addressed further below.

Related to the size of the shortage, is the **timing** of the removal of the shortage during the course of the day, as measured by the extent of the announced shortage which remains outstanding after each round of the Bank's three daily open-market operations. Since the Bank chooses the proportion of offers for assistance which it accepts – by "scaling" offers – this variable is partly structural and may, on occasion, contain a Bank signal. To a large extent, however, the amount of the shortage remaining will reflect yet a third factor, namely strategic behaviour on the part of certain money market participants.

7.2.3 Strategic behaviour

Profitable trading in the overnight market is largely a matter of *timing*. It was seen above that the overnight rate exhibits considerable intraday movement, tending either to rise or to fall as the day evolves. Banks will therefore form expectations about the intraday behaviour of the overnight rate and will want to time their trading accordingly. If a bank needs to borrow funds but expects the overnight rate to fall, for instance, it will tend to delay its bid for funds in the hope that cheaper funds can be obtained later in the day.

One potentially important piece of information regarding the likely behaviour of the overnight rate is the announced shortage, adjusted by the amount of reserves already supplied by the Bank in its money market operations. Albeit with considerable error, the remaining shortage represents the aggregate reserve need of the banking system which is still outstanding. If the remaining shortage is expected to be substantial as the day progresses, therefore, the aggregate demand for overnight funds will be expected to stay high and so the overnight rate will be expected to rise. Of course, this expectation will provide banks with a greater incentive to offer bills to the Bank of

England, which would tend to alleviate the shortage more quickly, countering the expected increase in the overnight rate. If the sterling money market were populated exclusively by identical, atomistic banks, each of which could participate equally in the daily operations of the Bank of England, a predictable relationship between the overnight rate and the proportion of the shortage which remains outstanding at any time during the day would be unlikely[18].

If the above characterisation of the sterling money market were correct, the proportion of the shortage outstanding would not contain information about the overnight rate. Previous chapters, however, have highlighted that the sterling money market does not correspond closely to this characterisation. The clearing banks, in particular, are considerably different from other banks, not least because they are substantially larger than most other banks. Two additional features also set the clearing banks apart - they are usually large net *borrowers* of overnight funds and they have recently – since the late 1980s – become very large holders of eligible bills. As the major holders of eligible bills, the clearing banks have become the dominant counterparties in the daily operations of the Bank of England: a substantial proportion of the bills which are sold to the Bank via the discount houses each day are thus offered by the clearing banks.

This institutional feature is important, for it may permit a clearing bank, through strategic behaviour, to gain precise information about its own reserve needs relative to those of other banks, enabling it to borrow cheaper overnight funds. A clearing bank may "position" itself so that its own borrowing need exceeds the forecast shortage by a comfortable margin[19]. By selling sufficient bills to the Bank to remove the entire shortage during the first round of operations, the bank is left in the position where it still needs to borrow a substantial amount of funds, whilst at the same time ensuring that the

[18]Imprecise information about the reserve position of other banks, and about the intentions of other banks as regards the sale of bills to the Bank, would make it extremely unlikely that any single bank could, through its own actions, influence the overnight rate.

[19]Given that the average size of the shortage is about £1000 million, this is clearly only possible for very large banks who, simply by virtue of their size, naturally have a substantial *absolute* short-term borrowing need.

aggregate demand for reserves (i.e. the remaining shortage) is zero. This then implies that the clearing bank's reserve deficit must be matched by a reserve surplus in some other banks, who do *not* possess this information and who *atomistically* compete to lend their surplus funds. By lowering its bid rate for overnight funds, the clearing bank can then meet its reserve deficit at a lower cost[20].

The possibility of strategic behaviour has a further consequence. Regular participants in the overnight funds market may then try to position themselves to take advantage of anticipated overnight rate movements. If, for instance, it were suspected that the overnight rate would be low on some particular day, it would become rational for all banks to try to borrow (temporarily) cheap overnight funds. This would tend to frustrate the strategy outlined above, whose success depends on the fact that it is not anticipated. For this reason, a clearing bank may deliberately choose *not* to offer bills to the Bank. Since it is not automatically the case that all the clearing banks (and other holders of eligible bills) always offer bills to the Bank, this withholding of bill offers may result in the shortage not being removed until fairly late in the day. Insofar as this succeeds in causing the overnight rate to be higher, it may also lead market participants to expect (incorrectly) that similar conditions will prevail on the subsequent day (or days).

In sum, strategic behaviour on the part of the large clearing banks would cause the overnight rate to vary positively with the proportion of the shortage outstanding (after the Bank's operations). Furthermore, it may also account for some of the weekday effects that were noted in the cursory examination of the spread in the previous section. Since the maturity of an overnight loan is longer before weekends and public holidays, the incentive to engage

[20]Donaldson (1993) has recently presented a model of interest rate determination in which larger banks are able to "corner" the market in short-term funds (cash), thereby becoming what he calls "strategic suppliers" of cash to other market participants - at rates substantially above normal market rates. Although it is employed by Donaldson to explain interest rate behaviour during banking crises, the model offers useful parallels to the manner in which strategic behaviour affects interest rate determination in the sterling overnight market.

in such behaviour increases at these times. The movement of the overnight rate before weekends and public holidays may thus also indicate strategic behaviour.

To some extent, however, these effects will also be the result of *self-fulfilling* beliefs rather than strategic behaviour: once it is believed that a large remaining shortage (say) will result in a higher overnight rate, those banks wanting to lend overnight funds will delay offers whilst those banks wanting to borrow overnight funds will immediately submit bids, bringing about the anticipated movement (and reinforcing the original belief). Even when the size (proportion) of the remaining shortage is *not* due to strategic behaviour, therefore, the overnight rate may react as if this were the case.

Could the relationship between the remaining shortage and the overnight rate be *entirely* attributable to self-fulfilling beliefs rather than to strategic behaviour? Examining the variance (volatility) of the overnight rate, and its determinants, may provide a useful method of distinguishing between these (competing) explanations. An integral part of the strategic behaviour described above is to make the timing of bill sales to the Bank a source of *misinformation* about the day-to-day movement of the overnight rate. Offers of bills to the Bank are thus withheld in order to generate uncertainty about the possible path of the overnight rate on subsequent days. This suggests that the variance of the overnight rate would exhibit time-varying behaviour, since successive days of a somewhat higher overnight rate (i.e. low variance) will be followed by a day (or more) on which rates are suddenly lower (i.e. high variance). If, on the other hand, the remaining shortage were simply a random variable (i.e. unaffected by strategic behaviour), always generating the same self-fulfilling response in the overnight rate, this time variation in the overnight rate should not be apparent.

Admittedly, the presence of time variation in the variance of the overnight rate may be due to any number of other, unidentified, factors. A much stronger case for the presence of strategic behaviour would exist if the variance of the overnight rate was significantly related to the liquidity premium on eligible bills (i.e. the spread between the yield on bills and equivalent instruments). As the previous chapter has argued at some length, the size of

this liquidity premium may penalise the holders of eligible bills and provide a strong incentive for them – i.e. the clearing banks – to generate cheaper funding in the overnight market.

The following two sections now trace the impact of the above factors on the level and volatility of the overnight rate. Given that reliable daily data on these factors could only be obtained for the sub-period from the beginning of January 1992 until the end of October 1993, this served as the period for analysis. This sub-period still represented a considerable number of data points, namely 464 daily observations, and also incorporates most of the period over which the overnight rate was at its most volatile.

7.3 Estimating a model of the spread

Recall that the dependent variable in the analysis is the daily spread between the overnight rate and the stop rate (SPREAD). The mean spread over the sub-period was about -5 basis points, with a minimum and maximum of -2.42 and 8.08 percentage points respectively (see table 7.8). As there is a strong presumption that the overnight rate does not deviate from the stop rate in any "long-run" sense – although it may exhibit short-run deviations – the day-to-day spread should be a stationary process. This was confirmed by a test for a unit root, which was conclusively rejected.

7.3.1 Autocorrelation

It was noted above that the average spread between the overnight rate and the stop rate exhibits positive autocorrelation. A plot of the autocorrelation and partial autocorrelation functions confirmed this, and suggested that the dynamics of the spread conformed to an autoregressive process of order one, i.e. AR(1). As a benchmark, therefore, the following model was estimated:

$$\text{SPREAD}_t = \theta_0 + \theta_1 \cdot \text{SPREAD}_{t-1} + \varepsilon_t \,. \tag{7.3}$$

The results of fitting such a model are shown in table I below.

Table I. Conditional mean of the spread		
Variable	Coefficient	t-Statistic
CONSTANT	-0.025	0.86
$SPREAD_{t-1}$	0.42	12.72

$R^2 = 0.173$.

The AR(1) representation of the spread is clearly significant and can account for about 17% of the variation in the spread. Note also that the (long-run) expected value of the spread, given by $E(SPREAD_t) = \frac{\theta_0}{1-\theta_1} = -0.04$ (i.e. 4 basis points) is close to the actual mean observed over the sample.

7.3.2 Other explanatory variables

The previous section identified three variables which may be expected to influence the spread between the overnight rate and the stop rate on a systematic basis, namely the overall size of the shortage, the amount of the shortage which remains outstanding during the course of the day, and weekends and public holidays. These variables were calculated as follows:

Size of the money market shortage

Although the shortage is announced at 9h45 each trading day, this is only a forecast figure and is frequently revised over the course of the day. These revisions are usually small in relation to the overall size of the shortage, however, and are more often upwards rather downwards[21]. It was thus decided to use the final revised figure for the shortage, published by the Bank at 14h00 each day, which is denoted by SIZE.

The average shortage over the sample period was £1168 million, with the minimum and the maximum shortage being −£200 (i.e. a surplus) and £3950 million respectively (table 7.6). It is noteworthy that the average size

[21]This is probably because the Bank prefers to understate rather than overstate the shortage. The latter could result in reserves being oversupplied, with the aggregate surplus of reserves resulting in a sharply lower overnight rate. This may require the Bank to supply and remove reserves on the same day - something which it has never dome.

of the shortage exhibits a weekly pattern, tending to be smallest on Tuesdays and highest on Fridays (table 7.7). This probably reflects the behaviour of currency holders, who tend to increase their currency holdings just before weekends. After the weekend, retailers then redeposit this currency with the banking system. Table 7.7. indicates that this tendency is even more marked before long weekends (i.e. weekends where either the adjoining Friday and/or the Monday is a bank holiday), with shortages on such occasions being larger than on normal Fridays.

Proportion of the shortage remaining

Having announced its forecast shortage, the Bank then invites its counterparties to offer bills for sale to the Bank. The total amount of bills purchased by the Bank is then immediately published, indicating the amount of the forecast shortage which remains. The proportions of the shortage still remaining after each round of open-market operations thus serve as useful proxies for the timing of the removal of the shortage, and were calculated as follows:

$$\text{REMAINDER}_s = 1 - \left[\frac{\sum_{t=9h45}^{s} \text{Operations}_t}{\text{Shortage}_s} \right],$$

where s denotes either 9h45, 12h00, 14h00 or 14h45, which are the intervention times of the Bank.

Figure 7.5 shows four scatterplots, each depicting the remaining shortage (REMAINDER) against the size of the shortage (SIZE) after each round of the Banks' operations. Two important tendencies are apparent from the figure. First, the proportion of the shortage outstanding after the first (9h45) round of operations is typically quite high, indicated by the clustering at the top of the plot. To some degree this reflects the fact that the Bank of England does not usually invite its counterparties to offer bills at 9h45 if the size of the forecast shortage is below about £750 million. Using this variable to proxy the rate of removal of the shortage, therefore, would produce misleading results, since a high outstanding shortage does not necessarily reflect a *choice* by the Bank's counterparties not to offer bills. A second feature of figure 7.5 is that smaller shortages tend to be removed more slowly than large shortages, which may give the impression that smaller shortages tend to produce tighter

money market conditions than large shortages. This would be misleading, however, since it is the actual amount of the shortage which is outstanding, rather than the proportional amount, which is relevant to the degree of reserve availability.

In the light of these problems, REMAINDER as measured at 12h00 was deemed to be the best proxy variable for the rate of removal of the shortage. Calculation of this variable over the sample indicated that, on average, 50% of the shortage remained outstanding at noon. And, as with the size of the shortage, this variable also varies over the days of the week (table 7.6.). The pattern in this case is reversed, however, with the average proportion of the shortage left outstanding at noon being lower at the beginning and at the end of the week[22].

Weekends and public holidays

On Fridays and before public holidays, an overnight loan effectively becomes a longer term loan, proportionally increasing the costs or profits accruing to overnight trades and thus also the tendency for strategic behaviour. To examine the impact of this on the overnight rate, the following dummy variable was constructed:

$$\text{WEEKEND}_t = \begin{cases} 0 & \text{if t+1 is a working day} \\ 1 & \text{otherwise} \end{cases} , \; t = 1, \ldots, T.$$

Essentially, therefore, the WEEKEND dummy variable simply takes on a value of unity on Fridays or before bank holidays (usually a Thursday in the UK) and a value of zero on all other days.

These three variables, SIZE, REMAINDER and WEEKEND were thus included in the basic regression model (7.3) above, which was then re-estimated over the sample period. The results are provided in table II below.

[22]Interestingly, the proportion of the shortage remaining at 12h00 on days before long weekends was even lower than on Fridays, a fact which we return to below in the discussion of the estimation results.

Table II. Conditional mean of the spread

	Variable	Coefficient	t-Statistic
Dynamics	CONSTANT	-1.18	-11.91
	$SPREAD_{t-1}$	0.36	9.73
Other effects	SIZE	0.00044	7.09
	REMAINDER	1.40	14.56
	WEEKEND	-0.33	3.93

$R^2 = 0.43$, $F_{4,454} = 87.47$

It is evident from table II that the addition of these three variables dramatically improves the explanatory power of the basic model, with the R^2 rising from 0.17 to 0.43. The t-statistics indicate that each of the explanatory variables is significant at the 1% level, whilst the F-statistic shows that the explanatory variables are jointly significant at the 1% level.

The coefficients are straightforward to interpret and their sign is consistent with the expected impact of the factors identified in the previous section. Both the size of the shortage, and the proportion of the shortage outstanding at noon, drive the overnight rate above the stop rate, whilst weekends and public holidays drive the overnight rate below the stop rate. More specifically, an increase of £100 million in the money market shortage tends to widen the spread between the overnight rate and the stop rate by about 5 basis points, whilst an increase of 10% in the proportion of the shortage remaining tends to raise the spread by about 15 basis points. The spread between the overnight rate and the stop rate also tends to be about 30 basis points wider on Fridays and before public holidays, albeit in the opposite direction to the other factors; that is, the overnight rate tends to fall further below the stop rate on these days.

What is relevant to determining the spread on any particular day, however, is the *combined impact* of these factors. This is also consistent with the expected impact of these factors as discussed above. From table 7.7 it is evident that an "average" day in the money market over the period had

an announced shortage of about £1100 million, 50% of which remained at 12h00. These values imply that on an average day the expected spread between the overnight rate and the stop rate is zero (given that the spread on the previous day was zero). An above average shortage of (say) £2000 million would, *ceteris paribus*, raise the expected spread to +40 basis points. If, additionally, an above average proportion of the shortage remains outstanding at 12h00 – for example 80% – this further increases the expected spread to +80 basis points. Alternatively, if the entire shortage is removed by 12h00 (i.e. REMAINDER = 0), and the following day is a weekend (i.e. WEEKEND = 1), the expected spread is −60 basis points.

Broadly, therefore, the size of the shortage may be interpreted as a structural factor which, *ceteris paribus*, tends to drive a positive spread between the expected overnight rate and the stop rate. This factor, however, does not operate in isolation. The proportion of the shortage remaining at 12h00 is crucial in determining the expected spread: the spread will be positive when the bulk of shortage is taken out late in the day and negative when the bulk of the shortage is taken out early in the day. Thus the relationship, if any, between the size of the shortage and the proportion remaining at 12h00 will be of some importance. An OLS regression of REMAINDER on SIZE revealed a significant negative correlation but an extremely low R^2: that is, *virtually none of the variation in the proportion of the shortage outstanding can be accounted for by the size of the shortage.* Larger shortages, therefore, do not translate straightforwardly into a higher or a lower overnight rate: what matters is the proportion of the shortage that is removed during the various operations of the Bank, which is *not predictable* from the size of the shortage. Instead, this will depend on the decisions of those institutions which hold the bulk of the eligible bills, namely the clearing banks.

A rather clear indication of strategic behaviour is given by the fact that the overnight rate on Fridays (and before public holidays) tends to be systematically lower than on other days. Since the overnight rate is *predictably lower* on Fridays, however, this behaviour would have limited consequences for market efficiency, since market participants could adjust their reserve management behaviour accordingly. A potentially stronger case for market

inefficiency could be made if factors such as the size of the shortage, the proportion of the shortage outstanding and, importantly, the liquidity premium on eligible bills, were found to increase the *volatility* (i.e. the conditional variance) of the overnight rate. This would suggest that market participants were facing a higher level of uncertainty due to the influence of these structural and strategic variables, and would carry the implication that these variables actually inhibited the efficiency of the overnight market in facilitating reserve management.

A natural starting point for the examination of this issue, then, is to test whether the variance of the spread was in fact constant over the sample period. If not, then further analysis of the variance of the spread, and its determinants, is justified.

7.3.3 Heteroscedasticity

Whenever the variance of the error, or residual, of a regression are not constant across observations, the regression is said to be heteroscedastic. Consequently, tests for heteroscedasticity typically take the form of Lagrange multiplier tests which utilise the residuals of the regression (see Godfrey 1988). The hypothesis being tested takes the following form:

$$H_0 : \sigma_t^2 = \sigma^2$$
$$H_A : \sigma_t^2 = \sigma^2 \cdot (\alpha_0 + \underline{\alpha}' \cdot z_t)$$

Essentially, the absence of heteroscedasticity corresponds to the result that the α's above are equal to zero. Testing this hypothesis is achieved by constructing an "artificial" regression in which the squared (normalised) residuals of the original regression are the dependent variable. The test statistic is then $T \cdot R^2$, where T is the number of observations and R^2 is the uncentered R-squared obtained in the regression, which is asymptotically distributed as chi-squared with q (the number of explanatory variables) degrees of freedom. Unfortunately this test is somewhat general and the particular

type of heteroscedasticity being tested for depends only upon the explanatory variables z_t which are included in this artificial regression.

A common test for heteroscedasticity in financial data is the ARCH test (Engle 1982), in which the explanatory variables z_t are simply the lagged squared errors, up to lag q. As will be shown in the following section, ARCH effects imply that the conditional variance (of the error) exhibits a specific form *time variation* in which periods of higher variance are followed by periods of lower variance. An ARCH test was therefore conducted using four lags[23]. The null hypothesis could be rejected at the 5% level[24], indicating the possibility of ARCH effects in the data.

Another common test for heteroscedasticity is that of Breusch and Pagan (1979), who include in z_t some or all of the explanatory variables of the original regression, plus any additional variables which are hypothesised to influence the variance. Such a test was therefore undertaken using three explanatory variables, namely SIZE, REMAINING and, lastly, the liquidity premium on eligible bills. The latter variable, which is simply the difference between the yield on eligible bills and LIBOR (this was denoted by S in the previous chapter, but will be denoted by PREMIUM in this chapter), was included under the presumption that strategic behaviour in the overnight funds market might increase with this premium, in turn affecting the variance of overnight rate. The null hypothesis could be easily rejected at the 1% level[25], indicating the presence of a relationship between the variance of the overnight rate and these explanatory variables.

7.4 The variance of the spread

The heteroscedasticity tests suggest that the variance of the regression errors of the model estimated in the previous section not only exhibits ARCH effects, but may also be systematically related to several other explanatory variables

[23]This choice was arbitrary.

[24]The test statistic was computed as 12.4.

[25]The test statistic was computed as 106.

in the model. This section thus briefly introduces the notion of ARCH, which may take a number of different forms. For this reason, several ARCH specifications are estimated, and a parsimonious specification chosen by using some simple tests recently developed by Engle and Ng (1993). It is then a straightforward matter to estimate a more general model of the conditional variance, incorporating those explanatory variables which might plausibly affect this variance.

7.4.1 Autoregressive Conditional Heteroscedasticity

The notion of ARCH is due to Engle (1982), and can be identified using the basic model (7.1) introduced at the outset of this chapter,

$$y_t = \underline{\theta}' \cdot \underline{x}_t + \varepsilon_t \,, \quad t = 1, \ldots, T \,.$$

Summarising the information available at time t by ψ_{t-1}, ARCH(q) may be expressed as

$$\varepsilon_t \mid \psi_{t-1} \sim N(0, h_t) \tag{7.4}$$

where

$$h_t = \omega + \beta_1 \cdot \varepsilon_{t-1}^2 + \ldots + \beta_q \cdot \varepsilon_{t-q}^2 \tag{7.5}$$

with $\omega > 0$ and $\beta_i \geq 0$, $i = 1, \ldots, q$ by assumption to ensure that the conditional variance h_t is positive. In words, ARCH(q) implies that the conditional variance of the *current* error is a linear, symmetric function of *past* errors up to lag q[26].

The distinguishing feature of (7.5) is that the conditional variance (i.e. volatility) varies in an *episodic* manner, with large (small) errors tending to cluster sequentially. The order of the lag, q, simply captures the length of time for which any given error will *persist* in conditioning the variance of subsequent errors. Essentially, therefore, and as its name conveys, ARCH is simply the specification of temporal dependence in the second moments (i.e

[26]Since the most recent lagged error is a function of the current information set, i.e. $\varepsilon_{t-1} = (y_{t-1} - \underline{\theta}' \cdot \underline{x}_{t-1})$, this means that the conditional variance also depends on the current information set.

in the variance) of a time series, and is fully analogous to the specification of some form of ARMA process for the conditional mean of a time series (as given, for example, by the model of equation 7.3)[27].

An important innovation to the basic ARCH(q) formulation is due to Bollerslev (1986), who noted that in many empirical applications this model assumed an *ad hoc* linear declining lag structure in the conditional variance equation to take account of the long persistence which was typically found. This assumption was invoked since free estimation of the lag structure would often violate the non-negativity constraints on the coefficients. Bollerslev thus proposed the introduction of the lagged conditional variance, h_{t-i}, $i = 1,\ldots,p$, into the variance equation (7.1). This extension, which is called *generalised* ARCH (GARCH(p,q)), has been found to be applicable in a large number of empirical studies (see Bollerslev, Chou and Kroner 1992) with only a single lag, i.e.

$$h_t = \omega + \alpha \cdot h_{t-1} + \beta \cdot \varepsilon_{t-1}^2 \,. \tag{7.6}$$

Further innovations to the GARCH formulation have often involved the explicit incorporation of some kind of non-linearity or asymmetry. Under the GARCH specification, past errors can only affect the current variance in a linear, symmetric fashion. In reality, however, this dependence between the errors and the variance may be non-linear, or asymmetric, and the GARCH specification should allow for this.

A simple example of an *asymmetric* GARCH (AGARCH) model is given by

$$h_t = \omega + \alpha \cdot h_{t-1} + \beta \cdot (\varepsilon_{t-1} + \gamma)^2 \tag{7.7}$$

---

[27]The estimation of ARCH models is straightforward via maximum likelihood methods. If the normality assumption of conventional OLS regression is maintained, the loglikelihood of the t^{th} observation is given by

$$logL_t = -\frac{1}{2}\left(log(h_t) + \frac{\varepsilon_t^2}{h_t}\right) .$$

The parameters $\underline{\theta}$, ω and β_i, $i = 1,\ldots,q$, can then be estimated recursively by applying any one of the several algorithms commonly applied to finding the maximum of the joint likelihood for some given sample (see Hamilton 1994 or Greene 1993).

Clearly if $\gamma < 0$ then negative errors have a greater effect on h_t than positive errors (the opposite is true if $\gamma > 0$). Another, more general, asymmetric model was introduced by Nelson (1990), namely an *exponential* GARCH (EGARCH) model:

$$log(h_t) = \omega + \alpha \cdot log(h_{t-1}) + \beta \cdot \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma \cdot \left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right) . \quad (7.8)$$

Here it is again the case that, if $\gamma < 0$, negative errors have a greater effect on the conditional variance than positive errors (vice versa if $\gamma > 0$).

Although both the AGARCH and the EGARCH specifications permit the past errors to influence the conditional variance in an asymmetric fashion, the EGARCH model also has two further advantages over other GARCH models. First, the exponential form implies that the conditional variance cannot become negative, with the result that *no non-negativity constraints* have to be placed on the coefficients: in principle, therefore, this permits a more flexible specification. A second advantage of the exponential form is that it allows *large* past errors to have a greater effect on the conditional variance than does the standard GARCH model, which may more accurately capture the type of ARCH which is present in the data.

Clearly, then, there are many potential models which may be used to specify the conditional variance, as is amply demonstrated by the literature reviews of Bollerslev, Chou and Kroner (1992) and Bera and Higgins (1993). How should a researcher discriminate, or choose, between these various alternatives?

7.4.2 Which GARCH?

Engle and Ng (1993) have recently suggested that alternative ARCH models may be simply viewed as different formulations of how past errors[28] affect

[28]In studies where the dependent variable is asset returns, the lagged error may be interpreted as a collective measure of "news" at time t, since it is the *unexplained* component in current returns. This interpretation does *not* carry over to the model in this chapter, where the dependent variable is the spread between the overnight rate and the stop rate.

current volatility. This is most easily seen by holding constant all information dated $t-2$ and earlier, which then leaves the current volatility h_t as a function of the lagged error ε_{t-1} only. The conditional variances of each type of GARCH may then be written as

$$h_t = \begin{cases} A + \beta \cdot \varepsilon_{t-1}^2 & \text{GARCH} \\ A + \beta \cdot (\varepsilon_{t-1} + \gamma)^2 & \text{AGARCH} \\ B \cdot exp\left[\frac{\beta+\gamma}{\sigma} \cdot \varepsilon_{t-1}\right] & \text{EGARCH } (\gamma > 0)\,, \end{cases} \tag{7.9}$$

where $A = \omega + \alpha \cdot \sigma^2$, $B = \sigma^{2\alpha} \cdot exp[\omega - \beta \cdot \sqrt{2/\pi}]$ and σ^2 is the *un*conditional variance.

The simple GARCH model is thus just a quadratic function of the past error, centred at $\varepsilon_{t-1} = 0$. The AGARCH model is also a quadratic function, but is centred at $\varepsilon_{t-1} = \gamma$. Finally, the EGARCH model is similar to a quadratic function centred at $\varepsilon_{t-1} = 0$, but increases exponentially in both directions (possibly with different slopes if $\gamma < 0$).

The three functions in (7.9) may be interpreted as "error impact" curves, since they trace the effect of the past error on the conditional volatility. In other words, they show the manner in which the different GARCH models capture ARCH effects. Consequently, they also highlight the potential shortcomings of these alternative models, and how they may *misspecify* the true ARCH which is present in the data. A simple GARCH model, for instance, does not distinguish between positive and negative shocks, and assumes that these shocks affect the variance through a quadratic relationship. Clearly, therefore, the model may have a *sign* bias, in that the actual ARCH effects in the data may be affecting the variance in an asymmetric fashion. Similarly, the model may have a *size* bias, overestimating the impact of small errors on the conditional variance whilst underestimating the impact of large shocks on the conditional variance. AGARCH and EGARCH models, whilst more flexible than the simple GARCH model, may contain similar biases.

For this reason, Engle and Ng (1993) develop three simple diagnostic tests which attempt to capture the bias of an ARCH specification with respect to a particular dataset. These tests, which are all very similar to Lagrange

multiplier tests, examine whether variables which *are* in the current information set but which are *not* included in the volatility model can predict the squared (normalised) residuals. The tests thus provide a useful way to discriminate between the potentially numerous GARCH specifications which might characterise the conditional variance of the empirical model of the previous section.

In the interests of continuity these tests are described in more detail in an appendix to this chapter. Three specifications for the conditional variance of the spread, namely the GARCH, AGARCH and EGARCH as specified in (7.6), (7.7) and (7.8) respectively, were then estimated in conjunction with the original model. The results, which are also reported in the appendix, reveal that the coefficients in the conditional variance equations of each of the three GARCH models are all significant at the 1% level. Potentially, therefore, any one of these would be found to be an acceptable specification if it were conducted in isolation. The fact that the γ coefficient is positive in the AGARCH model, however, suggests that there is some asymmetry between the impact of positive and negative lagged errors on the conditional variance, with positive shocks having a larger impact on the variance than negative shocks. This asymmetry, which is corroborated by the EGARCH model (since both α and γ are positive), suggests that either the AGARCH or the EGARCH would be preferable to the simple GARCH model.

The diagnostic test results reported in the appendix also reveal that the alternative GARCH models are all very similar in the manner in which they capture the ARCH effects present in the data. The test statistics indicate that in all three cases the null of no sign bias and no negative size bias are both accepted, while the null of no positive size bias is rejected. Essentially, therefore, all three models tend to *under*predict the impact of large positive (lagged) errors on the conditional variance, and all three models consequently fail the joint test. Since the ARCH effects in the data suggest that large positive errors have a considerable impact on the conditional variance, the EGARCH model seems, overall, to be the most well specified. This conclusion is supported by the loglikelihood as well as the results of the joint test, both

of which are somewhat less in the case of the EGARCH model[29].

The presence of ARCH effects (i.e. time variation in the conditional variance of the errors of the empirical model estimated in section 3) can, it was argued, be accounted for by strategic behaviour on the part of the clearing banks. They may also be accounted for by other factors, which have not been identified. To strengthen the case for the presence of strategic factors in the data, therefore, it is necessary to examine whether other variables directly affect the conditional variance.

7.4.3 Other explanatory variables

There are several variables whose impact on the variance of the spread is of interest. First, the size of the shortage (SIZE) may affect the variance. If this was found to be the case, it could be interpreted as a form of structural inefficiency: by increasing market participants' uncertainty regarding the day-to-day evolution of the overnight rate relative to the stop rate, larger shortages tend to increase the resource costs (or risks) of reserve management in the sterling overnight market. Second, the proportion of the shortage remaining at 12h00 (REMAINDER), the incidence of weekends and public holidays (WEEKEND) and the liquidity premium on eligible bills (PREMIUM) may also affect the variance. This could be construed as a different form of structural inefficiency, arising from the incentive for, and ability of, certain market participants to engage in strategic behaviour.

The impact of these variables on the variance of the spread can be straightforwardly assessed by reestimating the EGARCH model and including these variables in the conditional variance equation. Indeed, for purposes of comparison the EGARCH model estimated in the preceding section may be considered as a restricted form of this, more general, specification. The results,

[29]The results suggest, however, that an even better model would probably be one for which the relationship between the positive lagged errors and the conditional variance was even steeper than an exponential. This may reflect the few very large positive errors experienced in September 1992.

which were obtained by maximum likelihood estimation, are shown in tables III and IV below.

Table III. Conditional mean of the spread

	Variable	Coefficient	t-Statistic
Dynamics	CONSTANT	-1.14	-16.1
AR(1)	$SPREAD_{t-1}$	0.33	9.88
Other effects	SIZE	0.0004	8.69
	REMAINDER	1.33	16.57
	WEEKEND	-0.27	-3.64

Table IV. Conditional variance of the spread

	Variable	Coefficient	t-Statistic
Dynamics	CONSTANT	-1.7	-8.05
EGARCH	$VARIANCE_{t-1}$	0.182	2.4
	$exp\left\{\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}\right\}$	0.082	1.17
	$exp\left\{\frac{\lvert\varepsilon_{t-1}\rvert}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}}\right\}$	0.3	2.91
Other effects	SIZE	0.00028	2.88
	REMAINDER	0.51	3.00
	PREMIUM	0.65	5.42

The results in table III, for the conditional mean of the spread, do not differ much from those reported in table II: indeed, the dynamics as well as the original explanatory variables become *more* significant once the conditional variance is modelled explicitly. Table IV, which is of interest here, indicates that the GARCH dynamics and three out of the four explanatory variables in the conditional variance are significant at the 1% level - WEEKEND was not significant. These three variables were also found to be jointly significant

at the 1% level[30].

7.5 Concluding remarks

This chapter has confirmed that there are several features of the sterling money market which not only systematically affect the level of the overnight rate, but also systematically increase its variance. That is to say there are elements of the *microstructure*, or architecture, of the money market which, although they can successfully predict almost 50% of the daily variation between the overnight rate and the stop rate, also tend to make this prediction more uncertain. The "explanation" proffered here is that the current architecture of the money market and the Bank of England's operations is such that some money market participants have both the *incentive* and the *ability* to generate greater uncertainty about the likely path of the overnight rate. And, whilst this may have limited consequences in terms of the transmission of monetary policy to the wider economy, it surely affects the efficiency of the overnight market as the locus for sterling reserve management[31]. How, then, might the (micro)structure of the money market and the Bank's daily operations be changed to address these problems?

The size of money market shortages has clearly played a major role in the determination of both the level and variance of the spread between the overnight rate and the stop rate, and has done so on a number of fronts – both direct and indirect. For instance it led, in the early 1980s, to a

[30]This was evaluated using a likelihood ratio test, which compares the change in the maximum loglikelihood attained in the unrestricted versus the the restricted model (the simple EGARCH model without the additional variables). The maximum loglikelihood achieved in the unrestricted estimation was 17.64, as against -1.74 in the restricted model (see the table of results in the appendix). The test statistic is computed as 2 (17.64+1.74) = 38.76, which easily exceeds the chi-squared variate with 3 degrees of freedom at the 1% level.

[31]As noted at the outset of this chapter, daily trading in overnight funds *vastly* exceeds daily turnover on the London Stock Exchange. Were this market, and the prices established there, to exhibit similar behaviour, this would undoubtedly prompt the interest of both academics and regulators, and would certainly lead to debate about possible reform.

change in the financial instruments that were routinely purchased by the Bank (i.e. from Treasury bills to bank bills), a change which was accompanied by the imposition of restrictions on the issue of these instruments (see chapter 6). These restrictions have, in the presence of continuously large shortages, lowered the yield on eligible bills, with two consequences. First, eligible bills became less attractive as assets, which meant that a typical bank would have little incentive to hold bills for purposes of obtaining reserves from the Bank of England. Second, by concentrating the ownership of bills within a few, large banks, it then created the opportunity for these institutions to deploy bills in a *strategic* fashion. This linkage – between the liquidity premium on eligible bills and the volatility of the overnight rate – has been clearly demonstrated, and it is difficult to conceive of a factor besides strategic behaviour which could have resulted in such a link being present.

It is important to think about this situation as an "equilibrium": the unattractiveness of bills as *liquid* assets must be compensated by their utility as *strategic* assets. Of course, an alternative equilibrium is conceivable, namely one in which the Bank purchases assets which have a high liquidity value and a low strategic value. Government securities fit this description relatively well[32]. It is encouraging, therefore, that the Bank has recently begun to move in the direction of relying to a greater extent upon repurchase agreements (repo) based upon gilts, although only twice a month (see chapter 4). There is no good reason why the Bank could not rely exclusively (i.e. daily) upon gilt repos in its money market operations, a move which would go a long way towards eliminating the strategic behaviour discussed in this chapter.

Other refinements of the Bank's operations logically follow from this, and have been discussed at some length in Schnadt (1994a, 1994b and 1994c). It is not obvious, for example, why *three* open-market operations are necessary each day or why these operations have to be channelled through a narrow group of passive counterparties, i.e. discount houses. A single open

[32]Their liquidity value derives from the fact that they are virtually free of credit-risk, which would continue to be true even if the Bank purchased them. Their strategic value, on the other hand, would be low, as they are in large supply and relatively widely held.

market operation conducted (say) at midday, in which a broad group of financial institutions is invited to repo gilt-edged securities with the Bank, seems sufficient. Thereafter, financial institutions which still wished to obtain same-day reserves from the Bank could utilise a standing facility, which would again require the repo of gilts with the Bank, but which would attract a higher interest rate than that applied to the open-market transaction[33]. In the absence of totally *unexpected* variations in the reserve needs of a particular institution – or in the money market shortage –, therefore, the standing facility would not be utilised. The fact that it were available, however, would place a ceiling on the overnight rate that was equal to the implied cost of one-day funds under this facility[34].

In sum, it is not difficult to envisage comparatively minor adjustments in the operating procedures of the Bank which would stabilise overnight interest rates but leave the Bank's flexibility to influence money market interest rates unchanged.

[33]This rate could increase as the cut-off time for same-day value was approached.

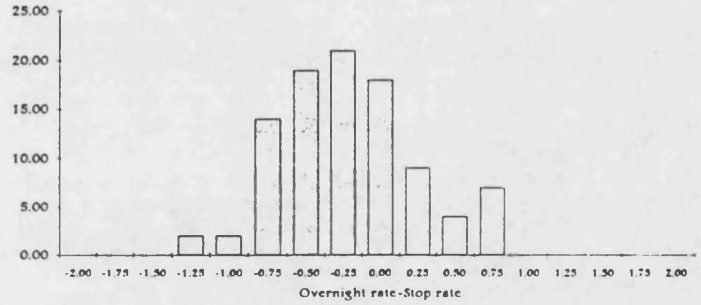
[34]The current regime of reserve accounting in the UK, under which clearing banks strive to meet a comparatively small reserve target on a daily basis may also be refined. Larger reserve requirements could be imposed, which had to be met over a longer period (thereby attaining greater overnight interest rate stability due to averaging) but which were remunerated at close to the market rate (thereby reducing their implicit cost on banks).

7.6 Figures and Tables

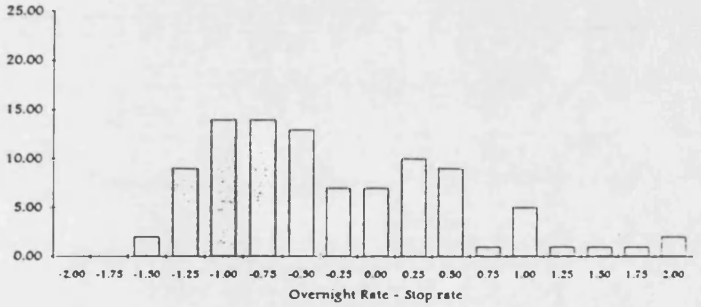
Figure 7.1

The intraday distribution of the overnight rate
(Four daily observations: Jan - May 1994)

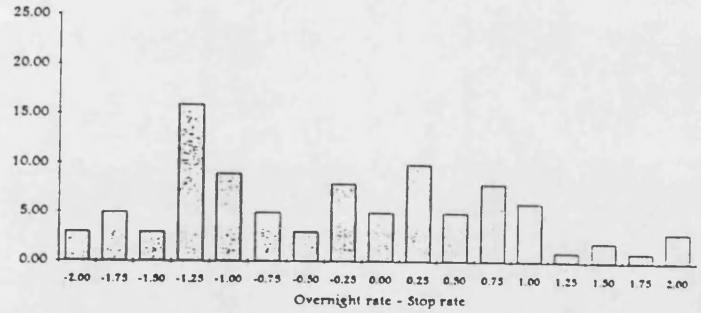
8:00 AM



11:00 AM



2:00 PM



CLOSE

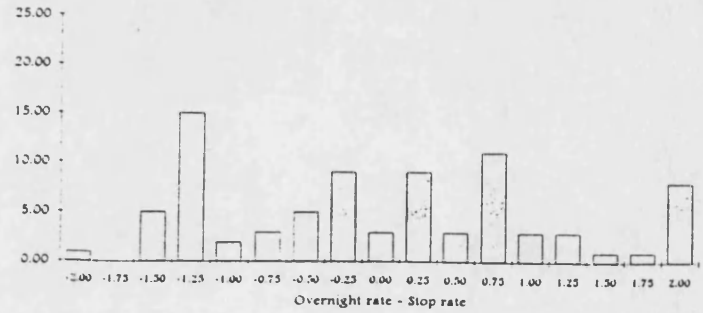


Figure 7.2

The effective overnight rate compared with seven day LIBOR and one month LIBOR (Daily observations: Jan 1992 - Oct 1994)

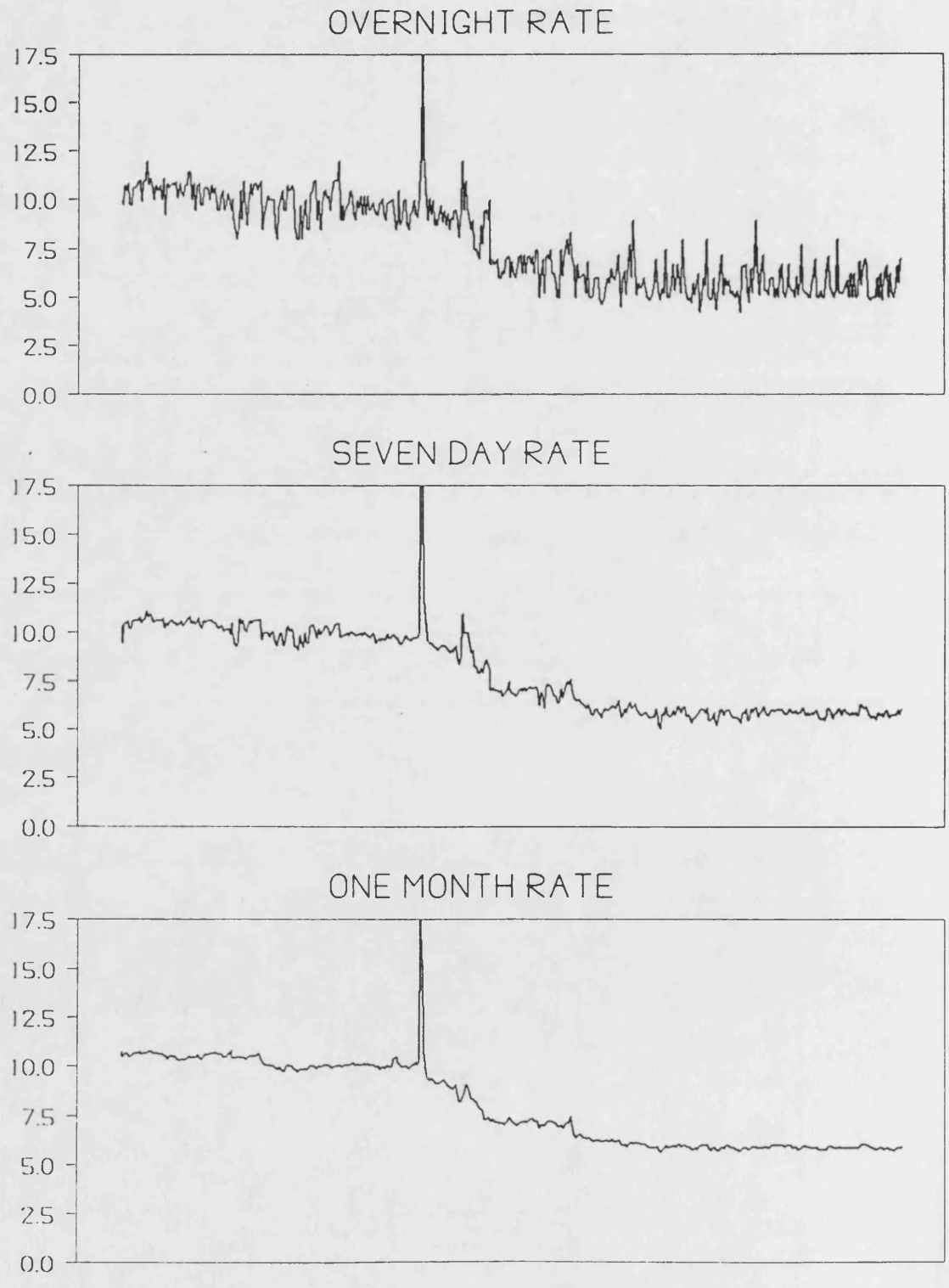


Figure 7.3

The spread between the effective overnight rate and the stop rate (Daily observations: Jan 1992 - Oct 1993)

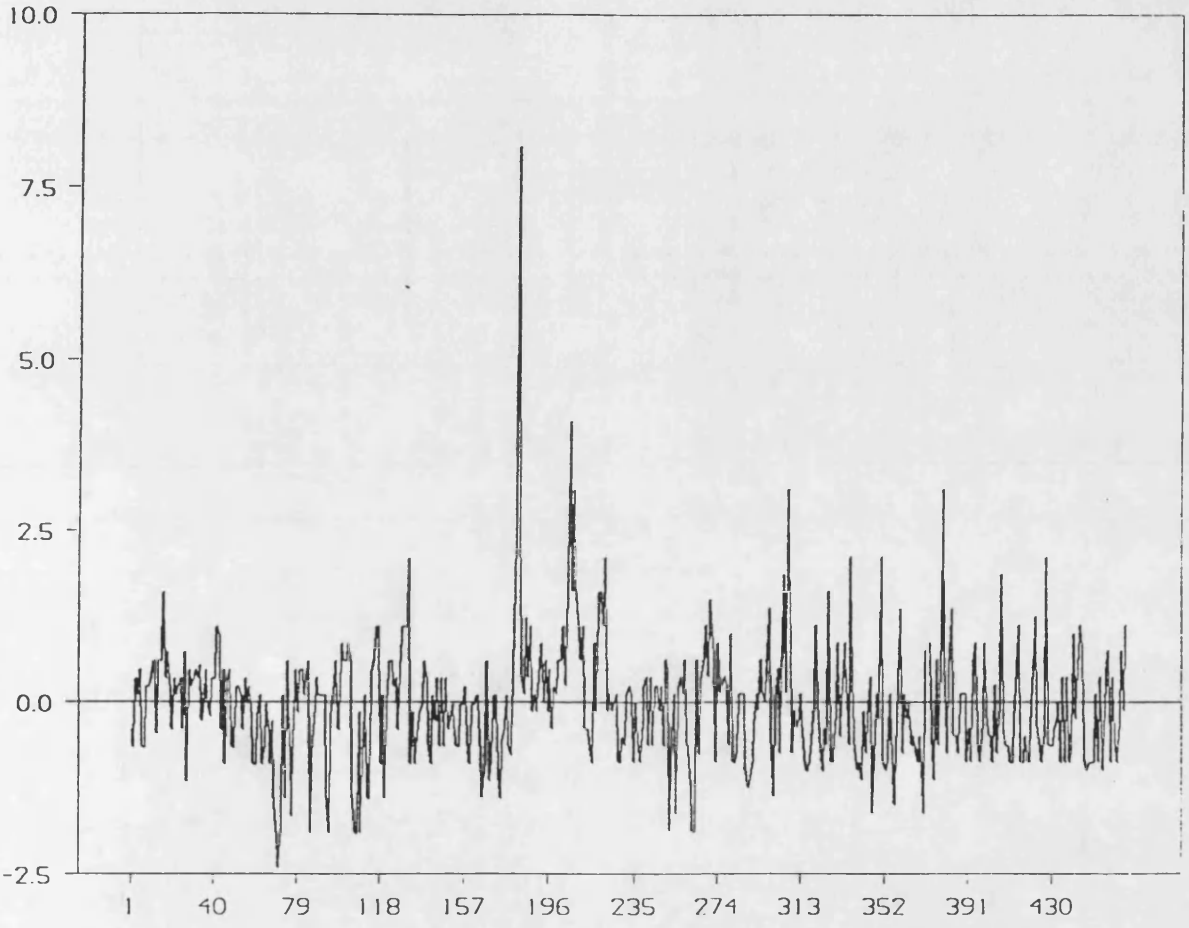
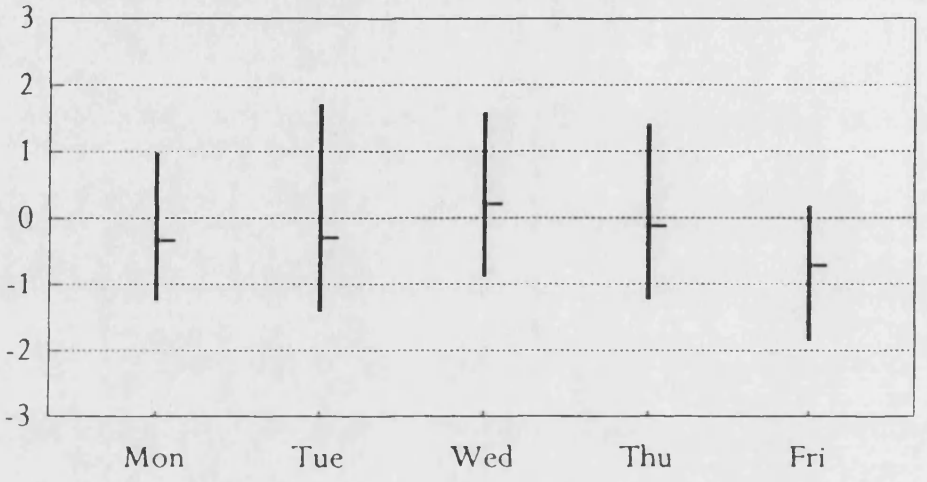


Figure 7.4

Day-of-the-week patterns in the spread

High, Low and Average



Variance and Volatility

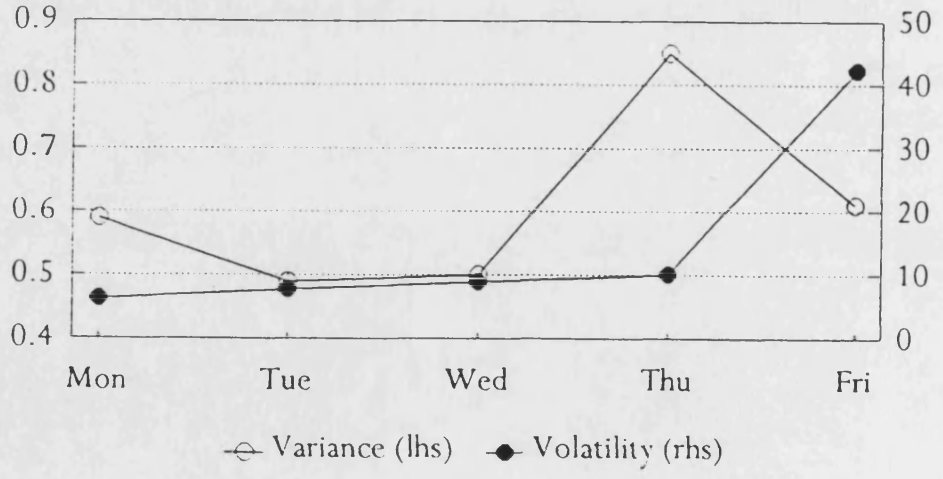


Figure 7.5

REMAINDER (y-axis) versus SIZE (x-axis)
after each round of money market operations

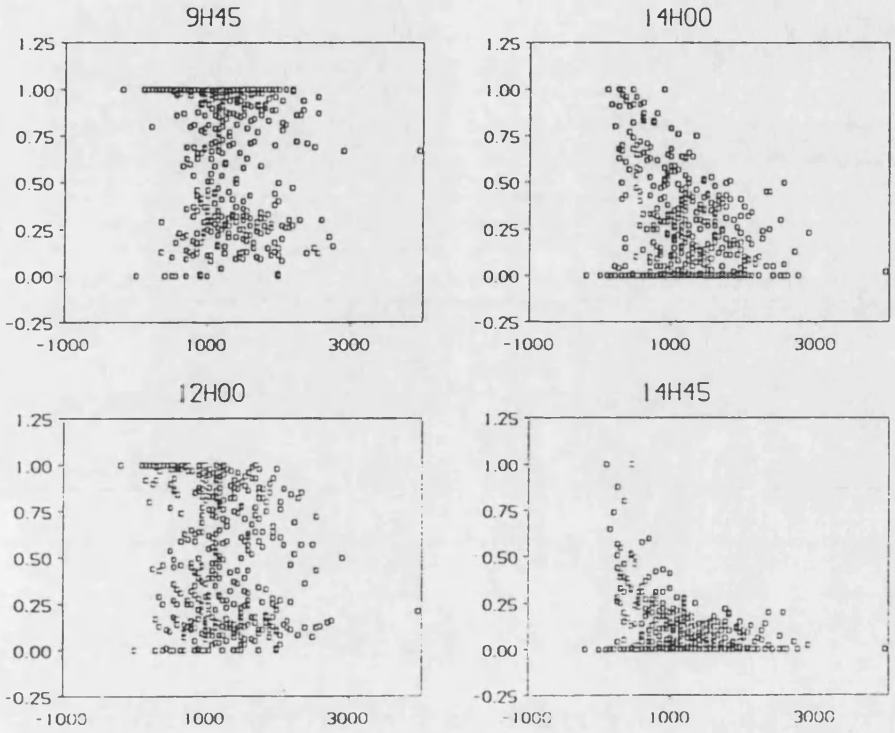


Table 7.1

Intraday behaviour of the effective overnight rate

Sample variance (Jan - May 1994)

Day	8 am	11 am	2 pm	Close
Mon	0.27	0.70	1.00	1.73
Tues	0.41	0.65	0.97	3.63
Wed	0.41	0.81	0.94	2.06
Thur	0.51	0.89	1.43	1.32
Fri	0.48	0.89	1.03	1.03
All	0.45	0.80	1.12	2.19

Table 7.2

Intraday behaviour of the effective overnight rate

Intraday changes (Jan - May 1994)

Change between ...								
8am-11am			11am-2pm			2pm-close		
μ	μ abs	σ^2	μ	μ abs	σ^2	μ	μ abs	σ^2
-0.59	0.49	0.43	-0.09	0.49	0.85	0.21	0.85	2.89

Table 7.3

Spread between effective overnight rate and stop rate (annual)

Year	sample mean	sample variance	sample volatility
1991	0.44	0.37	0.77
1992	-0.04	1.02	1.04
1993	-0.16	0.81	1.19
1994	-0.25	0.67	1.02
90-94	-0.05	0.75	1.01

Table 7.4

Spread between overnight rate extremes and stop rate (annual)

Year	Daily High		Daily Low		Volatility
	μ	σ^2	μ	σ^2	D
1991	1.43	1.99	2.37	3.09	9.41
1992	2.15	6.57	2.03	2.17	21.87
1993	1.66	6.23	1.72	1.43	17.21
1994	1.17	2.18	1.35	1.17	3.58
90-94	1.80	5.40	1.94	2.32	14.95

Table 7.5

Spread between effective overnight rate and stop rate (daily)

Year	Day	sample mean	sample variance
	Mon	0.07	0.18
	Tues	-0.07	0.69
1991	Wed	0.09	0.72
	Thur	0.14	0.94
	Fri	0.06	0.36
	Mon	-0.28	0.66
	Tues	-0.07	0.52
1992	Wed	0.11	0.92
	Thur	0.17	2.04
	Fri	-0.11	0.92
	Mon	-0.15	0.81
	Tues	-0.26	0.53
1993	Wed	0.14	0.76
	Thur	0.00	1.45
	Fri	-0.46	0.56
	Mon	-0.33	0.59
	Tues	-0.29	0.49
1994	Wed	0.21	0.50
	Thur	-0.12	0.85
	Fri	-0.71	0.61
	Mon	-0.33	0.59
	Tues	-0.29	0.49
91-94	Wed	0.21	0.50
	Thur	-0.12	0.85
	Fri	-0.71	0.61

Table 7.6

Spread between overnight rate extremes and stop rate (daily)

Year	Day	Daily High		Daily Low		Volatility
		μ	σ^2	μ	σ^2	D
	Mon	1.43	1.85	1.85	2.58	7.93
	Tues	1.02	1.79	2.85	3.48	9.38
1991	Wed	1.98	2.02	2.40	3.21	9.14
	Thur	1.98	2.50	2.38	3.34	12.78
	Fri	1.25	1.62	2.38	2.75	7.79
	Mon	1.19	1.94	2.43	2.24	7.18
	Tues	1.68	3.60	1.96	2.28	9.42
1992	Wed	2.00	3.47	1.74	2.23	10.46
	Thur	2.40	4.69	1.90	2.01	14.84
	Fri	3.49	12.93	2.11	2.13	68.19
	Mon	1.31	2.46	1.69	1.37	5.04
	Tues	1.18	2.07	1.83	1.51	4.47
1993	Wed	1.92	3.80	1.36	1.50	8.52
	Thur	1.58	1.94	1.53	1.36	4.29
	Fri	2.40	12.88	2.13	1.37	64.59
	Mon	0.98	1.30	1.27	1.16	2.27
	Tues	1.71	3.42	1.43	1.13	6.82
1994	Wed	1.59	1.41	0.93	1.34	2.76
	Thur	1.41	2.68	1.25	0.96	4.14
	Fri	0.18	0.87	1.87	1.13	1.88
	Mon	0.98	1.30	1.27	1.16	6.28
	Tues	1.71	3.42	1.43	1.13	7.68
91-94	Wed	1.59	1.41	0.93	1.34	8.78
	Thur	1.41	2.68	1.25	0.96	10.08
	Fri	0.18	0.87	1.87	1.13	42.21

Table 7.7

The money market shortage: size and timing of removal

January 1992 - October 1993

Day	Shortage SIZE (£mn)	Shortage REMAINDER at noon (%)
Mon	1249	42.3
Tues	918	57.1
Wed	1117	57.3
Thur	1176	51.0
Fri	1392	42.2
Weekend	1405	41.2
Average	1168	50.2

7.7 Appendix: Alternative GARCH models

7.7.1 Diagnostic tests

The three tests suggested by Engle and Ng (1993) are:

(i) Sign bias test

This test examines the null hypothesis that the particular GARCH specification adequately captures the impact of negative versus positive errors on the condititional variance predicted by the model. The explanatory variable which is used is

$$d_{t-1}^{-} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} > 0 . \end{cases}$$

If this dummy is significant, the model is clearly biased *against* finding that negative shocks have a different impact to positive shocks.

(ii) Negative size bias test

This test examines the null hypothesis that the GARCH specification adequately captures the impact of large versus small negative shocks on the conditional variance predicted by the model. Consequently, the explanatory variable used is

$$d_{t-1}^{-} \cdot \varepsilon_{t-1} = \begin{cases} \varepsilon_{t-1} & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} > 0 . \end{cases}$$

If this variable is significant, the model will tend to overpredict small negative shocks and underpredict large negative shocks.

(iii) Positive size bias test

This test examines the null hypothesis that the GARCH specification adequetely captures the impact of large versus small positive shocks on the conditional variance predicted by the model. Consequently, the explanatory variable used is

$$(1 - d_{t-1}^{-}) \cdot \varepsilon_{t-1} = \begin{cases} 0 & \text{if } \varepsilon_{t-1} < 0 \\ \varepsilon_{t-1} & \text{if } \varepsilon_{t-1} > 0 . \end{cases}$$

If this variable is significant, the model will tend to overpredict small positive shocks and underpredict large positive shocks.

Finally, a combined test may be conducted, which examines the joint significance of all of the three aforementioned variables on the conditional variance.

For details of how to construct the artificial regressions of the squared residuals on the explanatory variables consult Engle and Ng (1993, pp.1758-1760). The relevant test statistics are then given by the t-statistics obtained in these artificial regressions.

7.7.2 Results

	GARCH Specification		
	GARCH	**AGARCH**	**EGARCH**
Conditional mean (standard errors in parantheses)			
CONSTANT	-1.14 (0.024)	-1.14 (0.022)	-1.11 (0.030)
SPREAD_{t-1}	0.36 (0.035)	0.31 (0.029)	0.35 (0.034)
REMAINDER	1.29 (0.032)	1.29 (0.033)	1.27 (0.053)
SIZE	0.39 (0.017)	0.42 (0.016)	0.43 (0.034)
WEEKEND	-0.23 (0.060)	-0.24 (0.053)	-0.21 (0.080)
Conditional variance (standard errors in parantheses)			
ω	0.18 (0.013)	0.10 (0.012)	-0.74 (0.070)
β	0.18 (0.026)	0.25 (0.025)	0.20 (0.040)
α	0.56 (0.029)	0.49 (0.026)	0.39 (0.065)
γ		0.41 (0.030)	0.59 (0.080)
Loglikelihood	-18.80	-3.66	-1.74
Sign Bias test	-1.46	-1.49	-2.00
Negative Size Bias test	-0.76	-0.71	0.11
Positive Size Bias test	28.59	28.62	28.06
Joint test	351.4	351.0	335.0

Conclusion

Since its establishment 300 years ago the Bank of England has discounted short-term financial assets. Initially, this was purely for commercial reasons and did not really distinguish the Bank from other commercial banks. However, as the Bank's position evolved from a commercial bank to that of a central bank, its money market operations began to reflect this new status. In particular, the Bank's directors found that they could not refuse to discount assets: not, at least, without threatening a panic in the money market and a widespread "run" on the banking system. Thus the Bank's discounting operations gradually became a *routine* activity, with banks offering bills for sale to the Bank whenever the market discount rate rose up to the Bank's discount rate. Once it was routinely discounting bills, the Bank also began to appreciate that its discount rate was its primary policy instrument. Understandably, this instrument was first used in a reactive fashion, to counter reductions in the Bank's gold reserves. Over time, however, the use of the interest rate instrument has become more pro-active, to offset anticipated deviations from policy objectives[1].

[1]This is *not* to suggest that the interest rate instrument has always been succesfully deployed in meeting the policy objectives of central banks. Governments, in particular,

The history of many other central banks shows a similar development and has been discussed in *Part I* of this thesis. This has shown that, while the adoption of an interest rate instrument was a general phenomenon shared by all central banks, the particular money market operations adopted by individual central banks to support their interest rate instrument usually contained idiosyncratic elements. Furthermore, while the aim of central bank money market operations has remained broadly the same everywhere – the determination of money market interest rates – the structure of these operations has, in many cases, changed over the years.

Partly, these changes reflect the ever-changing financial environment in which such operations take place. The Bank of England's money market operations – the focus of much of this thesis – are no exception. For example, these operations initially comprised the outright purchase of prime quality bank bills, which were virtually the only money market assets at the time. When the issue of Treasury bills outstripped that of bank bills at the outset of world war one, the Bank purchased Treasury bills instead. Then, when the government's overfunding policy reversed this situation, the Bank once again purchased large quantities of bank bills.

Whilst the assets discounted by the Bank have changed from time to time, the Bank's counterparties – a small group of institutions known as discount houses – have not. This is noteworthy, if only because these counterparties arose through circumstance rather than design: early commercial banks did not wish to be openly reliant upon the Bank – a competitor – for their liquidity[2]. Although these circumstances have changed, this indirect arrangement for providing reserves has been retained for other reasons. In particular, the Bank has expressed a preference for making reserves availible through a so-called "market-oriented" arrangement as opposed to lending directly to a few large clearing banks. However, the analysis in *Part II* of this thesis suggests that the Bank's current operations are no longer fully consistent with this stated preference.

---

have exerted considerable influence over this instrument in order to pursue their own objectives. See Capie, Goodhart and Schnadt (1994).

[2]Of course, commercial banks still relied ***indirectly*** upon the Bank for liquidity.

The clearing banks now own the majority of eligible bills in issue, which makes them the Bank's *de facto* counterparties. What, then, is the purpose of the discount houses? This question has faced the houses for some time now, and several have already diversified their business in response to their diminished role in the Bank's operations. Clearly, however, this question is part of wider – and more important – issue regarding the Bank's money market operations: *should the Bank continue to base these operations upon the outright purchase of eligible bills*?

The stock of eligible bills is comparatively small relative to the Bank's daily purchases, so counterparties are permitted to offer bills over a fairly wide range of (unexpired) maturities. It has been shown (chapter 5) that this may cause very short-term interest rates to behave in a manner which actually frustrates the Bank's intentions. In particular, short-term rates may (temporarily) move in the *opposite* direction to an anticipated change in the official interest rate. Furthermore, it may cause short-term interest rates to become more volatile when market participants expect a change in the official rate. Second, eligible bills expose the Bank to credit risks, which it has sought to manage through the imposition of accepting limits on eligible acceptors. It has been illustrated (chapter 6) how eligible bills have, as a result, become "artificially scarce" and are consequently less attractive than other assets in terms of their yield. It is primarily for this reason that most eligible bills are now held by a small group of clearing banks, who then deploy their bill portfolios to generate cheap funding in the overnight market. In sum, the common symptom of these problems has been greater volatility in very short-term sterling interest rates, particularly overnight rates (chapter 7). This is disturbing, given the important role played by the overnight funds market in the reserve management of banks operating in the UK.

What is to be done? Fortunately, these problems point to a common solution, which has been suggested at various points in this thesis: *gilt repo*. Repurchase agreements involving gilt-edged securities were amongst the first open-market operations[3] of the Bank, and have recently again been adopted

[3]Recall, however, that they were originally used to drain reserves *from* the banking system, rather than to supply reserves *to* the banking system.

by the Bank as a permanent component of its money market operations. So far, however, these gilt repos have only been used in a supplementary fashion, twice a month, rather than on a day-to-day basis. There are a number of reasons why the Bank should, in future, supply reserves *exclusively* through gilt repo:

- *Maturity* – unlike outright transactions, in which the maturity of the funds lent depends upon the unexpired maturity of the security purchased, repurchase agreements would permit the Bank to lend reserves at a single maturity to all borrowers, irrespective of the unexpired maturity of the securities offered. Indeed, the exclusive use of repurchase agreements would permit the Bank to lend funds on an overnight basis, considerably reducing the scope for very short-term market interest rates to deviate from the Bank's interest rate in the manner described in chapter 5.

- *Credit risk* – unlike eligible bills, which involve some credit risks, gilts are essentially free from credit risk. By lending funds against the security of gilts, the Bank would reduce its credit-risk exposure arising from money market operations. Furthermore, the Bank would no longer have to specify which banks were considered as "eligible" acceptors, nor would it have to enforce accepting limits on these institutions as outlined in chapter 6.

- *Equity* – any financial institution which owned gilts could offer these to the Bank as security to borrow funds, greatly widening the potential counterparties in the Bank's money market operations. Although this would mean that discount houses would lose their status as official counterparties, this role has already been eroded in practice. More important, clearing banks would no longer play a disproportionately large role in the operations of the Bank, diminshing their capacity to engage in strategic behaviour of the kind discussed in chapter 7.

- *Simplicity* – the Bank's money market operations are unduly complicated as regards their objective, which is to determine money market

> interest rates by lending sufficient reserves to the banking system to cover the projected money market shortage. Not only does the Bank regularly lend funds through three different types of open market transactions (i.e. outright purchase of eligible bills, repo of eligible bills and repo of gilts), it also conducts as many as three open market operations each day. Inviting a gilt repo twice a day would be sufficient - and much simpler.

Although the Bank of England has recently made gilt repo a permanent feature of its money market operations, this change was prompted largely by an extraordinary event – namely a currency crisis in September 1992 – rather than by a decision on the part of the Bank that its operations were in need of revision. Many other features of the Bank's existing money market operations were similarly shaped by a reaction to critical events. The subsequent longevity of some of these features – for instance the discount houses – reflects the fact that these have continued to serve an intrinsically useful function during more "normal" times. Nonetheless, financial markets continue to evolve even during such periods. Being less perceptible, this evolution may give the false impression that there is no pressing need for change. This thesis has analysed this gradual development of the sterling money market in recent years and has traced its implications for short-term interest rates. The upshot is that the Bank of England should now complete the process it has initiated and adopt money market operations which can support its interest rate instrument for another 300 years.

References

Alworth, J.S. and C.E.V. Borio (1993) "Commercial paper markets: a survey", *BIS Economic Papers*, No 37, April.

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

Arrow, K.J., Harris, T. and J. Marschak (1951) "Optimal inventory policy", *Econometrica*, Volume 19, pp.250-272.

Artis, M.J. and M.K. Lewis (1981) *Monetary Control in the United Kingdom.* Phillip Allan: London.

Avriel, M. (1993) "Optimization tools for the financial manager's desk", in Zenios, S.A. (ed) *Financial Optimization.* Cambridge University Press: Cambridge, UK. pp.176-188.

Ayuso, J., Haldane, A.G. and F. Restoy (1994) "Volatility transmission along the money market yield curve", Banco de Espana, Servicio de Estudios, Documento de Trabajo No. 9403.

Balduzzi, P., Bertola, G. and S. Foresi (1993) "A model of target changes and the term structure of interest rates", National Bureau of Economic Research, *Working Paper Series*, Number 4347, April.

Baltensperger, E. (1980) "Alternative approaches to the theory of the banking firm", *Journal of Monetary Economics*, Volume 6, pp.1-37.

Bank of England (1961) "Commercial bills," *Bank of England Quarterly Bulletin*, Volume 1, Number 5, pp.26-31.

Bank of England (1962) "Bank liquidity in the United Kingdom," *Bank of England Quarterly Bulletin*, Volume 2, Number 4, pp.248-255.

Bank of England (1963) "The management of money day by day," *Bank of England Quarterly Bulletin*, Volume 3, Number 1.

Bank of England (1965) "The UK and US Treasury bill markets, " *Bank of England Quarterly Bulletin*, Volume 5, Number 4, pp.327-338.

Bank of England (1966) "Local authorities and the capital and money markets", *Bank of England Quarterly Bulletin*, Volume 6, Number 4, pp.337-347.

Bank of England (1967) "The London discount market: some historical notes," *Bank of England Quarterly Bulletin*, Volume 7, Number 2, pp.144-156.

Bank of England (1972) "Sterling certificates of deposit", *Bank of England Quarterly Bulletin*, Volume 12, Number 4, pp.487-495.

Bank of England (1980) "Methods of monetary control", *Bank of England Quarterly Bulletin*, Volume 20, Number 4, pp.428-429.

Bank of England (1981a) "Monetary control: next steps", *Bank of England Quarterly Bulletin*, Volume 21, Number 1, pp.38-39.

Bank of England (1981b) "Monetary control: provisions", *Bank of England Quarterly Bulletin*, Volume 21, Number 3, pp.347-349.

Bank of England (1981c) "The liquidity of banks", *Bank of England Quarterly Bulletin*, Volume 21, Number 1, pp.40-41.

Bank of England (1982a) "The role of the Bank of England in the money market," *Bank of England Quarterly Bulletin*, Volume 22, Number 1, pp.86-94.

Bank of England (1982b) "The supplementary special deposits scheme," *Bank of England Quarterly Bulletin*, Volume 22, Number 1, pp.74-84.

Bank of England (1982c) "A note on money market arbitrage", *Bank of England Quarterly Bulletin*, Volume 22, Number 2, pp.207-208.

Bank of England (1982d) "Bills of exchange: current issues in historical perspective," *Bank of England Quarterly Bulletin*, Volume 22, Number 4, pp.514-518.

Bank of England (1984) "Funding the public sector borrowing requirement: 1952-83," *Bank of England Quarterly Bulletin*, Volume 24, Number 4, pp.482-492.

Bank of England (1987a) "The early development of the sterling commercial paper market," *Bank of England Quarterly Bulletin*, Volume 27, Number 4, pp.527-535.

Bank of England (1987b) "Commercial paper markets: an international survey," *Bank of England Quarterly Bulletin*, Volume 27, Number 1, pp.46-53.

Bank of England (1988) *Bank of England Operations in the Money Market.* (The Red Book). October.

Bank of England (1993) "The UK approach to controlling risk in large-value payment systems," *Bank of England Quarterly Bulletin*, Volume 33, Number 4, pp.530-534.

Bank of England (1994) "The repo and secured loan facilities," *Bank of England Quarterly Bulletin*, Volume 34, Number 2, p.109.

Banque de France (1990). "Instruction No.3-90 Regarding Banque de France Intervention Procedures on the Money Market", Banque de France Official Publication, June.

Baring, F. (1797) *Observations on the Establishment of the Bank of England and on the Paper Circulation of the Country..* Reprinted in Collins, M. (ed) *Central Banking in History.* Volume I. Edward Elgar: London. pp.181-193.

Barro, R.J. (1989) "Interest rate targeting", *Journal of Monetary Economics*, Volume 23, pp.1-30.

Barsky, R.B., Mankiw, N.G., Miron, J.A. and D.N. Weill (1988) "The worldwide change in the behaviour of interest rates and prices in 1914", *European Economic Review*, Volume 32, pp.1123-1147.

Batten, D.S., Blackwell, M.P., Kim, I, Nocera, S.E. and Y. Ozeki (1990) "The

conduct of monetary policy in the major industrial countries: instruments and operating procedures", *International Monetary Fund Occasional Paper.* Washington D.C.

Beckhart, B.J. (1972) *The Federal Reserve System.* Columbia University Press: New York.

Bencivenga, V., Smith, B. and R. Starr (1992) "Liquidity of secondary capital markets: allocative efficiency and the maturity of the capital stock", Rochester Centre for Economic Research, *Working Paper*, Number 326.

Bera, A.K. and M.L. Higgins (1993) "ARCH models: properties, estimation and testing", *Journal of Economic Surveys*, Volume 7, Number 4, pp.305-362.

Bhattacharya, S. and D. Gale (1987) "Preference shocks, liquidity and central bank policy", in Barnett, W.A. and K.J. Singleton (eds) *New Approaches to Monetary Economics.* Cambridge University Press: Cambridge, UK. pp.69-88.

BIS (1993) *Payment Systems in the Group of Ten Countries.* Prepared by the Committee on Payment and Settlement Systems of the Central Banks of the Group of Ten countries, Bank for International Settlements, Basle, December.

Bollerslev, T. (1986) "Generalised autoregressive conditional heteroscedasticity", *Journal of Econometrics*, Volume 31, pp.307-327.

Bollerslev, T., Chou, R.Y. and K.F. Kroner (1992) "ARCH modeling in finance", *Journal of Econometrics*, Volume 52, pp.5-59.

Bopp, K.R. (1953) "Reichsbank operations 1876-1914", Manuscript, University of Pennsylvania.

Borio, C.E.V. and P. Van den Bergh (1993) "The nature and management of payment system risks: an international perspective", Bank for International Settlements, *Economic Papers*, Number 36, February.

Boughton, J.M. (1972) *Monetary Policy and the Federal Funds Market.* Duke University Press: Durham, N.C.

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

Breusch, T.S. and A.R. Pagan (1979) "A simple test for heteroscedasticity and random coefficient variation", *Econometrica*, Volume 50, pp.1287-1294.

British Banker's Association (1992) *Commercial Paper: London Market Guidelines.* September.

Brunner, K. and A.H. Meltzer (1964) *An Alternative Approach to the Monetary Mechanism.* A Staff Analysis for the Subcommittee on Domestic Finance of the House Committee on Banking and Currency. Washington D.C. August.

Brunner, A.D. and C.S. Lown (1993) "The effects of lower reserve requirements on money market volatility", *American Economic Review*, March, pp.199-205.

Burgess, W.R. (1946) *The Reserve Banks and the Money Market.* (Revised Edition). Harper and Brothers: New York.

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

Campbell, J.Y. (1987) "Money announcements, the demand for bank reserves, and the behaviour of the federal funds rate within the statement week", *Journal of Money, Credit and Banking*, Volume 19, Number 1, pp.56-67.

Capie, F., Goodhart, C.A.E. and N. Schnadt (1994) "The development of central banking", in *The Future of Central Banking.* Cambridge University Press: Cambridge, UK. Forthcoming.

Chen, A.H. and S.C. Mazumdar (1992) "An instantaneous control model of bank reserves and federal funds management", *Journal of Banking and Finance*, Volume 16, pp.1073-1095.

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

Clark, T.A. (1986) "Interest rate seasonals and the Federal Reserve", *Journal of Political Economy*, Volume 94, Number 1, pp.76-125.

Coleby, A.L. (1983) "The Bank's operational procedures for meeting monetary objectives", *Bank of England Quarterly Bulletin*, Volume 23, Number 2, pp.209-215.

Cosimano, T.F. (1987) "Reserve accounting and variability in the federal funds market", *Journal of Money, Credit and Banking*, Volume 19, Number 2, pp.198-209.

Cox, A.H. and R.F. Leach (1964a) "Defensive open-market operations and the reserve settlement periods of member banks", *Journal of Finance*, Volume 19, March, pp.76-93.

Cox, A.H. and R.F. Leach (1964b) "Open-market operations and reserve settlement periods: a proposed experiment", *Journal of Finance*, Volume 19, March, pp.76-93.

Cukierman, A. (1991) "Why does the Fed smooth interest rates?" in Belongia, M.T. (ed) *Monetary Policy on the 75th Anniversary of the Federal reserve System.* Kluwer: London, pp.111-144.

Cukierman, A. (1992) *Central Bank Strategy, Credibility and Independence: Theory and Evidence.* MIT Press: Cambridge, Mass.

Cuthbertson, K. (1992) "The expectations hypothesis of the term structure: the U.K. interbank market", University of Newcastle upon Tyne, *Discussion Paper* Number 92-04, June.

Dale, S. (1992) "The effect of changes in official money market interest rates since 1987", *Supplement to The Manchester School*, Volume LXI,

pp.76-94.

Dale, S., Haldane, A., Murphy, G. and N. Schnadt (1994) "On the determination of overnight interest rates in the UK", Manuscript, Bank of England.

Donaldson, R.G. (1993) "Financing banking crises: Lessons from the panic of 1907", *Journal of Monetary Economics*, Volume 31, pp.69-95.

Donaldson, R.G. (1994) "Sources of panics: evidence from the weekly data", *Journal of Monetary Economics*, Volume 32, Number 2, pp.277-306.

Dow, J.C.R. and I.D. Saville (1988) *A Critique of Monetary Policy: Theory and British Experience.* Clarendon Press: Oxford.

Dutkowsky, D.H. (1993) "Dynamic implicit cost and discount window borrowing: an empirical investigation", *Journal of Monetary Economics*, Volume 32, pp.105-120.

Dvoretsky, A., Kiefer, J. and J. Wolfowitz (1952) "The inventory problem I: cases of known distribution of demand", *Econometrica*, Volume 20, pp.187-222.

Dyl, E. and J.R. Hoffmeister (1985) "Efficiency and volatility in the Federal funds market", *Journal of Bank Research*, Winter, pp.234-239.

Edgeworth, F.Y. (1888) "The mathematical theory of banking", *Journal of the Royal Statistical Society*, Volume LI, Part I, march, pp.113-127.

Eichengreen, B. (1992) *Golden Fetters.* Oxford University Press: Oxford.

Engle, R.F. (1982) "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, Volume 50, Number 4, pp.987-1007.

Engle, R.F. and V.K. Ng (1993) "Measuring and testing the impact of news on volatility", *Journal of Finance*, Volume XLVIII, Number 5, December, pp.1749-1778.

Englund, P., Horngren, L. and S. Viotti (1989) "Discount window borrowing and money market interest rates", *Scandanavian Journal of Economics*, Volume 91, Number 3, pp.517-533.

Evanoff, D.E. (1990) "An empirical examination of bank reserve management behaviour", *Journal of Banking and Finance*, Volume 14, pp.131-143.

Evans, P. (1981) "Why have interest rates been so volatile?", Federal Reserve Bank of San Francisco *Economic Review*, Summer, pp.7-20.

Feinman, J.N. (1993a) "Estimating the Open-Market Desk's daily reaction function", *Journal of Money, Credit and Banking*, Volume 25, Number 2, pp.231-247.

Feinman, J.N. (1993b) "Reserve requirements: history, current practice and potential reform", *Federal Reserve Bulletin*, June, pp.569-589.

Flemming, J.S. and D.G. Barr (1989) "Modelling money market interest rates", *Bank of England Technical Paper Series*, 24.

Foot, D.K.W., Goodhart, C.A.E. and A.C. Hotson (1980) "Monetary base control", *Bank of England Quarterly Bulletin*, Volume 19, Number 2, pp.149-159.

Friedman, M. (1959) *A Program for Monetary Stability.* Fordham University Press: New York.

Frost, P.A. and T.J. Sargent (1970) "Money market rates, the discount rate, and borrowing from the Federal Reserve", *Journal of Money, Credit and Banking*, Volume 2, pp.56-82.

Frost, P.A. (1971) "Banks' demand for excess reserves", *Journal of Political Economy*, Volume 79, pp.805-825.

Garber, P. and S. Weisbrod (1990) "Banks in the market for liquidity", National Bureau of Economic Research, *Working Paper Series*, Number 3381, June.

Garfinkel, M. (1990) "The FOMC in 1989: walking a tightrope", *Federal*

Reserve Bank of St Louis Review, Number 72, pp.17-35.

Gillett Brothers (1952) *The Bill on London.* Chapman and Hall: London.

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

Goldfeld, S.M. and E.J. Kane (1966) "The determinants of member bank borrowing: an econometric study", *Journal of Finance*, Volume 21, pp.499-514.

Goodfriend, M. (1983) "Discount window borrowing, monetary policy, and the post-October 6, 1979 Federal Reserve operating procedure", *Journal of Monetary Economics*, Volume 12, September, pp.343-56.

Goodfriend, M. (1987) "Interest rate smoothing and price level trend-stationarity", *Journal of Monetary Economics*, Volume 19, May, pp.335-48.

Goodfriend, M. (1991) "Interest rates and the conduct of monetary policy", *Carnegie-Rochester Conference Series on Public Policy*, Volume 34, Spring, pp.7-30.

Goodfriend, M. (1993) "Interest rate policy and the inflation scare problem: 1979-1992", *Federal Reserve Bank of Richmond Economic Quarterly*, Volume 79, number 1, pp.1-24.

Goodfriend, M. and Hargreaves, M. (1983) "A historical assessment of the rationales and functions of reserve requirements", *Federal Reserve Bank of Richmond Economic Review*, March/April, pp.3-21.

Goodfriend, M and King, R.G. (1988) "Financial deregulation, monetary policy and central banking", *Federal Reserve bank of Richmond Economic Review*, Volume 74, Number 3, pp.3-22.

Goodfriend, M. and W. Whelpley (1986) "Federal funds: instrument of Federal Reserve policy", *Federal Reserve Bank of Richmond Economic Review*, Volume 72, September/October, pp.3-11.

Goodhart, C.A.E. (1969) *The New York Money Market and the Finance of Trade: 1900-1913.* Harvard University Press: Cambridge, Mass.

Goodhart, C.A.E. (1972) *The Business of Banking, 1891-1914.* Weidenfeld and Nicholson: London.

Goodhart, C.A.E. (1984) *Monetary Theory and Practice: The UK experience.* MacMillan: London.

Goodhart, C.A.E. (1987) *Money, Information and Uncertainty.* MacMillan: London.

Goodhart, C.A.E. (1989a) "Monetary Base", in Eatwell, M. , Milgate, M. and P. Newman (eds) *The New Palgrave: Money.* MacMillan: London. pp.206-211.

Goodhart, C.A.E. (1989b) "The conduct of monetary policy", *Economic Journal*, Volume 99, June, pp.293-346.

Greene, W.H. (1993) *Econometric Analysis.* 2nd edition. MacMillan: London.

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

Gumbel, E.J. (1958) *Statistics of Extremes.* Columbia University Press: New York

Hamdani, K, and S. Peristiani (1991) "A disaggregate analysis of discount window borrowing", *Federal Reserve Bank of New York Quarterly Review*, Volume 16, Number 2, pp. 52-62.

Hamilton, J.D. (1994) *Time Series Analysis.* Princeton University Press: Princeton.

Harvey, C.R. and R.D. Huang (1994) "The impact of the Federal Reserve bank's open market operations", National Bureau of Economic Research, *Working Paper Series*, Number 4663, February.

Hawtrey, R.G. (1938) *A Century of Bank Rate.* Longmans, Green and Co: London.

Harrington, R. (1991) "The London financial markets", in Green, C.J. and Llewellyn, D.T. (eds) *Surveys in Monetary Economics.* Blackwell: Oxford.

Hedley, W. (1986) *Bills of Exchange and Bankers' Documentary Credits.* Lloyd's of London Press: London.

Helfrich, R.T. (1976) "Trading in bankers' acceptances: a view from the acceptance desk of the Federal Reserve Bank of New York," Federal Reserve Bank of New York *Monthly Review*, February.

Hill, E. (1986) "Bankers acceptances," in Cook, T.Q. and T.D. Hahn (eds) *Money Market Instruments.* Federal Reserve Bank of Richmond, VA.

Ho, T.S.Y. and A. Saunders (1985) "A micro model of the federal funds market", *Journal of Finance*, Volume XL, Number 3, pp.977-987.

Holden, J.M. (1955) *The History of Negotiable Instruments in English Law.* Athlone Press: London.

Hols, M. and C.G. de Vries (1991) "The limiting distribution of extremal exchange rate returns", *Journal of Applied Econometrics*, Volume 6, Number 3, pp.287-302.

Howard, D.H. (1982) "The British banking system's demand for cash reserves", *Journal of Monetary Economics*, Volume 9, pp.21-41.

Hurn, A.S., T. Moody and V.A. Muscatelli (1993). "The term structure of interest rates in the London interbank market", *ICMM Discussion Paper*, University of Strathclyde.

James, J. (1991) *Richardson's Guide to Negotiable Instruments.* (Eighth Edition). Butterworths: London.

Jensen, F. and P. Parkinson, (1986) "Recent developments in the bankers acceptance market," *Federal Reserve Bulletin*, January, pp.1-12.

Johnson, N.L. and Kotz, S. (1970) *Continuous Univariate Distributions.* John-Wiley and Sons: New York.

Kasman, B. (1992) "A comparison of monetary operating procedures in six industrial countries", *Federal Reserve Bank of New York Quarterly Review*, Vol 17, Number 2, pp.5-24.

Kasriel, P.L. (1981) "Interest rate volatility in 1980", Federal Reserve bank of Chicago *Economic Perspectives*, January/February, pp.8-17.

King, W.T.C. (1936; reprinted 1973) *History of the London Discount Market.* Frank Cass: London.

King, W. (1947) "The changing discount market," *The Banker*, March.

King, W. (1947) "War and the money market," *Journal of the Institute of Bankers*, April.

Kiyotaki, N. and R. Wright (1989) "On money as a medium of exchange", *Journal of Political Economy*, Volume 97, Number 4.

Kneeshaw, J.T. and P. Van den Bergh (1989) "Changes in central bank money market operating procedures in the 1980s", Bank for International Settlements *Economic Papers*, Number 23, January.

Knobel, A. (1977) "The demand for reserves by commercial banks", *Journal of Money, Credit and Banking*, September, pp.32-47.

LaRoche, R.K. (1993) "Bankers acceptances," *Federal Reserve Bank of Richmond Economic Quarterly*, Volume 79, Number 1, Winter, pp.75-85.

Lasser, D.J. (1992) "The effect of contemporaneous reserve accounting on the market for Federal funds", *Journal of Banking and Finance*, Volume 16, pp.1047-1056.

Laurent, R.D. (1979) "Reserve requirements: are they lagged in the wrong direction?", *Journal of Money, Credit and Banking*, Volume 13, pp.301-310.

Law, R. (1965) "The resurgence of the commercial bill," *The Bankers' Magazine*, Volume 3, Number 6, pp.341-347.

Lippman, S.A. and J.J. McCall (1986) "An operational measure of liquidity", *American Economic Review*, Volume 76, Number 2, pp. 43-55.

Llewellyn, D.T. (1990). "Money market operations of the Bank of England and the determination of interest rates", in T.Bandyopadhyay and S.Ghatak (eds) *Current Issues in Monetary Economics.* Harvester Wheatsheaf: London.

Llewellyn, D.T. and B. Tew (1993) "The sterling money market", *Banking World*, January, pp.21-25.

Mackinnon, N. and M. Dowding (1989) *The Sterling Money Market.* IFR: London.

Mankiw, N.G. and J.A. Miron (1986) "The changing behaviour of the term structure of interest rates", *Quarterly Journal of Economics*, Volume 51, Number 2, pp.211-228.

Mankiw, N.G. and J.A. Miron (1991) "Should the Fed smooth interest rates: The case of seasonal monetary policy", *Carnegie-Rochestor Conference Series on Public Policy*, Volume 34.

Martensson, A.W. (1992) " The interest rate scale", Sverige Riksbank *Quarterly Review*, Number 4, pp.13-18.

McCallum, B.T. (1987) "The case for rules in the conduct of monetary policy: a concrete example", Federal Reserve bank of Richmond *Economic Review*, Number 73, September/October, pp.10-18.

McCallum, B.T. (1988) "Robustness properties of a rule for monetary policy", *Carnegie-Rochester Conference Series on Public Policy*, Number 29, Autumn, pp.173-203.

McCallum, B.T. (1989) *Monetary Economics: Theory and Policy.* MacMillan: New York.

Meek, P. (1985) *Open Market Operations.* Federal Reserve Bank of New York.

Meulendyke, A. (1992) "Monetary policy and open market operations during 1991", Federal Reserve Bank of New York *Quarterly Review*, Spring, pp. 72-95.

Mengle, D. (1986) "The discount window", in Cook, T.Q. and T.D. Rowe (eds) *Instruments of the Money Market.* Federal Reserve Bank of Richmond. pp.23-35.

Meulendyke, A. (1990) *U.S. Monetary Policy and Financial Markets.* Federal Reserve bank of New York.

Midland Bank (1961) "The Treasury bill: the story of an economist's invention," *Midland Bank Review*, February.

Minsky, H.P. (1957) "Central banking and money market changes", *Quarterly Journal of Economics*, Volume LXXI, Number 2, pp.171-187.

Miron, J.A. (1986) "Financial Panics, the seasonality of the nominal interest rate, and the founding of the Fed", *American Economic Review*, Volume 76, Number 1, pp.125-140.

Mishkin, R. (1991) "Asymmetric information and financial crises: a historical perspective", in G. Hubbard (ed) *Financial Markets and Financial Crises.* The University of Chicago Press: Chicago.

Morrison, G.R. (1966) *Liquidity Preferences of Commercial Banks*, University of Chicago Press: Chicago.

Nelson, D.B. (1991) "Conditional heteroscedasticity in asset returns: a new approach", *Econometrica* Volume 59, Number 2, pp.347-370.

Nishimura, S. (1971) *The Decline of Inland Bills of Exchange in the London Money Market 1855-1913.* Cambridge University Press: Cambridge.

Palgrave, R.H.I. (1903) *Bank Rate and the Money Market in England, France, Germany, Holland and Belgium 1844-1900.* John Murray: London.

Parkinson, M. (1980) "The extreme value method for estimating the variance of the rate of return", *Journal of Business*, Volume 53, Number 1, pp.61-65.

Pearce, D.K. (1993) "Discount window borrowing and Federal Reserve operating regimes", *Economic Inquiry*, Volume XXXI, pp.564-579.

Persitiani, S. (1991) "The model structure of discount window borrowing", *Journal of Money, Credit and Banking*, Volume 23, Number 1, pp.13-34.

Poole, W. (1968) "Commercial bank reserve mangement in a stochastic model: implications for monetary policy", *Journal of Finance*, Volume 23, pp.769-791.

Poole, W. (1970) "Optimal choice of monetary policy instruments in a simple stochastic macro model", *Quarterly Journal of Economics*, Volume 84, May, pp.197-216.

Poole, W. (1976) "A proposal for reforming bank reserve requirements in the United States", *Journal of Money, Credit and Banking*, Volume 8, pp.137-47.

Poole, W. (1990) "Central control of interest rates: a commentary", in Haraf, W.S. and Cagan, P. (eds) *Monetary Policy for a Changing Financial Environment.* AEI Press: New York.

Pulli, M. (1992) "Overnight market interest rates and banks' demand for reserves in Finland", *Bank of Finland Institute for Economic Research*, Series B:47.

Radcliffe Committee (1958) *Report of the Committee on the Working of the Monetary System.* Cmnd.827. London:HMSO.

Rasche, R.H. (1985) "Interest rate volatility and alternative monetary control procedures", Federal Reserve Bank of San Francisco *Economic Review*, Summer, pp.46-63.

Reid, M. (1982) *The Secondary Banking Crisis, 1973-75*, MacMillan: London.

Santomero, A.M. (1984) "Modeling the banking firm: a survey", *Journal of Money, Credit and Banking*, Volume 16, Number 4, pp.576-602.

Saunders, A. and T. Urich (1988) "The effects of shifts in monetary policy and reserve accounting regimes on bank reserve management behaviour in the Federal funds market", *Journal of Banking and Finance*, Volume 12, pp.523-535.

Saunders, A. (1994) *Financial Institutions Management: A Modern Perspective.* Irwin: Illinois.

Sayers, R.S. (1947) "Central banking in the light of recent British and American experience", *Quarterly Journal of Economics*, Volume LXIII, Number 1, pp.198-211.

Sayers, R.S. (1957) *Central Banking After Bagehot.* Clarendon Press: Oxford.

Sayers, R.S. (1976) *The Bank of England 1891 - 1944.* Cambridge University Press: Cambridge, UK.

Scammel, W.M. (1968) *The London Discount Market.* Elek Books:London.

Schnadt, N. (1994a) *The Domestic Money Markets of the UK, France Germany and the US.* Subject Report VII, Paper 1, City Research Project. London Business School. January.

Schnadt, N. (1994b) *European Monetary Union and the Sterling Money Market.* Subject Report VII, Paper 2, City Research Project. London Business School. January.

Schnadt, N. (1994c) "Reforming the sterling money market", *Banking World.* April.

Schnadt, N. and J. Whittaker (1992) "A suggestion for the operating practices of the European Central Bank", *Economies et Societés*, Série

Economie monétaire, Number 9, September-October, pp.133-145.

Schnadt, N. and J. Whittaker (1993a) "Inflation proof currency? The feasibility of variable commodity standards", *Journal of Money, Credit and Banking*, Volume 25, Number 2, pp.214-221.

Schnadt, N. and J. Whittaker (1993b) "The viability of an 'indirectly convertible' gold standard", *Southern Economic Journal*, Volume 60, Number 2, pp.495-500.

Schnadt, N. and J. Whittaker (1994a) "Is indirect convertibility impossible? A reply to the comment by L. Yeager, R. Greenfield, and W. Woolsey", *Journal of Money, Credit and Banking*, forthcoming.

Schnadt, N. and Whittaker (1994b) "Optimal money market behaviour and sterling interest rates", *The Manchester School*, forthcoming.

Smith, V. (1936; reprinted 1990) *The Rationale of Central Banking.* Liberty*Press*: Indianapolis.

Spindt, P.A. and J.R. Hoffmeister (1988) "The micromechanics of the federal funds market: implications for day-of-the-week effects in funds rate variability", *Journal of Financial and Quantitative Analysis*, Volume 23, Number 4, pp.401-415.

Sprague, O.M.W. (1910) "The Treasury and the Panic", *History of Crises under the National Banking System.* National Monetary Commission, US Senate, 61st Congress. Reprinted in Collins, M. (ed) *Central Banking in History.* Volume I. Edward Elgar: London. pp.323-327.

Sternlight, P.D. (1964) "Reserve settlement periods of member banks: comment", *Journal of Finance*, Volume 19, March, pp.94-98.

Stevens, E.J. (1991) "Is there any rationale for reserve requirements?", *Federal Reserve Bank of Cleveland Economic Review*, Number 3, pp.2-17.

Stigum, M. (1990) *The Money Market.* 3rd Edition. Irwin: Illinois.

Summers, L.H. (1991) "Planning for the next financial crisis", in Feldstein, M (ed) *The Risk of Economic Crisis.* Chicago University Press: Chicago.

Tallman, E.W. and J.R. Moen (1993) "Liquidity shocks and financial crises during the National Banking era", Federal Reserve Bank of Atlanta, *Working Paper Series*, Number 93-10, August.

Taus, E.T. (1943) *Central Banking Functions of the United States Treasury 1789 - 1941.* Columbia University Press: New York.

Taylor, M.P. (1992). "Modelling the yield curve", *Economic Journal*, Volume 102, May, pp. 524-37.

Timberlake, R.H. (1978; reprinted 1993) *Monetary Policy in the United States: An Intellectual and Institutional History.* University of Chicago Press: Chicago.

Treasury Bulletin (1992) "Government funding: working out official gilt sales".

Vickson, R.G. (1985) "Simple optimal policy for cash management: the average balance requirement case", *Journal of Financial and Quantitative Analysis*, Volume 20, Number 3, pp.353-369.

Waller, C.J. (1990) "Administering the window: a game-theoretic model of discount window borrowing", *Journal of Monetary Economics*, Volume 25, pp.273-287.

Walmsley, J. (1992) *The Foreign Exchange and Money Markets Guide.* Wiley: New York.

Walsh, C.E. (1990) "Issues in the choice of monetary policy operating procedures", in Haraf, W.S. and Cagan, P. (eds) *Monetary Policy for a Changing Financial Environment.* AEI Press: New York.

Warburg, P.M. (1911) "A United Reserve Bank of the United States", *Proceedings of the Academy of Political Science in the City of New York.*

Reprinted in Collins, M. (ed) *Central Banking in History.* Volume I. Edward Elgar: London. pp.323-327.

Wheelock, D.C. (1992) "Seasonal accomodation and the financial crises of the Great depression: did the Fed 'furnish an elastic currency?', Federal Reserve Bank of St. Louis *Review*, November/December, pp.3-18.

Wicker, E.R. (1965) "Federal Reserve monetary policy, 1922-33: a reinterpretation", *Journal of Political Economy*, Volume 73, August. pp.325-43.

Wicker, E.R. (1966) *Federal Reserve Monetary Policy 1917-33.* Random House: New York.

Wicker, E.R. (1992) "The US central banking experience: a historical perspective", in Fratianni, M.U. and D. Salvatore (eds) *Monetary Policy in Developed Economies.* North-Holland: Amsterdam.

Wilson, J.S.G. (1989) "The London money markets", (revised edition) *SUERF Papers on Monetary Policy and Financial Systems*, No. 6, Tilburg.