1

Introduction

Government securities are an essential item of equipment for an effective state. They enable governments to borrow to finance expenditure which cannot immediately be paid for out of taxation or accumulated savings. In the past, such expenditure was often for the conduct of wars. Britain was able to raise more money than France to fight the Napoleonic Wars because it had better arrangements for government borrowing. Governments which could not borrow have often resorted to creating money, which its citizens are forced to accept as payment, leading to serious inflation; that is why wars are frequently accompanied by inflation.[1]

At the end of the Napoleonic Wars, in 1815, the British national debt was 214% of gross domestic product, according to present-day estimates of GDP. By 1913, the ratio had come down to 28%.[2] There had been no significant price inflation in the meantime, the gold convertibility of the pound having been restored in 1821. The reduction in the debt ratio was achieved through the combination of economic growth and a sustained balanced budget. Britain's already-strong credit standing was further enhanced.

The First World War increased the debt/GDP ratio to 140%, and there was considerable inflation: the cost of living index rose by 120% during the war. Owing to deflationary policies and slow economic growth, and despite balanced budgets, the debt ratio did not fall back after the war: by 1939, it was 153%. Economic management during the Second World War relied

[1] On the Napoleonic Wars see Ferguson (2001, p. 180); also Bordo and White (1993), Sargent and Velde (1995) and Bernholz (2015, ch. 6), who presents a complete theory of inflation.

[2] Source: Bank of England, 'A millennium of data', www.bankofengland.co.uk/research/Pages/datasets/default.aspx, tables A29 and M6.

heavily on controls over prices and borrowing. The cost of living index rose by 31% during that war, and the debt/GDP ratio had risen to 259% in 1946.

Since the First World War, government borrowing has been a routine feature of peacetime economic management. Large amounts of government debt are outstanding, and large amounts need to be borrowed each year, not only to finance ongoing deficits but also to refinance maturing debts. Thus in the financial year 2016/17, official government bond sales were £147.6 billion, or about 7.5% of GDP. Roughly half of this borrowing was accounted for by the refinancing of maturing debt. Failure to sell these bonds would have led to creation of near-money assets of the same amount, increases in inflation and inflationary expectations, and the destruction of the government's monetary policy. For this reason, governments are reliant on the existence of markets in which their bonds can be sold, as are central banks, which could not otherwise achieve their price-stability objectives. Market liquidity – the ability to buy and sell easily – matters a lot. It makes government securities more attractive to investors; and it thereby enables governments to borrow more, and more cheaply.

This book is largely the story of how the market for UK government bonds – known as gilt-edged, or gilts – developed in the middle of the twentieth century, of how the monetary authorities tried to compensate for its deficiencies, and of how they overcame the unintended consequences of their actions. Specifically, it is about official intervention by the British monetary authorities in the secondary market for gilts from 1928 to 1972. Its main purpose is to describe how the intervention was conducted and to what ends.

The Bank of England, which was the agency responsible for government debt management, was motivated not only by immediate needs to sell gilts to finance the government, but also by the desire to maintain the liquidity of the gilt market. The latter became more difficult during and after the Second World War, when the quantity of gilts outstanding exploded, and the capacity of the commercial market-makers could not keep up. The Bank's operations, which were conducted in great secrecy, led to conflicts with monetary policy. These were partly resolved in 1971.

The Bank went to great lengths to ensure the continued presence of commercial market-makers. In 1931 it acted as lender of last resort to the Stock Exchange jobbers, and on several occasions in the 1950s and 1960s it subsidised them to keep them in business. During the 1950s and 1960s, the Bank played a much larger role itself as a market-maker in gilt-edged than has generally been appreciated.

The role of central banks as 'market-maker of last resort' has been discussed extensively in the context of the crisis of 2008–09. This study shows that, long before that, the Bank of England had acted as the market-maker of last resort – and perhaps at times of first resort – for a protracted period, and demonstrates that market microstructures can have important macroeconomic implications.[3]

There are already several accounts of debt management in Britain in the twentieth century,[4] to which the present account should be seen as a supplement. My main contribution is to say more about official operations in the secondary market, and to draw attention to the under-appreciated connection between the microstructure of the gilt-edged market and monetary policy.[5] It is based on looking at what the Bank of England did, as well as what it said.

The account begins in 1928, when the Treasury currency notes which had been introduced in 1914 were absorbed into the Bank of England note issue, and the consequent massive enlargement of the assets of the Issue Department of the Bank provided it for the first time with the resources to intervene in the gilt market on a substantial scale.[6] It ends in 1972, partly because the Competition and Credit Control programme of 1971, which

[3] Recent discussions of central banks acting as market-maker of last resort, in the context of financial crisis, can be found in Buiter and Sibert (2008, pp. 171–8), Tucker (2009) and Mehrling (2011).

[4] They include Wormell's book (1999) on 1900–32, Howson's paper (1988) on 1932–51, and Sayers' book (1956) on 1939–45. In addition, the successive histories of the Bank of England by Sayers, Fforde and Capie (1976, 1992 and 2010 respectively), Howson's accounts of British monetary policy from 1919–38 (1975) and 1945–51 (1993), Nevin's account of the mechanism of cheap money (1955), Dow's history of macroeconomic management from 1945–60 (1964), my own account of British monetary policy from 1951–59 (Allen, 2014), the accounts of the 1960s by Cohen (1971) and Tew (1974), and Needham's account of post-1967 monetary policy (2014) all have plenty to say about debt management. And there is the Bank of England's reporting of its own activities in its *Quarterly Bulletin* and elsewhere.

[5] The work of Ranald Michie (1999) and Bernard Attard (2000) on the history of the Stock Exchange has been invaluable in the writing of this book. Attard's paper describes the conditions in which the jobbers worked and of the relationships which developed within the Stock Exchange. It is partly based on his fascinating project on the jobbing system of the London Stock Exchange undertaken under the auspices of the Centre for Metropolitan History at London University, shortly after the Big Bang of 1986 ended single capacity in the Stock Exchange and, with it, the distinction between brokers and jobbers. Records of interviews with individual jobbers can be found at www.history.ac.uk/projects/research/jobbing (last accessed 28 March 2016).

[6] The accounts of the Bank of England are divided into two parts, the Banking Department and the Issue Department. The Issue Department's only liabilities can be Bank of England notes.

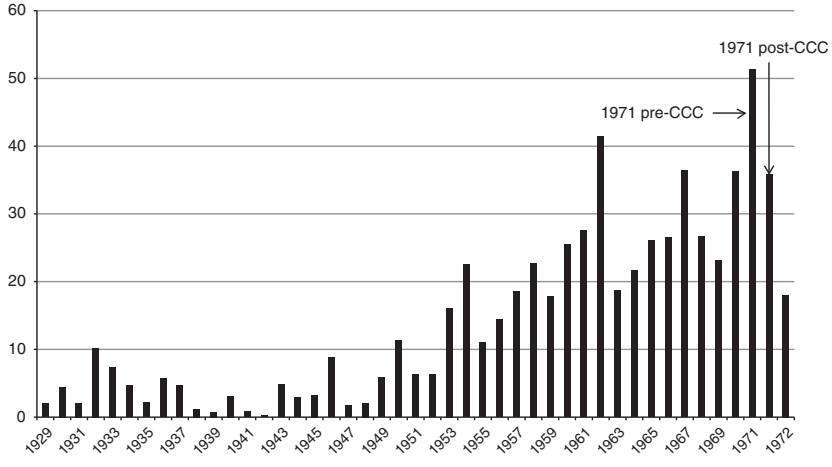

Figure 1.1 Issue Department Secondary Market Turnover with Market Counterparties (Percentage of Outstanding Gilts), 1929–72

curtailed the Bank's intervention in the gilt market, represents a natural break point, and partly because I joined the Bank of England staff in 1972, and do not believe in writing the history of events in which I was personally involved, even if only peripherally. Intervention nevertheless continued after 1972 until 1986, when Big Bang in the London Stock Exchange led to greatly increased liquidity, and ultimately made it possible for gilts to be sold by auction with no need for official intervention (Figure 1.1).[7]

[7] It was not until more than a decade after Big Bang that all primary market sales of gilts were conducted by auction, but it was the increase in market liquidity that Big Bang created that made auctions possible. For an early assessment of the liquidity effect of Big Bang, see Bank of England (1989).

2

Price and Quantity Discovery, Market-Making and Liquidity in the Gilt Market

Standard economic theory cannot readily accommodate the concept of market liquidity. In models of perfect competition, prices depend on the supply and demand schedules of the participants in the economy, none of whom is important enough to have a perceptible effect on the market price and all of whom therefore take prices as given: they are price-takers, not price-makers. In the models, as Kenneth Arrow pointed out, 'there is no one left over whose job it is to make a decision on price.'[1]

> The job is, in fact, entrusted to a *deus ex machina*: Walras' auctioneer is assumed to inform all traders of the prices at which all markets are going to clear. This always trustworthy information is supplied at zero cost. Traders do not have to wrestle with situations in which demands and supplies do not mesh; all can plan on facing perfectly elastic demand and supply schedules without fear of ever having their trading plans disappointed. All goods are perfectly 'liquid', their full market values being at any time instantaneously realizable. Money can be added to such models only by artifice.[2]

The lack of realism has serious consequences. According to one influential interpretation, the target of Keynes' attack on 'classical economics', and its inability to explain mass unemployment, was its assumption of instantaneous market-clearing, and its failure to explore the processes of price and quantity discovery, in particular in the labour market.[3] Much modern macroeconomic theory has been devoted to surmounting, or circumventing, the theoretical difficulty posed by the absence of a procedure to determine prices in models of a perfectly competitive

[1] Arrow (1959, p. 43).
[2] Leijonhufvud (1981, p. 6). The reference is to Léon Walras's *Éléments d'économie politique pure*, first published in 1874.
[3] Leijonhufvud (1968).

market.[4] Obviously, it is logically impossible to draw inferences about the optimality, or otherwise, of the quantity or price of market-making services provided in a free-market economy from theories that assume that such services are available at no cost.

In real-life financial markets, market-makers are the parties that are always ready to deal.[5] They fill, after a fashion, the vacancy identified by Arrow. Such was the structure of the gilt-edged market. Market-makers are willing to quote prices (bids and offers) at which they will buy and sell. They provide to inquirers, free of charge, options to buy or sell up to a certain amount at the quoted prices; if a market participant wants to buy or sell more than that amount, then he or she will have to find additional bids or offers, which may be less attractive. The term 'market liquidity' refers to the ease with which large amounts of a particular asset can be bought or sold; 'ease' embraces both the amount of time it takes to complete the transaction, and how close the transaction price is to the price ruling in the market just before the transaction was undertaken.

Market liquidity depends on the amounts for which market-makers are willing to quote, the number of market-makers, and the spread between the bid and offer prices, which provides the reward which the market-makers receive for their services. The market is not in equilibrium as long as the market-makers are holding unwanted positions, but it is in a kind of near-equilibrium as long as the market-makers' positions are not too far away from what they want. The near-equilibrium is continually disturbed as new bids and offers are made, including, in the case of gilts, new issues by the government. It is also disturbed when new information emerges which affects the valuation of the asset in question: for example increases in Bank rate often led to immediate large falls in gilt prices. Of course the market-makers are exposed to risk: if they have a positive inventory of an asset whose price falls, they will lose money; likewise if they have a negative inventory of an asset whose price rises (they can acquire a negative

[4] Backhouse and Boianovsky (2013) provide an excellent account of the work. Kregel (1995) notes that the accounts of price formation developed by Walras and Marshall in the nineteenth century reflect the contemporary methods of trading employed in the Paris and London stock exchanges, respectively; the Paris exchange used a procedure akin to a periodic *tâtonnement*, whereas trading in the London exchange was continuous (as is common practice today), with temporal gaps between buying and selling orders being bridged by the intervention of professional jobbers. He concludes that the difference does not lead to theoretical diversity: 'There thus appears to be a substantial similarity between Marshall and Walras' (p. 463).

[5] Foucault, Pagano and Roell (2013) give a lucid partial-equilibrium account of the economics of market-making and market liquidity.

inventory by borrowing an asset and then selling it, leaving themselves obliged to buy the asset back and return it to the lender). The spread between bid and offer prices includes a charge for bearing these risks.

Plainly the behaviour of market-makers depends on the anticipated behaviour of other market participants. If market-makers believe that others are willing to buy and sell substantial amounts of the financial asset in question in response to small price changes, they will feel more confident in quoting prices themselves. Thus market liquidity depends not only on the market-makers themselves, but also on the community of active dealers.[6] Indeed, the distinction between market-makers and active dealers is often unclear.

It is possible to imagine a near-perfect government securities market in which the government, or any other party, can sell as many securities as it wishes, at a time of its choosing, and at a price very close to the price prevailing before the sale. Such a market has existed in the United States for many years, perhaps since the 1970s, and in the United Kingdom after Big Bang in the Stock Exchange in 1986.[7] This book, however, is concerned with the period 1928–72, when the UK government securities market was nowhere near perfect. The characteristics of the market at that time, compared with the imaginary ideal, had seriously adverse macroeconomic consequences.

There is no comprehensive body of evidence on the liquidity of the gilt-edged market in the period. No continuous records survive of the amounts for which the market-makers' bids and offers were good. As regards bid-offer price spreads, until November 1965 the *Financial Times* published two closing prices for each gilt-edged stock; these may be presumed to have been bids and offers reported at the end of the trading day.[8] The spreads as at (or near) 11 September each year (date chosen at random) from 1945–65 are shown in Figure 2.1, calculated as a percentage of the price of the stock in question. A tendency for spreads to widen is observable, except in the case of short gilts.

The evidence given to the Radcliffe Committee on the working of the monetary system, and to the Parker Tribunal on the alleged Bank rate leak of 1957, provides a lot of information on the liquidity of the gilt-edged market

[6] Hicks (1989, p. 10) talks of an 'inside market' between buyers and sellers.
[7] It has not always existed in the United States: see Garbade (2012), and Box 8.1.
[8] The words 'stock' and 'bond' are used interchangeably in this book. Gilts were normally known as 'stock' in British parlance during the period under review, except when they were in the form of bearer instruments, when they were known as 'bonds'. In American parlance, 'stock' denotes equity.

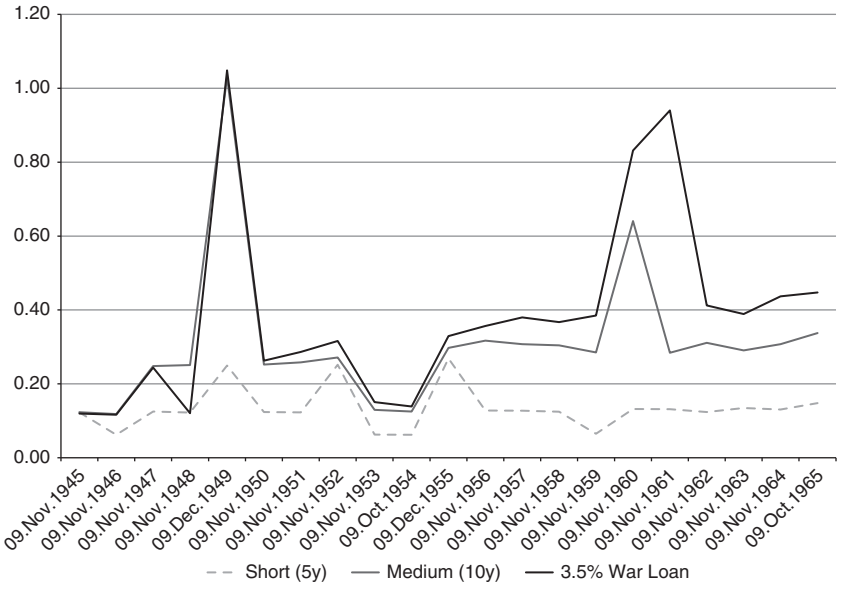

Figure 2.1 Dealing Spreads Quoted in the *Financial Times*, Around 11 September, 1945–65 (%)

in the late 1950s. The internal records of the Bank of England, and of the Government Brokers, Mullens and Co, give qualitative indications of how it developed in the 1960s. The Bank's archives contain detailed quantitative information on the Issue Department's transactions, and on the discount houses' holdings of gilts, which I have transcribed onto spreadsheets and made available on the internet.[9] The gilt prices which were published each day in the *Financial Times* and *The Times* newspapers can be found in their digital archives. And in 1964, the Stock Exchange began to collect and publish statistics of turnover in gilts. Turnover is not the same as liquidity, but it is suggestive. This book describes, among other things, how the Bank of England became the principal market-maker in gilts in the 1960s. The share of official transactions in total turnover is a revealing indicator of how far it had progressed by the mid-late 1960s, and of how far it withdrew from market-making in 1971, when the conflict with monetary policy had become intolerable.

[9] The data are available at cambridge.org, niesr.ac.uk, bankofengland.co.uk, eh.net and researchgate.net. See Appendix B for more information on sources.

Market-makers supply liquidity by quoting prices, or limit orders, at which investors can trade. Market orders – orders to deal at the best available price in the market – are executed against standing limit orders, and 'effectively decrease the available trading options, and, as such, consume liquidity.'[10] At least from the 1950s onwards, the Bank of England seems to have executed its transactions by responding to bids and offers from the jobbers, thus providing liquidity – e.g. it made tap stocks available at prices which were known in the market.[11]

The work of Benos and Wetherilt suggests a measure of liquidity provision which can be applied to the Bank of England's activities in the gilt market. If the Bank systematically sold gilts when yields fell, and bought them when yields rose, it would be supplying liquidity. In Benos and Wetherilt's language, it would for example be contributing offers of gilts to the market at times when offers were being consumed by others because demand was rising. If the Bank's purchases and sales were unrelated to yield changes, it would be a consumer of liquidity; and if the Bank were systematically to sell when yields rose and to buy when yields fell, it would be a destroyer of liquidity. The scale of its liquidity supply or destruction can be measured by the amount it bought or sold for a given yield change, and this can be estimated by regression analysis; this is done in Chapter 13.

[10] Benos and Wetherilt (2012, p. 345).

[11] Confusingly, the word 'tap' has two different meanings in the history of the gilt market. 'Tap stocks' in and just after the Second World War were gilts issued continuously at a fixed yield, directly to investors, in response to the flow of demand. 'Tap stocks' in peacetime were stocks of which the Issue Department held a large amount as a result of its underwriting activity, and which it was willing to sell in response to bids from the jobbers in the Stock Exchange. The reference here is to tap stocks in the latter sense.

3

Government Securities and the Structure of the Stock Exchange

Government securities had been dealt in on the Stock Exchange from the inception of a permanent national debt after the Glorious Revolution of 1688.[1] From 1909 until Big Bang in 1986, the London Stock Exchange insisted on 'single capacity'. Member firms had to be either brokers, who could deal with ultimate investors but could not trade with those investors from their own portfolios, or jobbers, who could trade from their own portfolios but only through brokers. Single capacity was made obligatory in 1909, and it was supplemented in 1912 by a rule requiring brokers to charge their clients according to a minimum scale of commissions. These rules were regarded as a means of ensuring that investors could have an impartial source of advice from brokers, who could not offload unwanted positions of their own onto their clients, and that prices quoted to clients were based on all the bids and offers in the market, not just those available to an individual firm.[2]

Although gilts could be and sometimes were dealt in outside the Stock Exchange, the Bank of England dealt nearly exclusively on the London Stock Exchange, and, not being a member, was obliged to deal through a broker. Its broker, the Government Broker, was the senior partner of Mullens and Co, a stockbroking firm.[3] The structure is illustrated in Figure 3.1.

The market-makers were thus the jobbers. Until the late 1960s, jobbing firms had to be partnerships, and the partners had unlimited liability. This

[1] Morgan and Thomas (1962, ch. 1).
[2] Morgan and Thomas (1962, pp. 145–7, 153–4), Kynaston (1983, pp. 252–62), Michie (1999, pp. 115–21) and Attard (2000, pp. 8–9).
[3] Mullens and Co had previously been known as Mullens, Marshall and Co, and Mullens, Marshall, Steer Lawford and Co. Wainwright (1990) provides a history of the firm and of the various individual Government Brokers.