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).



https://doi.org/10.1017/9781108605830.003 Published online by Cambridge University Press

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.

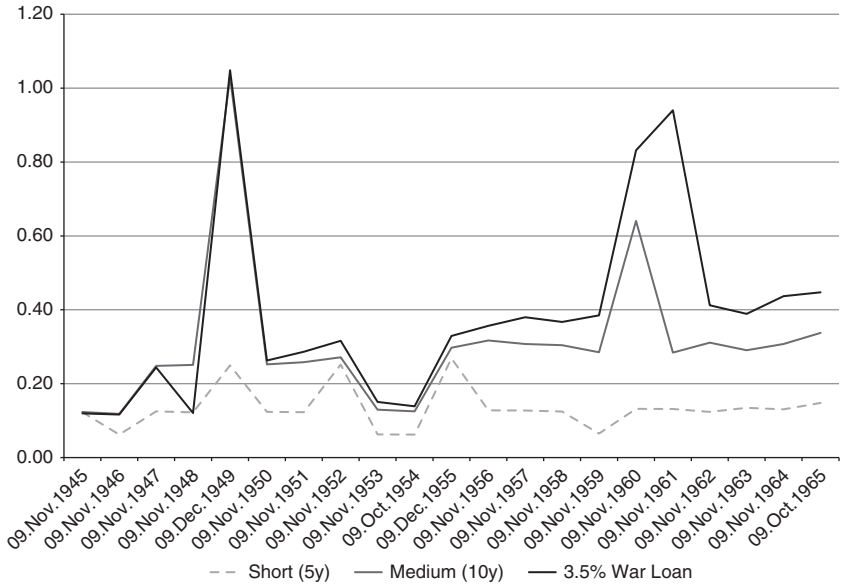


Figure 2.1 Dealing Spreads Quoted in the *Financial Times*, Around 11 September, 1945–65 (%)

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.

https://doi.org/10.1017/9781108605830.003 Published online by Cambridge University Press

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.