

4. Credit, money and central banks*

We now enter macroeconomics. It may seem bizarre to start off by dealing with the topic of credit and money. Students who receive instruction in orthodox economics are used to tackling money as an afterthought, once all the real phenomena have been taken care of. The ordering of topics proposed here is nonetheless perfectly legitimate. The principle of effective demand, which is an essential feature of post-Keynesian economics, with causality running from investment to saving, is best understood within the context of a macroeconomic explanation of a monetary production economy. Although investment is not constrained by saving, production needs to be financed, and this is why it is preferable to start with the monetary dimension of macroeconomics before proceeding to an explanation of employment and growth. Indeed, in Chapter 3, we already had to touch upon the borrowing possibilities of the firm in its attempt to expand production. Although one firm may conceivably be able to avoid borrowing, we shall see that, from a macroeconomic point of view, firms overall must get into debt for expansion to proceed. This chapter is concerned with domestic matters; money in an international context will be dealt with in Chapter 7.

4.1 BACKGROUND INFORMATION

4.1.1 Endogenous Money to the Forefront

Going over the post-Keynesian view of money and credit will give us, once more, the opportunity to take note of the internal coherence of post-Keynesian economics. Post-Keynesians of various strands hold views on monetary matters that have a common core, whether they are Sraffians, French and Italian circuitists, Kaleckians, Fundamentalist Keynesians or neo-chartalists – the branch of post-Keynesian monetary economics that has attracted considerable attention on the worldwide web by putting forth specific theoretical and policy proposals known as Modern Monetary Theory (MMT). All these economists agree that the supply of money is endogenous and demand-led. Indeed, an Italian post-Keynesian, Paolo Sylos Labini (1949, p. 240), argued, more than 60 years ago, that money has been endogenous for over 200 years. Wynne Godley (2012, p. 91) has illustrated the thesis of endogenous money supply with a rather ironic statement: ‘Governments can no more “control” stocks of either bank money or cash than a gardener can control the direction of a hosepipe by grabbing at the water jet.’ This is in clear contrast and opposition to the quantity theory of money, monetarism, and what can be found in nearly all contemporary mainstream textbooks, where the supply of money is exogenous and depicted as a vertical line in money and interest rate diagrams.

The endogeneity of the supply of money, and hence the reversed causation between money and income, has been most clearly presented in an article by Kaldor (1970b), who was responding to the claims of Milton Friedman. The endogenous money thesis was further developed in books by Kaldor (1982) and Moore (1988), where the latter opposed Horizontalists and Verticalists – those who think that the money supply is endogenous and those who believe that it is exogenous. Other important post-Keynesian contributions to the theory of endogenous money at the time included those of the French economist Jacques Le Bourva (1992), as early as 1959 and 1962, Weintraub and Davidson (1973), Eichner (1986a) and Wray (1990). Indeed, Cambridge post-Keynesians all expressed clearly their rejection of the quantity theory of money and their adhesion to the endogenous money hypothesis, as can be ascertained by the contributions of Robinson (1956, chs 22–3; 1970), Kahn (1972, chs 4 and 7), Tony Cramp (1971), Godley and Cripps (1983), and Kaldor and Trevithick (1981). John Hicks (1982) can also be put in that camp. I myself engaged with this literature when I wrote two surveys on the post-Keynesian views of money and credit (Lavoie, 1984; 1985b), spurred in part by the critiques of traditional theory by the French circuitists.

While a considerable amount of space was devoted in the past to arguments and statements supporting the claims of money supply endogeneity, this is no longer required. There are essentially two reasons for this. First, many central banks have changed the way they implement monetary policy. Their behaviour is now much more transparent, and no veil is hiding the way they are conducting monetary operations. In particular, the interest-rate targeting procedures of central banks are now explicit, corresponding to what several post-Keynesians had been claiming all along. It has thus become much more difficult to argue against the view that the money supply is endogenous and demand-led. Indeed, even researchers at well-known financial institutions are now giving their full support to post-Keynesian monetary theory (Sheard, 2013). These changes in operating procedures have been followed by changes in monetary theory within orthodox economics. While these changes are only entering a limited number of mainstream textbooks, the more advanced macroeconomic models do make room for endogenous money.

This is the case of the so-called New Consensus model, also called the New Neoclassical Synthesis, which arose in the late 1990s and early 2000s, and which combines the hyper-rational behaviour of real business cycle theory with the rigidities of New Keynesian theory (as in the DSGE models). In these models, actual interest rates are determined by a central bank reaction function, and the supply of money is assumed to adjust to the demand for money at the set interest rate. From the standpoint of endogenous money theory, as argued in the book edited by Fontana and Setterfield (2009a), there is a great deal of similarity with the monetary views that post-Keynesians have been advocating for more than 40 years, although the mechanisms underlying the rest of the New Consensus models are entirely different from those that post-Keynesians have in mind, as explained by Sebastian Dullien (2010; 2011; 2017). For instance, both post-Keynesians and New Consensus authors believe that higher real interest rates will slow down the economy. But while post-Keynesians argue that this is because high interest rates are likely to discourage investment or to have a detrimental impact on income distribution, New Consensus authors think that high interest rates ensure higher future income out of present savings, thus diminishing the need to work now to acquire assets, so that the reduction in the supply of labour leads to a reduction in current output.

In addition, while money is endogenous in New Consensus models, it plays no role, being only a money of account.

4.1.2 Historical Background

New Consensus authors usually do not attribute the rediscovery of endogenous money theory to the works of post-Keynesians, paying reverence instead to Fisher Black (1970), the co-author of the Black–Scholes formula now so much used in finance for the calculation of the prices of derivatives. Black described a world without central bank money. More knowledgeable New Consensus authors also sometimes link their monetary views to the endogenous view of money that was pursued in the early twentieth century by authors from the Austrian tradition, namely Ludwig von Mises, Friedrich Hayek and Joseph Schumpeter (Lakowski-Laguette, 2016), as well as their Swedish predecessor, Knut Wicksell (Seccareccia, 1994; Bellofiore, 2013). Indeed, Cambridge authors were also exposed to these views, since Hayek taught Kaldor when they were at the London School of Economics in the early 1930s, while Kahn translated Wicksell from German into English. But this Austrian view of endogenous money also nearly disappeared, with the advent of the exogenous money supply found in Keynes's *General Theory*, the rise of Friedman's monetarism and his advocacy of money supply growth targets, and the full acceptance of the money-multiplier story where bank reserves, controlled by the central bank, are said to allow the creation of credit and money by banks.

The vagaries of the thesis of endogenous money should not entirely surprise students of the history of economic thought, as they know that economic theory is infected by fashions and fads. Ideas that have long been abandoned suddenly reappear. As was pointed out by Cramp (1971, p.62), 'Economic ideas, the Cambridge economist D.H. Henderson once remarked, move in circles: stand in one place long enough, and you will see discarded ideas come round again . . . And nowhere is this more true than in respect to monetary theory and the associated theory of monetary policy.'

Most of the modern monetary controversies between mainstream economics (monetarism in particular) and post-Keynesian economics can be brought back to the Currency and Banking schools' debates of the early nineteenth century (Arnon, 2011). Ricardo and the Currency School argued that only coins and Bank of England notes could be considered as money, that this stock of money determined aggregate demand, and that aggregate demand determined the price level, thus giving support to the quantity theory of money. The Banking School, with John Fullarton, Thomas Tooke and John Stuart Mill, argued instead that the definition of money was much more complicated, that aggregate demand determined the stock of money, and that if controls were needed to influence prices, these controls should be imposed on credit aggregates. The Banking School also put forth the 'reflux principle', arguing that if too many banknotes were created, its holders would bring them back to the issuer, and hence the excess would disappear. The Banking School thus supported endogenous money, reversed causality and the need to focus attention on credit instead of money aggregates, just as modern post-Keynesians do (Panico, 1988; Wray, 1990).

Another well-publicized debate occurred in the late 1950s, during the hearings of the Radcliffe Commission in England. Two opposite views were then expressed, that of academics and that of central bankers. Academics defended the mainstream view,

advocating money supply control through the provision of reserves by the monetary authorities, claiming that the velocity of money and the money multiplier were constant or predictable variables, and arguing that causality ran from money to prices. Central bankers, on the other hand, supported by post-Keynesians Kaldor and Kahn, argued instead that the operating tool of central banks was interest rates, having only an indirect effect on monetary aggregates. They further claimed that the velocity of money was unstable, appealing instead to the importance of ‘general liquidity’. They added that monetary policy had only a moderate effect on inflation, because it depended on many other factors, and evoked the possible necessity of credit controls. The Radcliffe Commission dismissed the mainstream view and endorsed those of central bankers and of post-Keynesians. This choice certainly looks like the right one now, as argued by central banker Bindseil (2004b), but at the time the Commission was lambasted by mainstream Keynesians such as Samuelson (1969b, p. 7), who called it ‘one of the most sterile operations of all time’, opening the gates for the revival of the quantity theory of money and the advent of Friedman’s monetarism in the 1970s. Thereafter, as recalled by a high official at the Bank of England, ‘in separate acts of folly a quarter of a century or so ago, the monetary authorities sought to hide the fact that they were setting rates’ (Tucker, 2004, p. 369). This was the case not only in the UK, but all over the world.

Until 1970, it was not perfectly clear that post-Keynesians held a theory of money that was much different from that of the mainstream. Until then the main criticisms against the quantity theory of money were based on the instability of the velocity of money or that of the money multiplier (Minsky, 1957; Kaldor, 1964a). Only Kahn (1972, ch. 7), in his 1958 submission to the Radcliffe Commission, and Robinson (1956), in her book *Accumulation of Capital*, had a clear understanding of what was at stake. Unfortunately, Robinson’s exposition of modern post-Keynesian monetary theory is contained in chapters at the end of her book, after a complex description of growth theory, value theory and technical change, so that readers were too exhausted by then to pay any attention to her monetary theory. But as a direct response to the rise of monetarism and its widespread acceptance by central bankers and academia, the more important issue of reversed causality was simultaneously brought to the forefront by post-Keynesian authors, as already mentioned, and by a few central bankers (Holmes, 1969; Lombra and Torto, 1973; Goodhart, 1984; and now McLeay et al., 2014; Ábel et al., 2016; Bundesbank, 2017).

Over the last 30 years, there have been in my view three major contributions to our understanding of monetary economics. First, Rochon (1999) has provided an extended survey of post-Keynesian monetary economics, showing in particular the similarities and differences with the works of some New Keynesian authors, in particular the New Paradigm Keynesians (Stiglitz and Greenwald, 2003) – a comparison which is still valid today. Second, while circuitist authors emphasized the links between producers and commercial banks, and post-Keynesians underlined the links between commercial banks and the central bank, Wray (1998; 2012) has focused attention on the links between the government, the central bank and the clearing and settlement system, thus giving rise to what is now called Modern Monetary Theory (MMT). Third, Godley (1996; 1999a) has provided a framework that allows us to study together the real and the financial sides of the economy, thus greatly enhancing the comprehensiveness of post-Keynesian theory, and providing a true study of monetary production economies, giving rise to what is now called the post-Keynesian stock–flow consistent approach – the SFC approach. Besides

these post-Keynesian contributions, I should also mention the works of Bindseil (2004a; 2004b; 2014), who has produced an enlightening revision of our understanding of monetary policy implementation by central banks over the last 100 years.

4.2 THE MAIN CLAIMS

I have argued that post-Keynesian monetary economics has a common core. But of course, as is the case in mainstream economics, there is a degree of heterogeneity among the views held. In particular, for many years a debate occurred between ‘horizontalist’ and ‘structuralist’ post-Keynesians, as they were called. I think it is fair to say that most of the first exponents of money endogeneity were ‘horizontalists’: Robinson, Kahn, Le Bourva, Kaldor, Weintraub, Moore, Godley, as well as the French circuitists (Parguez and Seccareccia, 2000). However, as various authors examined in more detail the alternative monetary theory put forward by these post-Keynesian authors, some of their claims and simplifications were questioned. This internal critique came to be known, following Robert Pollin (1991), as the ‘structuralist’ view, because it was claimed by the critics that money endogeneity arose mainly from structural changes in financial conditions, and only meagrely from the accommodative behaviour of central banks (hence the name ‘accommodationists’ sometimes attributed to ‘horizontalists’). Critics included authors as diverse as Chick, Dow, Peter Howells, Le Héron, Palley and Wray, many of whom got their inspiration from the previous works of Minsky (1957) and Stephen Rousseas (1986). The disputes between the structuralists and the horizontalists generated an enormous amount of literature, followed by many attempts to produce syntheses of the two views (Rochon and Rossi, 2017).

Moore has called the controversy ‘a tempest in a teapot’ (1991, p.405) or ‘a quintessential storm in a teacup’ (2001, p. 13). I prefer to be more positive and accept Fontana’s (2003, p. 309) assessment of it, and say instead that ‘structuralists took over where the accommodationists had stopped’. The structuralists brought some clarifications and provided new details to the basic horizontalist story. For instance, they insisted that spreads between interest rates could quickly vary, in particular due to sudden changes in default assessments, the degree of uncertainty, confidence and liquidity preference. Wray (2006a, p.271) now believes that ‘for the most part . . . this particular debate was at best a result of misunderstanding’, wishing that ‘it had died a more timely death’. To a large extent, the debate has petered out, because many of the structuralist arguments regarding central bank behaviour simply cannot hold up in light of the recent developments in monetary policy implementation, notably the changes that have transformed murky central bank procedures into more transparent ones. I have given my own views on this in previous works (Lavoie, 1996d; 2006d), so those looking for further details are referred to these works and those of Rochon (1999; 2001) and Wray (2006a), although some of the debates will be evoked later in the chapter.

4.2.1 Credit and Money

Despite all this, I take the risk of devising three tables that attempt to summarize the main differences in monetary theory between the post-Keynesian school on the one

hand and mainstream economics on the other hand. As already pointed out, the concept of a demand-led endogenous supply of money is dear to post-Keynesians and makes their views distinct from the mainstream ones. However, since neo-Austrians and New Consensus economists now also recognize that money is endogenous, this cannot be the only distinctive feature of post-Keynesian monetary economics. Table 4.1 provides some additional general distinctions. Post-Keynesians, especially since the advent of the SFC approach, pay particular attention to the counterparties of the stock of money, especially the loans or credits granted by the banking system. By contrast, mainstream authors often just throw a stock of money into their models, as if it were dropped from a helicopter, as Friedman once famously said, or like manna from heaven. This explains in part why the main concern of post-Keynesians about the financial system resides in the availability of credit for productive activities and in the stock of debts held by the various agents, as debt can generate financial instability, whereas mainstream authors focus on the real balance effects of wealth or of money balances, thinking that they will stabilize the economic system.

Another key difference between post-Keynesians and mainstream authors is related to one of the presuppositions discussed in Chapter 1. The focus on production rather than exchange is a fundamental feature of heterodox economics relative to orthodox economics. This is reflected in explanations of why money is needed. In the mainstream tale, money arose as a facilitator of exchange relations, as an improvement over barter invented by economizing individuals. From this view follows the mainstream story, whereby commodity money and coins emerged spontaneously, with goldsmiths eventually realizing that they could issue promissory notes in excess of the gold entrusted to them, thus leading to fractional reserve banking. In the mainstream story, commodity money, based on silver or gold, emerged first, followed later by coins and tokens, with the final appearance of scriptural money such as bank deposits, which became accepted as a convention. Many mainstream economists wish for the return to a system where scriptural money would behave as if it were scarce commodity money. Because of the disastrous consequences of the Global Financial Crisis, some heterodox economists also seem to favour this approach, in an attempt to restrict the credit-creating ability of banks, advocating schemes similar to full-reserve banking (Laina, 2019), sometimes under the name of ‘sovereign money’ (Dyson et al., 2016), but these proposals have generated

Table 4.1 Main features, money and credit

Features	PK School	Mainstream
The supply of money is . . .	endogenous and demand-led	exogenous
Money . . .	has counterpart entries	falls from a helicopter
The main concern is with . . .	debts, credits	assets, money
Money is tied to . . .	production and social relations	private exchange
Monetary causality	Credits make deposits	Deposits allow credits
Banks are . . .	creators of credit flows	merely financial intermediaries
Reserve mechanism at work, if any	Divisor	Multiplier
Credit rationing is due to . . .	lack of confidence	asymmetric information

heated responses from post-Keynesians who emphasize the necessary elasticity of credit (Fontana and Sawyer, 2017; Nersisyan and Wray, 2017).

Post-Keynesians theorize a monetary system that has been developed by bankers for centuries, based on scriptural means of payment. Indeed, economic historians argue that scriptural credit money, providing a general unit of account and tracking credits and debts, preceded fiat money and coins (Innes, 1913; Copeland, 1974; Courbis et al., 1991). Money originated as a vehicle to settle debts. A proof of the above is that most of the so-called modern financial innovations, based on scriptural manipulations, were known since antiquity and were in practice just before and during the Renaissance. In the post-Keynesian approach, money is a social relation, with two somewhat different justifications. The first one says that credit money requires a property-based society, where pledges based on legal property – collateral – permit the expansion of loan contracts (Heinsohn and Steiger, 1983; de Soto, 2000). The second justification is based on the tax-driven approach, also called the chartalist view following Knapp (1924), which is at the heart of the MMT reconstruction of monetary theory. It says that the general acceptance of a non-metallic form of money is due to the fact that the state requires taxes to be paid in this medium (Wray, 2000). The usefulness of chartal money is derived from the state's authority to impose and collect taxes.

Keynes (1930a, pp. 4–5) himself takes the middle ground, claiming that modern money is 'a creation of the State', and 'beyond the possibility of dispute, chartalist', meaning that the state not only enforces the delivery of money contracts but also decides 'what it is that must be delivered as a lawful or customary discharge of a contract which has been concluded in terms of the money-of-account', making only a brief footnote reference to tax liabilities. Davidson (1994, p. 223) concurs, pointing out that 'the internal medium of contractual settlement is not only whatever is declared to be legal tender by the State, but also anything the State or the Central Bank undertakes to accept from the public in payment of obligations'. As to Kregel (2019), he seems to give major importance to confidence in bank money and the essential nature of the clearing house, while downplaying the role of tax liabilities. His views look similar to those of the French Institutional theories of money, which rely on trust, based on ethical confidence, convertibility and the security of payments (Alary et al., 2020).

Also from Table 4.1 we can see that, for post-Keynesians, the main causality runs from credits to deposits, meaning that bank deposits are created the moment a new bank credit is granted. There is reversed causation. This also ties in with the statement that banks are not merely financial intermediaries that would arbitrage between short-term and long-term assets, as many mainstream economists still believe. Banks, in the post-Keynesian view, are creators of credit and of purchasing power – a vision also found in Schumpeter (1934). When banks increase their assets and liabilities by granting a loan, they create new means of payment. These can be used to increase aggregate demand. This goes beyond the mainstream belief that banks merely transfer funds from patient individuals, who decide to save more and accumulate deposits, towards impatient individuals, who wish to spend more than their income and take loans to do so.

The peculiar credit-creating role of banks has been reaffirmed recently by several post-Keynesian authors (Bertocco, 2011; Bianco and Sardoni, 2018; Bouguelli, 2018; 2020; Lavoie, 2019b), thus rejecting Tobin's (1963) so-called new view, which became highly influential among orthodox authors and which claimed that banks were subjected

to the same constraints as other financial institutions, and hence had to be considered as just another financial intermediary. The post-Keynesian view of banks as credit-creating institutions that are to be clearly distinguished from other financial institutions has recently been given full support by some lawyers (Rahmatian, 2020, pp. 73–4) and a number of central bankers (McLeay et al., 2014; Jakab and Kumhoff, 2015; 2019; Ábel et al., 2016; Borio, 2019).

Furthermore, the causality associated with bank reserves is also reversed, with reserves being endogenous and demand-led, thus being a fraction (the divisor) of deposits, instead of deposits being a multiple of bank reserves (when these exist, which is no longer the case in several countries). Post-Keynesians thus reject the standard money-multiplier story, also known as the fractional-reserve banking theory of credit creation. This will be discussed in more detail in the next section. Finally, for post-Keynesians, if credit rationing occurs, it is mainly because of a lack of confidence on the part of the banking or financial system, and not so much because of asymmetric information. Indeed, in a world where so much information is now at the disposal of anybody with access to the Internet, it would be surprising to have asymmetric information as the main cause of such a recurrent phenomenon. Credit rationing, that is, credit restraints, will also be discussed later.

4.2.2 Interest Rates

Table 4.2 focuses on the characteristics of interest rates. Post-Keynesians view interest rates as a distribution variable that, to some extent, can be controlled by the monetary authorities. Thus the overall level of interest rates is not exclusively the result of market forces. The overall level of the rate of interest depends partly on policy choices about the way income will be distributed between borrowers and lenders. It is clear that there is a base rate of interest that is exogenous, in the sense that the central bank is able to set the base rate at the level of its choice. The central bank is a price-maker and a quantity-taker, at least under normal circumstances, as will be shown later. The base rate is the target interest rate of the central bank. This base interest rate is usually a short-term interest rate. Thus what can be considered to be an exogenous rate is this short-term rate. In the past it used to be the one-month or three-month yield on Treasury bills. Nowadays the target interest rate is the overnight rate – the federal funds rate in the USA, ESTER (Euro short-term rate) in Europe. In normal times, all short-term interest rates follow very closely the evolution of the overnight rate, and the latter is very close to the target rate set by the central bank. In unusual times, the evolution of short-term interest rates

Table 4.2 Main features, interest rates

Features	Post-Keynesian School	Mainstream
Interest rates . . .	are distribution variables	arise from market laws
Base rates . . .	are set by the central bank	are influenced by market forces
Liquidity preference . . .	determines the differentials relative to the base rate	determines the interest rate
The natural rate . . .	takes multiple values or does not exist	is unique, based on thrift and productivity

on private assets may diverge from that of the overnight rate and from the interest rates on government assets. Mainstream authors believe that market forces, as measured by longer-term interest rates, are the determinant of the base rate, while some structuralist post-Keynesians argue that these market forces constrain the ability of central banks to set their target of choice.

The relationship between the short-term target interest rate and long-term interest rates, especially rates on securities issued by the private sector, is much looser. Thus, while the short-term interest rate can be considered for all practical matters as being determined exogenously, this is far from the case for other rates, including bank lending rates and in particular long-term rates. Liquidity preference is reflected in the differentials between all these other interest rates, including bank lending rates, and the target interest rate set by the central bank. Liquidity preference, in contrast to the textbook model of interest rate determination, does not affect the base rate, unless we extend the concept of liquidity preference to the interest-setting behaviour of the central bank.

The last row of Table 4.2 deals with the notion of the natural rate of interest. This is the rate of interest that would exist if there were no money. It is determined by the confrontation of the productivity of capital and the time preference of economic agents. As Hubert Henderson, a friend of Keynes, put it: 'The rate of interest (like a price) is determined at the point of intersection of a demand curve and a supply curve . . . Behind the demand curve is the productivity of capital for investment; behind the supply curve is the disposition and ability to save' (Keynes, 1973, xxix, p.226). The natural rate is the key element of the New Consensus model. For orthodox authors, the long-term interest rate, deflated by expected price inflation, is the closest measure of this natural rate of interest. Thus short-term interest rates cannot deviate too much from the long-term rate, as otherwise the actual real rate of interest will diverge excessively from the natural rate of interest, thus creating disequilibria in the economy, in particular changes in inflation rates, as Wicksell would argue. Wicksellians thus contend that central banks are heavily constrained in their choice of the base rate of interest, and that the natural rate of interest, through the long-term interest rate and other market-determined interest rates, causes the base rate – an argument that is similar to the critique that Pollin (1991) was addressing to horizontalist post-Keynesians. Indeed, as Smithin (1994, p. 112) crucially points out, 'the usual argument that central banks *cannot* affect the real rate is based on the assumption that there *is* a natural rate from which the financial real rate must not deviate'.

Thus, as Mario Seccareccia (1994, p.70) points out, we cannot just assert that 'it is money-supply endogeneity which fundamentally distinguishes the neoclassical from the post-Keynesian conception of money, one would like to think that there is substantially more than the endogeneity/exogeneity issue that separates them'. What truly distinguishes (horizontalist) post-Keynesians from orthodox dissenters (for instance New Consensus authors), who also recognize the endogeneity of the money supply, is the post-Keynesian rejection of the concept of the natural rate of interest (Rogers, 1989). This is a point that I first made fleetingly (Lavoie, 1985b), and then in more detail (Lavoie, 1997). Keynes (1936, p.243) himself pointed out that he was 'no longer of the opinion that the concept of a "natural rate of interest" . . . has anything very useful or significant to contribute'. Post-Keynesians either completely reject the validity of the natural rate of interest, or consider that there exists a multiplicity of such natural rates of interest (Lavoie and

Seccareccia, 2019; Levrero, 2021). The belief or disbelief in the validity of the concept of the natural rate of interest is what allows or disallows the central bank to set the standard for all interest rates and pursue full-employment policies instead of focusing exclusively on inflation targeting. As Smithin (1996, p. 93) again puts it,

in the absence of a natural rate of interest, it can be argued that central bank control over short real rates will ultimately influence the entire structure of interest rates in the economy, including long rates . . . Eventually, the real economy must adjust to the policy-determined interest rate, rather than vice-versa. This is therefore the precise opposite to the natural rate doctrine.

4.2.3 Macroeconomic Implications

Table 4.3 summarizes the main macroeconomic implications of the two approaches to monetary theory. Post-Keynesians truly adopt the monetary analysis advocated by Schumpeter (1954, p. 277) and his student Minsky, by contrast with the real analysis that dominates orthodox theory. As already pointed out, post-Keynesians are concerned with a monetized production economy, where money is neither neutral nor an inessential veil. As a result, post-Keynesians believe that a restrictive monetary policy will have negative consequences on an economy, both in the short run and in the long run, meaning that it is likely to raise unemployment rates and possibly reduce real growth rates. Similarly, financial disturbances are likely to have both short- and long-run effects. Ironically, central bankers, imbued with the assumption of money neutrality and New Consensus dogma, used to swear that anti-inflation austerity policies could not have any negative long-term effects on the economy and that the best that central bankers could do to help the economy achieve its full potential was to make sure that inflation rates remained low and stable. Their tune changed after the 2008 financial crisis, as central bankers started arguing that the crisis had reduced the growth rate of potential output, something that was deemed impossible before the crisis. Anti-inflation policies, when real growth rates are still low and when unemployment rates are still high, are thus justified by orthodox authors on the grounds that the natural rate of unemployment is now higher than it was before the crisis.

Post-Keynesians hold two further claims of reversed causality, as shown in Table 4.3. Observing the statistical relationship between money aggregates and price inflation, post-Keynesians attribute the growth in money aggregates to the growth in output and

Table 4.3 Main features, macro implications

Features	Post-Keynesian School	Mainstream
Schumpeter's distinction	Monetary analysis	Real analysis
Financial disturbances . . .	have effects both in short and long run	have effects only in the short run
Inflation causality	The growth in money aggregates is mainly caused by the growth in output and prices	Price inflation is caused by an excess supply of money
Macro causality	Investment determines saving	Saving (loanable funds) determines investment

prices, thus objecting to the mainstream assertion that price inflation is a monetary phenomenon. This will be discussed in more detail in Chapter 8. Finally, as said at the beginning of the chapter, post-Keynesians argue that investment is not constrained by previous saving, but that instead investment determines saving – a claim that many regard as the key presupposition of post-Keynesianism. This reversed macroeconomic causality is evidently closely related to the monetized production economy, where banks can grant loans without disposing of previously acquired deposits. As Kregel (1973, pp. 159–60) put it, ‘the availability of finance in excess of earnings (and without reference to savings) allows investment to be a truly independent and autonomous variable in the system’.

It is perhaps worth dwelling on this last point. In Chapter 3, when dealing with the firm at the microeconomic level, we argued that individual firms faced financial constraints. Here, at the macroeconomic level, we argue that investment drives saving. At first sight, it would seem that a contradiction arises from these two statements. But this is not the case. Finance is not saving. The fact that households are saving or that firms are accumulating retained earnings does not mean that there will be more investment. This is a point that Keynes tried very hard to get across. We have already mentioned the paradox of thrift, which will be developed further in Chapter 5. An increase in the propensity to save will lead only to larger stocks of unsold goods in the very short run, and it will lead to a reduced level of output, with no increase in saving in the short run, when production is readjusted to aggregate demand. The fact that various individuals wish to save more does not mean in any way that firms will wish or will be able to invest more. As Keynes (1973, xiv, p. 222) put it in response to his critics and introducing his finance motive,

the public can save *ex ante* and *ex post* and *ex* anything else until they are blue in the face, without alleviating the problem in the least . . . The investment market can become congested through shortage of cash. It can never become congested through shortage of saving. This is the most fundamental of my conclusions in the field.

What ‘cash’ means here is that firms are able to invest if they hold money balances or if banks are ready to grant advances. More saving by the public will not help at all. In fact it will make matters worse.

What if prices were flexible? Would this make any difference to the paradox of thrift? Keynes (1930a, pp. 176–7) dealt with this when providing his ‘banana parable’. Keynes’s answer was essentially that if a thrift campaign got going, the reduced demand for bananas would lead to a fall in the price of bananas and to monetary losses for banana producers. Virtue would seem to be rewarded: households would accumulate their saving as deposits in bank accounts. But these deposits would then be the counterpart of the loans that banks would have to make to the banana producers so that they could cover their losses. The increased saving by banana consumers would be compensated by the financial loss of banana producers, and there would be no increase in the net wealth of the community. The saving, once again, would not help investment in the least. It is rather unlikely that loss-making producers would believe that current household saving would give rise to future consumption and invest in new facilities (Skidelsky, 1986, pp. 323–5).

4.3 UNDERSTANDING ENDOGENOUS MONEY

When students enter a post-Keynesian class in monetary economics, their minds have been so much distorted by the neoclassical fallacy of an exogenous supplied stock of money that they find it difficult to understand even the simplest story about demand-led endogenous money. The purpose of this long section is to present elementary features of a modern financial system that will help to explain the meaning of an endogenous money supply. To do so, the T-accounts of banks and central banks, where assets must by necessity balance with liabilities, will be presented in a systematic way, starting from the simplest pure credit economy, with a single bank and without central banks and outside currency. Complications will be gradually introduced, such as competing private banks, a central bank and its reserve requirements, and then, at a later stage, the state with its financial requirements and its issues of government bonds. Recent developments in banking, such as capital adequacy ratios, zero-reserve requirements and electronic money, will also be discussed within the framework of the T-accounts.

4.3.1 A Pure Credit Economy, with a Single Bank

We start with a pure credit economy, in which there is a single bank engaged in credit and debit operations, and where the unit of account has already been determined by the state. There is no other financial institution, and producing firms are forbidden to issue shares, bonds or commercial paper. This single bank may be a private bank, set up by some private entrepreneurs, with the imprimatur of the state that grants it a charter, or it may be an institution set up by the state itself. Whatever the case may be, we assume that the state has no budget, and hence that it spends nothing and taxes nothing. There is no currency money issued by the state, no reserve requirements, no financial markets and no central bank. The economy is also a closed one: there are no foreign transactions, no foreign reserves and no exchange rates. Neither gold nor silver is considered as reserve assets. Finally, in this ‘perfect’ world, selected borrowers never default, as in all DSGE models before the 2008 debacle.

Within such a financial system, all financial transactions would have to transit through this single bank. There could be no leaks, and no inflows, apart from the new credits granted by the bank. Everyone would have an account at the single bank. That account could be either a debit, in which case the bank would make a loan B to the individual household or firm, or it could be a credit, in which case the individual household or firm would be holding a bank deposit, D , that is, money. Because there are no leaks, whatever amount of credit exists must be equal to the amount of debit; that is, $B = D$.

In a sense, today we are nearly in the situation described above. Many countries have no compulsory bank reserves, for instance there are no reserve requirements in Canada since 1994 and the US Federal Reserve has entirely removed its reserve requirements in March 2020 (Ihrig and Wolla, 2020, p.25); few transactions are made with cash, most of them being done through electronic transfers, credit cards and debit cards. There are several banks, but all transactions are eventually netted in a single institution – the clearing house. Furthermore, in some countries, all banks ‘move forward in step’ as Keynes (1930a, p. 26) put it, so that ‘there is no limit to the amount of bank-money which the banks can safely create’.

Let us now assume that some agents, most likely firms, but perhaps also household consumers, wish to increase the amounts they borrow. Where are these loans going to come from? In the mainstream story, starting from fully loaned banks, new credits can be granted only when banks are the recipients of new deposits, a situation that occurs when the central bank purchases government bonds (from banks or from the general public) on the open market, thus giving rise to the creation of excess reserves. But in this pure credit economy there are neither government bonds nor a central bank. How can the banks create new loans?

The post-Keynesian answer to this query is rather simple. Loans are created *ex nihilo*, at the stroke of a pen, or by punching a key on the computer, as long as the borrower is creditworthy, that is, as long as the borrower can show some collateral. The main limit to this process is given by the amounts of loans that can be granted to creditworthy borrowers. This depends on the willingness of borrowers to borrow, on the amount of collateral they can show, and on the willingness of banks to grant creditworthy status to their customers. The last factor may be influenced by the liquidity preference of the bank, to be defined in a later section. In a sense, loans are not truly created *ex nihilo*, since they generally require collateral.

The simplest arrangement for banks and their customers is to set up credit lines, or overdrafts, which Keynes (1930a, p. 41) defined as ‘an arrangement with the bank that an account may be in debit at any time up to an amount not exceeding an agreed figure, interest being paid not on the agreed maximum debit, but on the actual average debit’. When the credit line is pulled on, the additional loan awarded to the borrower has an immediate counterpart in the liabilities of the bank, by the creation of an equivalent additional deposit. Thus, in Table 4.4, both loans and deposits would simultaneously increase by the same amount. The additional deposits would then change hands, as the borrower would presumably use the newly created deposit to pay for some goods or services.

In our modern world, just as in Keynes’s time, it is not necessary to have money deposits to be able to spend. ‘A customer of a bank may draw a cheque against his deposits, thus diminishing his *credit* with the bank; but he may, equally well, draw a cheque against his overdraft, thus increasing his *debit* with the bank’ (Keynes, 1930a, p. 41). The same point was made by Robinson (1956, p. 19) when she said that ‘overdraft facilities are just as good a source of purchasing power as a bank balance’. Similarly, recipients of cheques may use the funds to increase their credit balances at the bank, that is, their deposits; or they may use the funds to reduce their debit balances, that is, the used portion of their credit line.

This is what currently occurs with the use of debit cards. For instance, when customers use their debit cards to purchase goods, they may hold no bank deposits (no money), and hence the used portion of their credit line will be increased as the transaction occurs. Similarly, some sellers may also be in a debit position *vis-à-vis* the bank, and hence the payments received will be automatically applied to reduce the amounts due, as shown in Table 4.5. In these circumstances, no change whatsoever will occur in the amount of

Table 4.4 Simple balance sheet of the unique bank

Assets	Liabilities
Loans <i>B</i>	Deposits <i>D</i>

Table 4.5 *Payments through debit balances*

Bank assets	Bank liabilities
Debit position of purchaser +100	
Debit position of seller -100	

Table 4.6 *Taking explicit account of credit lines*

Bank assets	Bank liabilities
Loans (used overdraft facilities)	Deposits
Unused overdraft facilities (potential loans)	Unused overdraft facilities (potential deposits)

deposits held by each transactor. This is a clear example of the reflux principle emphasized by the Banking School and Kaldor (Kaldor and Trevithick, 1981).

Agents who desire to spend can thus do so in two ways. Either they spend by depleting their money balances (the bank deposits), or they maintain their money balances while increasing their debit at the bank. As Keynes pointed out, the potential for purchasing goods and services, or what he calls the cash facilities, is made up of two components, the money deposits held and the unused overdraft, that is, the portion of the credit line that has not yet been drawn on. ‘Properly speaking, unused overdraft facilities – since they represent a liability of the bank – ought . . . to appear on both sides of the account’ (Keynes, 1930a, p. 42). If we were to keep track of these unused overdraft facilities, as some regulators suggest, then bank accounting would look as shown in Table 4.6.

Although unused overdraft facilities are still considered as off-balance-sheet loan commitments, some countries, namely the USA, have some statistical data on these (Moore, 1988, p. 25). Unused credit lines at the time represented approximately half of the narrowly defined money stock, and twice the amount of used overdraft facilities. In the UK, ‘about 60 per cent of overdraft facilities are in use at any one time’ (Howells, 2010, p. 171). It is thus clear that those who have credit-line arrangements have access to an endogenous source of credit-money. Indeed, the Bank for International Settlements (BIS) considers that formal standby facilities and credit lines of banks must be taken into consideration when assessing minimum liquidity requirements (BIS, 2013a, p. 32).

4.3.2 A Pure Credit Economy, with Privately Issued Banknotes

Let us now assume that the customers of the unique bank would like to benefit from the convenience offered by banknotes when making their transactions. In other words, rather than having all transactions going through a scriptural system of accounts, some transactions, presumably the smaller ones, could be made without the bank being an intermediate. In our modern financial systems, we are accustomed to banknotes being backed by the central bank or by the Treasury of the central government. Here, neither of these institutions exists. Where will the banknotes come from? They will be issued by the private bank.

Table 4.7 Balance sheet of the unique bank, with banknotes

Assets	Liabilities
Loans B	Deposits $D' = D - H$ Banknotes H

In the past, privately issued banknotes were quite common. The state would grant some banks the right to issue them. Indeed, some economists believe that banks, and not the government, ought to issue banknotes, a position known as free banking. In our pure credit economy, banknotes would be issued by our unique bank, on demand. In other words, whenever a customer desires to have a deposit transformed into banknotes, these would be created by the bank. Banknotes are purely endogenous. Their supply is demand-led. The new T-account of the bank would look as in Table 4.7.

Banknotes issued by the unique bank are a liability of the bank. The bank transforms one kind of liability, the deposits, into another kind of liability, the banknotes. There is no limit to the amount of banknotes that can so be created. If customers of the bank were to bring back the banknotes, they could be exchanged only for deposits at the very same bank. There cannot be an excessive creation of banknotes, in line with the reflux principle underlined by Kaldor, Robinson (1956, p. 227) and Le Bourva (1992). It could happen, however, that credit is granted to finance inflationary expenditures or wage increases. This explains why the best-known exponent of the reflux principle – Thomas Tooke, who was in favour of free enterprise in all aspects of economic life – argued in his later work that while banknotes need not be regulated, credit and loans ought to be (Arnon, 1993).

It should be noted that banks should be highly favourable to the issue of private banknotes. Whereas banknotes carry no interest rates, bank deposits do or can easily do so. As a consequence, for a given mark-up between the loan and the deposit interest rates, the higher the proportion of money in the form of banknotes, the larger the profits of banks. As a result, one should expect banks to favour any technological change that would transform bank deposits into privately issued banknotes. This is the case of smart cards, electronic wallets or electronic purses.

4.3.3 The Profits and Bad Loans of Banks

One issue that has been ignored up to now is that of bank profits. Obviously, in a perfect world where financial transactions could be conducted at no cost, and where there would be no default risk on the part of borrowers, the condition of zero profit for the bank would be equivalent to a situation where the interest rate charged is equal to the rate of interest paid on deposits. In the real world, borrowers are sometimes unable to face their debt commitments and must default on their loans. Banks must thus set a spread between the lending and the deposit rates, to compensate for this risk. In addition, banks, like all firms, must pay salaries to their employees, they must service the cost of their fixed equipment, and they must turn a profit for their owners. Indeed, banks, like all firms, have a certain target rate of return, and the differential between the lending and the deposit rate (on top of service charges, now a key component of bank profits) will be such that banks

achieve this target rate of return on the capital of their owners in normal times, net of the losses due to loan defaults.

If all the profits are redistributed to the households who own the banks, then the simple equality of Table 4.4 will still be relevant. This can be seen with the help of Table 4.8. Assume that the rate of interest on loans is i_B , while the rate on interest on deposits, which is lower, is i_D . By the end of a year, unless interest payments have been made by the borrowers or loans have been reimbursed, outstanding loans will now amount to $B(1 + i_B)$. On the liability side, outstanding deposits will have increased to the amount $D(1 + i_D)$. This implies, assuming that in the initial state deposits and loans were of equal amounts, that the profits of the banks are now of an amount $D(i_B - i_D)$. These profits, as shown in the first line of Table 4.8, will now be added to the bank's own funds – to their equity capital. If all the profits are redistributed as dividends, which are now held by the bank owners as bank deposits, then we are back to the second line of Table 4.8, where deposits and loans are once more equal, as they were in Table 4.4.

It should also be clear, even though loans and deposits appear to grow at a rate equal to the rate of interest on loans, i_B , that there is no need for loans and money deposits to grow at that rate. For instance, if the households who are the owners of the banks decide not to save their dividends in the form of deposits, but rather decide to spend them all on consumption goods, the loans due by the non-financial firms will be reduced by an equivalent amount, and hence there will be no growth whatsoever in the amount of outstanding loans. Thus, although post-Keynesians keep saying that 'loans make deposits', implying that granting a new loan will lead to the creation of new money deposits, readers need also be aware that if households decide to increase their spending, the money balances of the purchasers will flow into the money balances of the sellers. This will allow the producers selling their wares to use their newly acquired bank deposits to reduce the size of their debt *vis-à-vis* the bank, and thus reduce the amount of outstanding loans, if the producers prefer to reduce their money balances. Thus there is a possible two-way link between bank loans and bank deposits. To sum up, the operating profits of the bank are added to the own funds of the bank, while the reimbursement of a bank loan is accompanied by a reduction in bank deposits.

In general one would expect the administrators of the bank to retain part of the profits, so as to increase their own funds and constitute a buffer against expected losses. Retained profits are added to the funds initially put up by the owners of the bank when starting business. Broadly speaking, it is the bank's net worth, although things in real life are a bit more complicated (see Fullwiler, 2013, p.174). The own capital of the bank constitutes a liability to itself. It represents the funds that the firm owes to its owners. In general, the bank's own funds play a role similar to deposits in the hands of the owners. The own funds, just like the deposits or the credits, are an accounting entry, but in contrast

Table 4.8 *Banks with own funds distributed as dividends*

Assets	Liabilities
Loans $B(1 + i_B)$	Deposits $D(1 + i_D)$
	Own Funds (equity capital) $OF = D(i_B - i_D)$
Loans $B(1 + i_B)$	Deposits $D(1 + i_B)$

Table 4.9 Banks with own funds and loan defaults

Assets	Liabilities
Loans B	Deposits D Own funds OF
Loans $B' = B - BLWO$	Deposits D Own funds $OF' = OF - BLWO$

to deposits, they cannot be drawn down by the owners. They are reduced whenever the bank pays out dividends or whenever a borrower defaults on a loan. In the latter case, an identical amount is deducted from the loan assets and from the own funds liabilities when the bad loans need to be written off (that is, when the accountants of the bank consider that the borrowers are unable to service the interest payments on their loan and will never be able to pay back the loan).

This is shown in Table 4.9, under the assumption that an amount $BLWO$ of bad loans is being written off. When there are too many bad loans, the amount of own funds, that is, the net worth of the bank, can become negative. This occurs when the amount of defaulting loans exceeds the amount of own funds of the bank, thus reducing the value of assets below that of liabilities, in which case the bank becomes insolvent. This, however, has no impact on the day-to-day operations of the bank, unless someone takes notice of the insolvency. As Neilson (2019, p. 73) puts it, ‘one can quite easily be liquid but not solvent.... As long as lenders are willing to put up cash, [borrowers] are able to make their payments’. In other words, while insolvent banks should legally go bankrupt, they are forced to do so only if the insolvency or the danger of insolvency makes the bank illiquid, that is, unable to cover deposit withdrawals and unable to settle in the payment system. During the Great Depression of the 1930s, concerns about US banks being insolvent generated hundreds of bank runs (where clients withdraw their deposits and where lenders recall their loans and refuse to do business). In Europe and in Canada many banks were no doubt insolvent, but no one cried ‘fire’ and so there was no stampede to the exit – no bank run – and banks eventually recovered when the economic situation improved.

In the case of a unique bank, with a total monopoly on financial assets and liabilities, it is difficult to conceive why insolvency should create an illiquidity problem. The unique bank *is* the payment system, and clients can withdraw only those banknotes issued by the bank itself. Agents have nowhere else to put their savings. The situation of central banks is quite analogous to that of our unique bank, and hence fears of possible default by an insolvent central bank seem unwarranted in general. There are exceptions to this claim, however, for instance if a central bank has borrowed funds in a foreign currency, or if a country were to leave the eurozone while its central bank held a large debit position in the TARGET2 system.

4.3.4 Bank Liquidity, Bank Solvency and the Capital Adequacy Ratio

Another interesting feature of the pure credit economy, with no government bonds on the books of commercial banks, which is indeed representative of several financial systems in the world, is that the liquidity of a bank is rather difficult to assess from the standard

point of view. In the standard view, the liquidity of a bank is measured by the ratio of its safe to total assets, or else by the ratio of its liquid to illiquid assets. This standard view of liquidity has already been criticized by those who point out that, through liability management, it is (nearly) always possible for large banks to obtain the funds that they need to settle their accounts (Moore, 1988, p. 33).

In the pure credit economy, there are no safe assets since banks hold no government bonds and since there are neither reserves nor central bank cash. The lending behaviour or the liquidity of a bank thus cannot be based in general on the proportion of cash or government bonds that the bank holds in its asset portfolio. Such a ratio would always be equal to zero. The only option left, then, is to measure the liquidity of a bank by the proportion of its own funds. It is the own funds of a bank, rather than its reserves or safe assets, that may play a key role in a theory of endogenous money (de Boyer, 1998). The relevant ratio to measure risk would be the *B/OF* ratio, that is, the ratio of loans to own funds. Alternatively, if the real assets of the banks are taken into account, it would be the asset to own funds ratio.

Lately, this ratio, or rather its inverse, the own funds to assets ratio, has become the subject of intensive scrutiny, under the guidance of the BIS. The BIS has designed a 'capital adequacy ratio' that private banks ought to respect, under the guidance of central banks or their supervising agencies. The minimum ratio suggested by the BIS used to be 8 per cent, but it is now quite complex. Roughly speaking, it is the ratio of the own funds of the banks (their capital) to a weighted measure of their assets, the weights being based on a conventional assessment of the risks associated with each kind of assets, and even off-balance items (BIS, 2013b).

Some economists have argued that the maximum asset to own funds ratios that arise from the imposition of minimal capital adequacy ratios may replace the role of reserve multipliers in a world without reserve requirements (Dow, 1996, p. 499; Nell, 2003, p. 196; Descamps and Soichot, 2003). For Borio and Disyatat (2010, p. 77), 'the main exogenous constraint on the expansion of credit is minimum capital requirements'. This new multiplier would be equal to the allowed assets to own funds ratio. For instance, it has been argued that the incapacity of the Japanese economy to move out of economic stagnation since the 1990s has been due to the low net worth of the Japanese banks, following the huge losses that these banks had to absorb as a consequence of substantial defaults on loans (mainly related to land and construction speculation). Because of their low net worth, it was said that Japanese banks were prevented from granting new loans because of binding capital adequacy ratios. A similar argument has been made with regard to European and American banks since the 2008 financial crisis.

A few objections can be made in this respect. It should be noted first that the capital adequacy ratios have been set in such a manner that they would only be binding for the most risky banks. Second, solvent and profitable banks accumulate retained earnings that are added to their net worth, and they should have no trouble in inducing economic agents to either forego deposits or take loans to purchase newly issued bank shares, thus allowing these banks to improve their own funds to assets ratio. Third, as I was told by a central banker, the BIS rules gave flexibility in how to evaluate the riskiness of assets, making it possible for accountants to devise means to transform the capital adequacy ratios of banks into good-looking ratios. Fourth, a bank can grant loans to investors for the purpose of purchasing newly-issued shares of the bank (Werner, 2016;

Rahmatian, 2020, p.80). The reluctance of banks to accept strict capital adequacy ratios may have more to do with their desire to achieve high rates of return on equity than with their fear of being unable to grant new loans. Indeed, Panagopoulos and Spiliotis (2017, p.139) conclude from their empirical analysis, that ‘the equity multiplier is operational and reversed (loans cause equity)’.

What happens, however, if all banks incur large losses, pushing the asset to own funds ratio beyond its maximum value, but without being insolvent, so that no private agent would be willing to buy bank equity? In such a situation, there is only one way out. The government or the central bank has the responsibility to purchase new equities issued by banks, thus proceeding to partial or full nationalization. If such action were taken by the central bank, the banks would increase both their own funds and their reserves at the central bank. This would allow them, as we shall see in later sections, either to reduce their borrowing from the central bank or to acquire safe assets such as government securities. They could then resume their business of granting loans to all creditworthy customers.

Some post-Keynesian authors argue that when banks grant new loans they are automatically reducing their liquidity and hence their liquidity preference, since the ratio of loans to own funds is immediately rising (Wray, 1995; Fontana, 2009, p.103). This would be one possible cause of an upward-sloping supply curve of credit. That it is true, in some sense, can be seen immediately from the first line of Table 4.9. At the very moment that a new loan is granted, the bank commits itself to a more illiquid position. The amount of loans rises while that of own funds remains the same. Thus, at the very moment when a new loan has been granted, the bank is in a more risky position.

This situation is, however, only a temporary one, for the larger stock of loans and deposits will allow the bank to rake up additional net interest revenues (unless the new loans are being defaulted on in unusual proportions), as shown in Table 4.8. These additional revenues, when they are due and integrated with the retained earnings, will thus bring the *B/OF* ratio back to its initial level. At the end of the year, the balance sheet of the bank has increased in size, but the liquidity of the bank may remain the same. In other words, at the macroeconomic level, there is no upward pressure on interest rates charged on loans when additional loans are granted.

4.3.5 A Pure Credit Economy, with Two Sets of Banks

What happens when there are two banks or more? New institutions will need to be created: a clearing house and an interbank market. Let us still assume that there is no central bank and no government expenditures. Let us further assume that there are two banks, one specialized in making loans to corporations, and the other specialized in collecting deposits from households. In a sense, this corresponds to the institutional framework of many financial systems. In the USA, for instance, banks located in New York have specialized in making large loans to big business, while the so-called country banks specialize in collecting deposits. Similarly in France, for a long time specialization was institutionalized, with business banks (*banques d'affaires*) and deposit banks (*banques de dépôt*) (Marchal and Poulon, 1987). Godley and Cripps (1983, p.77) call them loan banks and deposit banks.

With this division of the banking business, it is impossible for each bank to reach an approximate equality between loans and deposits. The deposit bank (Bank D)

Table 4.10 Two banks in a pure credit economy

Bank B		Bank D	
Assets	Liabilities	Assets	Liabilities
Loans to non-financial agents	Deposits	Loans to non-financial agents	Deposits
	Funds owed to Bank D	Advances made to Bank B	
	Own funds		Own funds

consistently has excess deposits, while the business bank (Bank B) continuously has an excess of loans over deposits. In other words, at the end of each day, Bank B realizes that the payments made by its customers in favour of the customers of Bank D are of an amount that exceeds the cheques drawn the other way. The positive balances of Bank D at the clearing house are exactly matched by the negative balances experienced by Bank B. At the end of the day, to settle the payments that have cleared, Bank B must obtain a loan from Bank D. Thus, over time, Bank B is indebted *vis-à-vis* Bank D, and Bank D holds an asset against Bank B, as shown in Table 4.10. In a system with only two banks, it cannot be otherwise. Still, if Bank D consents to grant loans to Bank B, the accounts at the clearing house will balance at the end of each day, and such a situation can perpetuate itself. As Godley and Cripps (1983, p. 77) say, 'once a system of inter-bank credits is in existence, there is no logical or institutional constraint on the extent to which the whole banking system can supply additional loans, thereby simultaneously expanding the stock of money held by the public'.

The two banks B and D, or the two sets of banks, need only make sure that they agree on a rate of interest that will be profitable to both of them. In other words, the interbank interest rate, that is, the rate of interest charged by Bank D to Bank B on the amounts due, must be somewhere in between the (low) rate of interest on deposits that Bank D is paying to its depositors, and the (high) rate of interest on loans that Bank B is charging to its borrowers. If the interbank rate is set in an appropriate fashion, the rate of return of both banks will be the same, and hence the ratio of own funds to assets of both banks will be the same. Similarly, the loans to own funds ratio will also be the same in both banks, where loans now include the loans made to other banks. Both banks will also engage in longer-term arrangements, discussed next.

4.3.6 Certificates of Deposits and Securitization

Although the above direct lending arrangements between banks are perfectly legitimate, some observers of the banking scene may find it rather worrying that some banks are heavily in debt in respect of other banks or other financial institutions. Overnight lending is only for a day, and needs to be renewed daily. It is thus safer to look for longer-term arrangements. Besides the obvious solution of attracting more deposits, and especially more time deposits, a series of arrangements has been designed, two of which we shall briefly outline.

An option for business banks, that is, the banks that specialize in lending to firms, is to issue certificates of deposit (CDs). Bank B, the business bank, is attracting an insufficient

Table 4.11 Two banks, certificates of deposit

Bank B		Bank D	
Assets	Liabilities	Assets	Liabilities
Loans to non-financial agents	Deposits	Loans to non-financial agents	Deposits
	Sold CDs	Purchased CDs	
	Own funds		Own funds

amount of deposits. Bank B may thus issue certificates of deposit, which will be purchased by Bank D, which has positive balances. The issued certificates of deposit would thus appear on the liabilities side of the balance sheet of Bank B, replacing the amounts due to Bank D, while the purchased certificates of deposit would appear on the assets side of the balance sheet of Bank D, taking the place of the loans made to Bank B. This is shown in Table 4.11, which is barely different from Table 4.10. But now Bank B no longer ‘borrows’ from Bank D, or so it appears from a legal point of view, and in any case the CD secures funds for more than one night.

The above presentation makes clear that credit relations are based on creditworthiness. As long as Bank D believes that Bank B is able to provide interest payments on its commitments, there is no reason for Bank D to refuse extending loans to Bank B, or purchasing certificates of deposit from Bank B. The same occurs at the level of the customers of a bank. Banks grant loans and renew lines of credit as long as they have faith in the ability of the borrower to make interest payments. Similarly, depositors have no hesitation to leave their deposits at a bank as long as they believe that their orders to transfer these deposits will be honoured. The creditworthiness of a bank thus ultimately depends on the creditworthiness of its borrowers and the confidence of its depositors. Creditworthiness and trust are the key elements of the financial system.

It should be emphasized that there are no safe assets in this pure credit economy. When deposit banks lend to business banks, collateral in the form of risk-free government assets cannot be provided since they do not exist by definition. The creditworthiness of the loans granted by the banks is the only collateral available. Conventions based on trust and confidence rule the banking system.

The sale of certificates of deposit used to be the main tool of liability management by banks. Broadly speaking, liability management refers to the ability of banks to increase their lending activity by borrowing funds that appear on the liability side of their balance sheet, without being forced to sell some of their marketable assets – mainly Treasury bills. This gives rise to another fashionable arrangement, ‘securitization’. It should be made clear that securitization has two different meanings. In the case of mortgages, securitization in its old meaning implied that banks would keep the mortgage on its books on the asset side, but that the bank, instead of looking for deposits, would issue securities based on these loans, with the mortgage-based securities now appearing on its liability side. Securitization in this first sense is thus part of liability management. To finance their loans, instead of looking for new deposits or issuing certificates of deposit (CDs) that are of a relatively short duration, banks issue long-term bonds that have the additional advantage that they help them cover term risk. This kind of arrangement was first put

in place in nineteenth-century Germany, where the securities are known as *Pfandbrief* bonds, and it was used in the USA as a means to help mortgage issuers, the thrift industry and home owners, in the hope of lowering mortgage interest rates. Only mortgage loans conforming to strict norms and insured by some government agency were eligible for such conversions, so there was very little risk for the investors. In and of itself, this first kind of securitization is of little consequence and has been practised successfully for decades – and still is.

Securitization is now mainly understood in its second meaning, which is related to asset management. Securitization may be defined as the transformation of an asset that was not previously marketable into a marketable one. In other words, the securitization of an asset implies that this asset can now be sold on some market. Securitized assets go beyond mortgages and include student loans, car loans and even credit card balances. A typical example of securitization is the sale of a set of loans, previously granted by a commercial bank or a mortgage bank, to a special purpose entity, or special investment vehicle. The loans are then transformed into securities (the asset-based securities (ABS) and the collateralized debt obligations (CDO)) with the help of investment banks – in the case of the USA these were the Wall Street banks (Morgan Stanley, Bear Stearns, Lehman Brothers and Goldman Sachs) – which also act as the underwriter. Part of these long-term assets are ultimately financed by short-term assets – asset-based commercial paper (ABCP). The buyers are other financial institutions, non-bank financial intermediaries such as money market funds, mutual funds, pension funds, trust companies and insurance companies, which collect vast amounts of savings from households and from non-financial firms with large money balances. This process, ‘originate and distribute’, as is well known now, contributed to the 2006 housing crash in the USA and to the 2008 debacle of the Wall Street banks. Minsky (1991) considered securitization as another example of endogenous money, whereby financial innovation helps to create portfolio assets that appear to take the characteristics of money as a safe and liquid store of wealth.

Once again, this kind of arrangement arises as a result of specialization. Other bank or non-bank financial intermediaries specialize in collecting long-term savings from households and corporations. The lender banks specialize in finding borrowers that seem to be creditworthy. This specialization creates an imbalance in the balance sheets of banks and financial intermediaries, similar to the one shown in the top part of Table 4.12. This imbalance could be solved by banks issuing certificates of deposit, which would be bought by financial intermediaries, but securitization quickly became the new fad in finance, allowing financial wizards to show their skills.

In the case of securitization, banks typically sell a portion of their loans to a financial intermediary that has collected a large amount of time deposits or other liabilities of a similar nature. The loans thus disappear from the balance sheet of the bank, and appear on that of the financial intermediary. In the example in Table 4.12, 70 per cent of the loans end up being securitized. The net result for the bank is a reduction in the size of its balance sheet: the loan made to a non-financial institution is gone, but so is its liability to financial intermediaries. The bank is forgoing the future interest revenues to be obtained from the borrower, but it has collected up-front fees when initially granting the loan and selling it. In addition, the bank is now in a better position with regard to its capital adequacy ratios, as previously discussed. It has collected fees when making the loan,

Table 4.12 *Securitization*

Bank B		Financial intermediary FI	
Assets	Liabilities	Assets	Liabilities
Loans to non-financial agents +100	Deposits +30		Deposits +70
	Funds owed to FI +70	Advances made to Bank B +70	
Loans to non-financial agents +30	Deposits +30	Securitized loans (basic paper) +70	Asset-based commercial paper (derivative paper) +70

and it can repeat the operation without having to worry about the BIS-imposed capital adequacy ratios, thus circumventing them.

Securitization, just like lending between banks, requires confidence. Bunches of loans can be repackaged and sold only as long as the purchasers of these loans are confident that the loans will be repaid and the interest payments will be made.

4.3.7 Multiple Banks, the Clearing House, the Central Bank and Repos

Things are just slightly more complicated in a multi-bank system. Payment orders will clear at a clearing house, usually run by the association of domestic bankers, or else the central bank will act as the clearing house. In the first case, the clearing system is likely to be a deferred net settlement (DNS) system, where payments are cleared on a net multilateral basis, but where final settlement occurs only at the very end of the day; in the second case, it is likely to be a system based on real-time gross settlement (RTGS), where participants need to hold sufficient balances at the central bank for each payment, with settlement occurring at the same time as the payment is made, and with clearing balances being moved from one bank account to another at the central bank. Obviously, the DNS system requires no or few reserves at the central bank, while the RTGS system necessitates much larger reserves (or else large daytime overdrafts from the central bank, as in the US system). There are also hybrid systems, as in Canada's large-value transfer system (LVTS), where payment is final the moment it clears while netting occurs on a multilateral basis. It is this system that we describe below, run with a private clearing house, and with the central bank playing a role only at the very end of the day.

A single bank may at the same time owe funds to a bank and be owed funds by another bank. The clearing house is designed to net out these balances, and bring together all the main participants to the clearing system. By the end of the trading day, each participant bank knows its clearing-house balance, that is, the amounts that it can lend to those in deficit or the amounts that it must borrow from those in surplus. The clearing house may then act as a broker between the deficit and the surplus banks. An example with four banks and their gross flows of payment is provided in Table 4.13. What should be noted is that, as long as commercial banks issue their own banknotes and as long as the transactions involve only commercial banks, and not the central bank, whenever a bank

Table 4.13 *The clearing house in a multi-bank system*

Owed to → Owed by ↓	Bank A	Bank B	Bank C	Bank D	Σ amounts owed by (debits)	Σ amounts owed to (credits)	Clearing balances
Bank A		15	20	20	55	60	+5
Bank B	30		40	35	105	90	-15
Bank C	20	50		10	80	85	+5
Bank D	10	25	25		60	65	+5
Σ amounts owed to	60	90	85	65	300	300	0

has a clearing-house deficit balance, there is another bank, or a group of banks, that has an identical surplus balance. In other words, the net overall balance of all banks taken together under these conditions is zero at all times. In the case illustrated by Table 4.13, Bank B (the business bank) would need to borrow 15 monetary units from the other three banks on the interbank market by the end of the day, in order to be able to settle its payments.

If one of the other three banks (say Bank A) declines to lend its surplus funds, then for final settlement to occur, Bank B will need to borrow funds from the central bank, while Bank A will have no choice but to deposit its surplus funds at the central bank. These are the so-called liquidity facilities of the central bank: the borrowing and the deposit facilities. The central bank acts as the lender of last resort, making a loan to ensure that payments are indeed settled.

In reality, during the course of the day, banks will estimate their position at the clearing house, and, instead of borrowing on the overnight market (for one night), may decide to engage in repurchase agreements, also called repos or RPs. For instance, when Bank B sees that it will end in a negative position at the clearing house, it will offer to sell some of the securities that it has on its balance sheet, but promise to buy them back – say the next day, in a week or in two weeks – at a pre-agreed price. From the point of view of the organization granting the liquidity – the party that buys the security and promises to sell it back – this agreement is called a reverse repo. This is tantamount to a loan, secured on collateral. Indeed, this is how the BIS (2013a, p.41) views repos. A one-week repurchase agreement is equivalent to a one-week collateralized loan. But with a repo, the security is held and owned for the duration of the repo by the lender, not by the borrower. Since capital adequacy ratios attach no risk to this kind of asset, loans based on repos can totally circumvent the capital adequacy requirements and have thus become highly popular. In the example of Table 4.13, if Bank B manages to sell 15 units worth of securities to the other banks, it will manage to bring its clearing-house balance back to zero (Rochon and Rossi, 2004).

In reality, repos are often based on government securities, mostly Treasury bills, but we have not yet introduced the government and the central bank, so here we assume that they are based on securities issued by private agents! While the system of repos seems to be highly useful, we shall deal later in this chapter with its possible drawbacks.

4.3.8 The Central Bank in an Overdraft Economy

An interesting classification, underlined by several post-Keynesian authors, is the distinction between ‘overdraft economies’ and ‘auto-economies’, a distinction first made by John Hicks (1974, p. 51). In the auto-economy, agents sell their liquid assets to finance new ventures, or they issue new bonds or shares. For this reason, these economies are often called ‘financial-markets economies’, but we shall call them ‘asset-based economies’ to underline the fact that firms in such economies are said to own the financial resources required to make their investment expenditures, while banks are said to sell their liquid assets (mainly Treasury bills) to obtain central bank money.

In the overdraft economy, by contrast, firms or households pull on their lines of credit with commercial banks when they require new financing means. The same distinction applies to the financial sector. In an overdraft economy, commercial banks need to take advances from the central bank. As a result, when the focus of the analysis is on the balance sheet of the central bank, the distinction between an overdraft economy and an asset-based economy depends on whether the central bank has claims over the domestic financial sector. ‘The overdraft economy is thus defined by a double level of indebtedness: that of the firms to the banks and that of the banks to the central bank’ (Renversez, 1996, p. 475). This distinction will be quite useful in the discussions that follow.

Let us start, then, with the simpler of the two systems, the overdraft system. The overdraft system is an extension of the pure credit economy, to which a central bank is added. In the overdraft system, the operations of central government, beyond those of the central bank, may still be assumed away, and this is mainly why the overdraft system is easier to describe than the asset-based system.

Let us then assume the existence of a central bank, and that of a network of commercial banks, consolidated into a single conglomerate for simplification. In this more realistic financial system, we assume that private banks can no longer issue banknotes. Only the central bank can issue banknotes. Suppose that we start from the situation described by Table 4.4. The private bank conglomerate has loans on the asset side, and deposits on the liability side (own funds are set aside for simplification). Its depositors now wish to split their money holdings into bank deposits and banknotes. How will the banks provide their customers with the banknotes issued by the central bank?

The mainstream answer, provided within the framework of an asset-based financial system, is that the banks must sell to the central bank some of the government securities they hold, thus obtaining the central bank banknotes they need. These banknotes, which are said to be part of the money supply, are also part of what is called high-powered money, that is, money issued by the central bank, or central bank money. But here, in this pure overdraft economy, there is no government sector (beyond the central bank) and there are no government securities lying around. And we assume, in line with present institutions (in normal times), that central banks just will not buy any privately issued asset. Still, banks are required to obtain the banknotes, for their customers will lose all faith in the banking system if it cannot provide the banknotes they want. How can the banks obtain the banknotes?

If banks cannot sell any asset to the central bank, the only way they can obtain the banknotes is by borrowing them from the central bank. This is what used to be known

Table 4.14 *Overdraft economy, with banknotes*

Commercial banks		Central bank	
Assets	Liabilities	Assets	Liabilities
Loans B	Deposits $D' = D - H$ Funds H borrowed from central bank	Advances H made to commercial banks	Banknotes H

as a ‘discount window operation’, whereby banks borrow from the central bank, now known as a ‘borrowing standing facility’. This is shown in Table 4.14. The amount borrowed is exactly equal to the required amount of banknotes, that is, the amount of central bank money. The central bank has a monopoly over the provision of banknotes. As long as there is a demand for central bank banknotes by their customers, commercial banks are forced to go into debt *vis-à-vis* the central bank. Commercial banks cannot but be indebted *vis-à-vis* the central bank.

4.3.9 Liability Management and the Overdraft Economy

The overdraft economy is the ultimate example of liability management. Several authors have argued that liability management is the latest stage in the historical development of banking systems, where money is seen as an evolutionary process (Chick, 1986). Before the advent of liability management, banks would passively wait for deposits, and only expand their lending activity if new depositors emerged. The attracted deposits were taken as a pool of funds, available for lending.

There is, however, another view of liability management – the radical view, akin to what Rochon and Rossi (2013) call the revolutionary view of endogenous money. According to this new view, liability management is not an innovation that would have transformed the process of banking intermediation. Rather, liability management is a permanent feature. Banks are perpetually engaged in passive liability management, as they must first consent to loans, and later search for funds to finance the currency drain (deposits that become banknotes). All overdraft systems are compelled to practise liability management as a logical necessity. Any adjustment is made on the liability side, simply because no adjustment from the asset side is possible. Banks as a whole, when they are in need of banknotes for their customers, or in need of compulsory reserves, as we shall soon see, cannot get them by selling liquid assets to the central bank, either because banks just do not hold Treasury bills or because the central bank declines to purchase government securities, as was the case in many emerging countries and with the European Central Bank until May 2010. Banks must obtain all their high-powered money by borrowing it from the central bank. The argument, to be found in the traditional view, that liability management would be a new phase in the development of financial systems thus does not seem to be a correct assessment of the actual evolution of financial systems throughout the world.

An interesting feature of the overdraft economy is that it shows clearly that money and high-powered money are endogenous variables that cannot be under the control of the

Table 4.15 *Overdraft economy, with compulsory reserves*

Commercial banks		Central bank	
Assets	Liabilities	Assets	Liabilities
Loans B	Deposits D		
Reserves H	Funds H borrowed from central bank	Advances H made to commercial banks	Deposits of banks (reserves) H

central bank. In the present overdraft economy, with a demand for central bank banknotes, the banknotes must be provided and the central bank has no choice but to provide the commercial banks with the loans they ask for (unless it wishes to create chaos in the economy, as ATMs get empty). The central bank is left, however, with a powerful tool: setting the rate of interest at which the commercial banks will be forced to borrow the required amounts of banknotes.

The situation is identical when compulsory reserves are taken into consideration. Suppose that we are in an economy where customers wish to use only scriptural money and no banknotes. Is it still possible for the central bank to force indebtedness on the part of private banks? The obvious solution is for the central bank to impose compulsory reserve requirements. It does not matter whether the reserves are imposed on deposits, as they are in most countries, or on loans or other assets, as they used to be in France and in other European countries and as is now suggested by Palley (2006). Reserve requirements have consequences similar to those of central bank banknotes on the accounting structure of banks. Again, because banks in an overdraft system have no assets to sell to the central bank, banks have no choice but to borrow the required reserves, at the rate of interest charged by the central bank. As shown in Table 4.15, the adjustment to the compulsory reserve requirements is made through the liability side.

4.3.10 The Overdraft Economy, with Two Banks or Sets of Banks

Finally, let us consider the case of two banks, or two sets of banks, within an overdraft economy. Let us sweep away, again for simplification, all the complications associated with compulsory reserves and banknotes issued by the central bank, as well as the own funds of banks. In a previous section, we also considered the case of two sets of banks, but in the absence of a central bank. All discrepancies in the net claims of each bank had to be made good by banks borrowing or lending funds to each other, usually with the help of some clearing agent, the clearing house. Let us now consider the case where the clearing house is the central bank, with no interbank market.

The advantage of such a system is that the commercial banks need not enter into contracts with each other. In other words, the risk of lending to another bank is now taken over by a public institution, the central bank. Suppose again that there are two kinds of bank, the business bank and the deposit bank. The business bank will consistently run negative balances at the clearing house, while the deposit bank will consistently accumulate surpluses. The clearing house is now the central bank, and what the central bank can do, in contrast to the private clearing house, is itself be the counterpart to the required

lending and borrowing operations of the banks when accounts have to be settled at the end of the day.

This is shown in Table 4.16. The discrepancy, for Bank B, between loans and deposits is exactly equal to the discrepancy, for Bank D, between deposits and loans. And this discrepancy is exactly balanced on the books of the central bank. Provided there is only a small difference between the penalty rate charged on negative settlement balances (the advances provided by the central bank to banks showing a deficiency of funds) and the rate of interest offered by the central bank on the positive settlement balances (the rate of interest on the surplus funds that banks with excess funds deposit at the central bank), there is no incentive for banks to look for the private arrangements described in Table 4.10, such as overnight lending. In other words, provided the central bank is content with making a small profit when running the clearing house, there is no need for commercial banks to settle their accounts between themselves before relying on the facilities of the central bank for final settlement.

An interesting feature of this overdraft system without an overnight market and with a central bank is that it clearly shows that the amount of high-powered money (here excluding banknotes and including only the amount of deposits held at the central bank – the reserves) has no relationship whatsoever with the total amount of money, or money deposits, in the system. This amount of high-powered money depends mainly on how extensive are the specializations of the private banks into loan-making and deposit-attracting activities. When all banks move together in step in their lending and deposit businesses, the required amount of high-powered money is quite low, and may even approach zero. When banks specialize heavily, the amount of high-powered money will be large relative to economic activity.

The amount of outstanding reserves, relative to economic activity, will also rise when the overnight market breaks down, because participants lose confidence in each other. This was precisely the case of the eurozone from 2010 to 2012. Investors moved their deposits from banks situated in the south towards banks situated in the north of the eurozone. Because of the lack of confidence in the solvency of banks situated in the south (Greece, Spain, Portugal, Italy), the overnight market was not functioning properly, so that banks in the north (mainly Germany) preferred to hold their positive settlement

Table 4.16 *Overdraft system, with two banks and no overnight market*

Bank B		Bank D		Central bank	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
Loans to non-financial agents +100	Deposits of non-financial agents +80	Loans to non-financial agents +70	Deposits of non-financial agents +90		
	Advances from the central bank (negative settlement balance) +20	Deposits at central bank (positive settlement balance) +20		Advances to Bank B +20	Deposits of Bank D +20

balances as deposits at the European Central Bank (ECB) rather than lending them on the overnight market. Something similar happened in the USA in September 2008, as banks became reluctant to lend overnight on the uncollateralized fed funds market, leading to an increase in deposits at the Fed (by banks with positive settlement balances) and in advances made by the Fed (to banks with negative settlement balances).

4.3.11 The Central Bank in an Asset-based Economy: The Neo-chartalist View

Let us now abandon the overdraft economy for a while and deal with the asset-based financial system. In an asset-based economy, there are large stocks of accumulated public debt. In the past, central governments have run public deficits, and as a result there is an outstanding stock of government securities, which is held by the central bank and by private agents, non-financial and financial ones, banks in particular.

New debt is issued when past issues have come to maturity and the central government is unable to reimburse the debt-holders: this is the case of the rollover. New debt is also issued when the government runs a deficit. There are two views with regard to the financing of government deficits. According to the first view, the Treasury, the fiscal arm of government, is best seen as drawing cheques from its account at the central bank, and selling securities to the central bank in order to replenish its bank account at the central bank. This view is endorsed in particular by the post-Keynesian advocates of Modern Monetary Theory – also known as the neo-chartalists (Wray, 1998; 2012; Mosler and Forstater, 1999; Bell, 2000). According to the second view, it is best to imagine that the Treasury sells securities to commercial banks (or more precisely to primary dealers). This second view – let us call it the post-chartalist view for taxonomic purposes – sees government expenditures in a light akin to that of expenditures by private firms: firms need to borrow from banks to make their expenditures. The bonds issued by government and purchased by banks play a role similar to that of the advances made by banks to firms (Lavoie, 2003d).

Let us start with the web-popular neo-chartalist view. Let us suppose that the government is running a deficit, say equal to 100 units, and hence issues Treasury bills (short-run government securities) to that amount, the bills being bought by the central bank, in line with the neo-chartalist view. Table 4.17 illustrates the neo-chartalist story. The first row of Table 4.17 shows the impact of the initial sale of Treasury bills on the accounts of the central bank. This sale has no impact whatsoever on the money supply and the private economy as long as the newly acquired government deposits are not spent in the

Table 4.17 *The neo-chartalist story: Treasury bills are first sold to the central bank*

Central bank		Commercial banks	
Assets	Liabilities	Assets	Liabilities
Treasury bills +100	Government deposits +100		
Treasury bills +100	Deposits of banks +100	Reserves +100	Household deposits +100
Treasury bills +19	Banknotes +10 Deposits of banks +9	Treasury bills +81 Reserves + 9	Household deposits +90

economy. Once the expenditures of government have actually occurred, the deposits of government at the central bank are now held by households or firms as deposits at private banks. But since the cheques were drawn on the central bank, the banks, the customers of which received the cheques, see their clearing balances at the clearing house increase by 100 units. If nothing else happens, by the end of the day, these balances will be deposited at the central bank, where they will enlarge bank reserves at the central bank by 100 units, as shown in the second row of Table 4.17.

Banks now dispose of excess reserves of 100 units. In the mainstream story, these 100 units of excess reserves would allow the commercial banks to provide new loans and start the money-multiplier process. The post-Keynesian story at this stage is entirely different, however. In the post-Keynesian view, banks provide loans first, and search for reserves later. Banks do not wait for excess reserves to be provided like manna from heaven. They grant loans whenever a creditworthy customer shows up or if they find one. It follows that, when banks wind up with excess reserves, they have already granted all the loans they could have made. The fact that they now have excess reserves, or positive settlement balances, does not make their potential customers more creditworthy.

What, then, will banks do with their excess reserves – their extra deposits at the central bank? Banks that hold these additional deposits will try to lend them to banks that have an insufficient amount of reserves or that are in a deficit position within the clearing house: they will try to lend them on the overnight market. But, as we have seen in a previous subsection, the overall net position of banks in the clearing house cannot but be zero as long as there is no transaction involving the government or the central bank. This means that, when the government spends through its account at the central bank, the overall net position of banks in the clearing house is necessarily positive. Taken as a whole, banks have no reason to borrow reserves. As a consequence, the overnight rate (the federal funds rate in the USA) must fall (Mosler, 1995). To stop it from falling, and to keep the overnight rate at its target level, the central bank must engage in a counter-vailing action. There are several possibilities, including the textbook one, which is to pursue open-market operations, with the central bank selling government securities. Open-market operations occur on secondary markets; that is, they deal with second-hand securities, which have already been issued and sold.

If there are no reserve requirements and if there is no demand for additional central bank banknotes by the bank customers, banks will use the entire reserve amount of 100 units to purchase Treasury bills. In the present case, let us assume that households now wish to hold 10 extra units of banknotes, and keep only 90 units in the form of bank deposits. Let us also assume that there is compulsory reserve ratio of 10 per cent on deposits. Banks will thus need to use 10 units of their reserves to acquire the banknotes and will need to keep nine units to fulfil their reserve requirements. The bank can thus trade the remaining 81 units of reserves by purchasing 81 units of Treasury bills from the central bank, as shown in the last row of Table 4.17.

4.3.12 The Central Bank in an Asset-based Economy: The Post-chartalist View

Let us suppose again that the government is running a deficit, say equal to 100 units, and hence issues Treasury bills to that amount, the bills being bought by a commercial bank, in line with what we called the post-chartalist view (we omit the complication brought

Table 4.18 *The post-chartalist story: Treasury bills are first sold to the commercial banks*

Central bank		Commercial banks	
Assets	Liabilities	Assets	Liabilities
		Treasury bills +100	Government deposits +100
		Treasury bills +100	Household deposits +100
Treasury bills +19	Banknotes +10 Deposits of banks +9	Treasury bills +81 Reserves +9	Household deposits +90

about by the fact that the bills are more realistically purchased by dealers, who are often affiliated with a bank, and who require bank advances). The counterpart to this purchase is the deposits that are now credited to government. This is shown in the first row of Table 4.18. In the second stage, however, the deposits will revert to households as soon as the planned government expenditures are made good, as shown in the second row. We can assume again that that households wish to keep 90 per cent of their additional money holdings in the form of bank deposits, and 10 per cent in the form of banknotes issued by the central bank.

The question arises again as to how the commercial bank will be able to obtain their required reserves and the banknotes demanded by its customers. In this asset-based financial system, banks need simply to sell to the central bank some of the Treasury bills that are part of their assets. In the present case, since households desire 10 units of additional banknotes and with banks being submitted to a 10 per cent reserve ratio on the remaining 90 units of deposits, banks will be selling 19 units worth of Treasury bills to the central bank on secondary markets, as shown in the third row of Table 4.18. Is the final situation any different from the one described by the neo-chartalists? The answer is no; ultimately it is not. The last row of Table 4.18 is identical to the last row of Table 4.17. Whether the Treasury bills used to finance government expenditures are initially sold to the central bank or to commercial banks makes no difference whatsoever to the final requirements of the banking system. Thus, the fact that most central banks cannot purchase government securities on the primary market, only being allowed to do so on secondary markets, seems to be a useless institutional constraint, as MMT advocates would claim.

Which view best describes the financial relationship between government and the banking system: the neo-chartalist view of Table 4.17 or the post-chartalist view of Table 4.18? It really does not matter. Each view may best correspond to the existing institutional arrangements. In Canada, the central bank makes direct purchases of up to 20 per cent of the securities newly issued by the federal government. The neo-chartalist view would thus seem to apply to that country. By contrast, the post-chartalist view seems more appropriate for the eurozone and the USA. In the eurozone, until the sovereign debt crisis that started in 2010, the European Central Bank declined to purchase any government security, either directly from governments (it is prohibited from doing so) or indirectly from financial institutions. Eurozone governments must sell their securities to commercial banks, and these banks must then take advances from the European Central Bank to acquire reserves or banknotes, as described in Tables 4.14 and 4.15. In the USA, there are also restrictions

on direct purchases of government securities by the Federal Reserve – the central bank in the USA (Akhtar, 1997, p. 37).

In reality, things are even more complicated, and correspond neither to what we called the neo-chartalist story nor to the post-chartalist story. First, the newly issued securities are purchased by dealers, which are not necessarily banks. Second, most of the funds so acquired by the government are normally repatriated on the books of the central bank, so that the financial sector must borrow funds from the central bank to settle its position at the clearing house. The resulting balance sheet is thus some intermediate step between the neo-chartalist and the post-chartalist stories (Lavoie, 2013a, p. 15).

MMT authors have attracted a considerable amount of attention, but also of puzzlement, by making statements such as ‘the Treasury does not “need” to borrow in order to deficit-spend’ (Wray, 1998, p. 117), ‘taxes do not finance spending’ (Forstater and Mosler, 2005, p. 538) or ‘the federal government is not dependent on revenue from taxes or borrowing to finance its spending’ (Kelton, 2020, p. 9). These statements rely on the neo-chartalist description of Table 4.17. But they implicitly omit the first transaction of this table, which is that the Treasury must first borrow funds, or rather sell bonds or bills to the central bank. Neo-chartalists rationalize this omission and their surprising statements by resorting to the consolidation of the government’s financial activities with the central bank’s operations. By doing so, the activities of the central bank are turned over to the Treasury. It is said that the sale of government securities to the central bank has no impact on the economy as long as the government does not spend the proceeds of this sale – certainly a correct claim – so that consolidation is justified. A number of post-Keynesian economists counter this by arguing that consolidation and its associated description of how governments finance their expenditures are counterfactual and likely to create confusion (Armstrong, 2020; Fiebiger, 2016a; Lavoie, 2013a).

4.3.13 The Government in an Overdraft Economy

MMT authors argue that consolidating the government and the central bank and making use of the neo-chartalist process shown in Table 4.17 is justified by the fact that ‘even after adding the self-imposed constraints and going through the minute details of Fed-Treasury operations, we find that the basic claims made in the much simplified model hold’ (Wray, 2012, p. 108). The case of overdraft economies, where central banks are prohibited by law or by convention to make any purchases of government securities on both the primary and secondary markets, illustrates the importance of taking into account operational and institutional reality. In overdraft economies, an institutional configuration encouraged by the IMF, the central bank acts as a lender of last resort to the banks, but not as a purchaser of last resort for the government securities. For instance, the constitutional act of the central bank of Chile says that ‘no public expenditure or credit of whatsoever nature may be financed with loans granted, directly or indirectly, by the Bank’ (Banco Central de Chile, 2016, p. 17). Several central banks from emerging economies face legal constraints or even bans on bond purchases. If MMT is truly ‘a *description* of how a modern fiat currency works’ (Kelton, 2020, p. 233), then it must also consider this third configuration.

The self-imposed constraints of such an overdraft economy will have repercussions on the ability of the government to pursue expansionary policies and handling its debt,

despite the fact that such a framework seems to fit the requirements of what Wray and Nersisyan (2020, p.262) define as a *sovereign currency*, where the country, as is Chile, is a *currency issuer* rather than a *currency user*: the State chooses the money of account; the State imposes taxes that must be paid in the money of account; the State issues bonds in its own currency; the central bank issues the currency that must be accepted for payment; the country is on a flexible exchange rate. MMT authors argue that such a sovereign state ‘can always meet its obligations denominated in its own currency’ and ‘can set the price/interest rate on any obligation it issues’ (ibid., p.263). But we can see, with the help of Table 4.19, that this will not necessarily be the case.

Again, we assume that the government is running a deficit of 100 units, financing it through the issue of securities, which in this overdraft economy can only be bought by the commercial banking sector. We retain the assumption that the central bank is the bank of the government, so that the proceeds of the bond sale are repatriated at the central bank. This puts the banks in a negative position at the clearing house, forcing them to take advances from the central bank, or else to enter into a repo agreement with the central bank, using the Treasury bills as collateral for the repos. This corresponds to the second row of Table 4.19. Thus, in contrast to the neo-chartalist and post-chartalist stories, government deficit-spending will tend to raise overnight interest rates, unless the central bank proceeds to liquidity-providing operations. The government then uses its deposits at the central bank to pay households, who get bank deposits, transforming 10 per cent of their new balances into banknotes, while the banks must keep a 10 per cent reserve to deposit ratio, as we assumed in the previous two tables. This leaves us with the third row of Table 4.19, where in contrast to the last row of Tables 4.17 and 4.18, the central bank holds no government securities.

Within this configuration, as exemplified by Chile, there is no workaround the self-imposed constraint. While it is true that the central bank keeps some control over short-term rates of interest, through the rate that it sets on its loans to the banks, within this configuration the central has little or no control over long-term rates if financial markets are determined to speculate against the sovereign bonds. Providing advances or repos – of short or long duration – to the financial markets is not the same thing as purchasing bills and bonds outright on the secondary markets. Interest rates will rise if banks or bond dealers are reluctant to purchase the securities when they are newly issued and when investors refuse to rollover the securities that come to maturity. Within this overdraft configuration, the central bank can only *encourage* financial markets to keep

Table 4.19 *An anti-chartalist story: Treasury bills are never sold to the central bank*

Central bank		Commercial banks	
Assets	Liabilities	Assets	Liabilities
		Treasury bills +100	Government deposits +100
Advance (repos) +100	Government deposits +100	Treasury bills +100	Advance (repos) +100
Advances (repos) +19	Banknotes +10 Deposits of banks +9	Treasury bills +100 Reserves +9	Household deposits +90 Advances (repos) +19

or purchase the securities since it does not act as a purchaser of last resort (Oreiro and Costa Santos, 2020).

The arch-typical example was the eurozone crisis, between 2010 and 2012. Financial markets took into consideration the fact that, *by convention*, the European Central Bank (ECB) only provided repos based on rated collateral and did not make outright purchases of sovereign debt on secondary markets. When it became clear with the Greek case that the ECB was willing to abandon the convention and use its constitutional right to engage in outright open-market operations, the financial markets then realized that the ECB was itself reluctant to do so in full force, in part because Bundesbank officials still insisted that the convention ought to be upheld, and so the crisis extended to other countries and kept dragging on. The financial crisis only came to a halt when the new President of the ECB, Mario Draghi, insisted that the central bank from then on would purchase a ‘whatever it takes’ amount of sovereign bonds. MMT advocates argue that the crisis occurred because financial markets took note that countries within the eurozone did not have their own currency and were *currency users*. I would argue instead that the crisis occurred because financial investors took note of the constraints imposed by the conventions tied to the overdraft configuration of the eurozone system.

4.3.14 Reverse Operations and Transfers of Government Deposits

In previous sections we saw that commercial banks as a whole have a net zero position at the clearing house as long as there are no transactions involving the central bank. Such transactions will occur if commercial banks need to acquire reserves or banknotes issued by the central bank. They will also occur whenever payments are moving in and out of the government account at the central bank; that is, whenever the government is paying for its expenditures by drawing on its deposits at the central bank or when tax payments are made by the private sector and the payment is deposited at the government’s account at the central bank. In a fixed exchange rate regime, as will be discussed in Chapter 7, the interventions of the central bank on foreign exchange markets also generate transactions between the central bank and the commercial banks. A payment involving the central bank and a commercial bank will also occur when a government security matures and the security has to be redeemed. Whenever such transactions occur, they affect the overall net clearing position of banks at the clearing house, and hence the amount of bank reserves at the disposal of the banking system. These various transactions and their effects on reserves and on the overnight interest rate are shown in Table 4.20.

As Peter Howells (2010, p. 169) says, ‘in the course of a normal working day, there will be large spontaneous flows between the public and the private sectors’. These payments can involve large sums of money and thus can have a sizeable impact on the amount of outstanding reserves. If the central bank did not intervene, these payments would generate huge daily fluctuations in the overnight rate of interest. As Bindseil (2004a, p. 17), an ECB and formerly Bundesbank central banker, puts it, ‘if the central bank does not systematically neutralize exogenous money market shocks, short-term rates would regularly be either zero or be extremely high’. Thus, in contrast to what is believed by mainstream authors, the main role of central banks is essentially a defensive one, as we shall see in more detail in the next subsection. The central bank must assess the size and the

Table 4.20 Impact of various transactions on reserves and overnight interest rates, with possible compensatory measures

Action	Reserves (settlement balances)	Overnight interest rates	Possible compensating measures
Citizens and corporations pay their taxes, which are deposited at the account of the government at the central bank	Fall (banks make a payment outflow involving the central bank)	Tend to rise	Central bank buys government securities or other financial assets Central bank provides advances to banks
Central bank sells banknotes to banks			Central bank provides repos to banks or dealers (liquidity-creating reverse operations)
Government sells securities to the private sector and deposits the proceeds in its account at the central bank			Central bank moves government deposits at the central bank towards government accounts at banks
Central bank sells foreign currency on foreign exchange markets			
Banks wish to acquire more reserves			
Government purchases goods and services out of its account at the central bank	Rise (banks receive a payment inflow involving the central bank)	Tend to fall	Central bank sells government or central bank securities Central bank provides reverse repos to banks or dealers (liquidity-absorbing reverse operations)
Government securities mature and central bank redeems them out of government account at the central bank			Central bank moves government deposits at banks towards government account at central bank
Central bank buys foreign currency on foreign exchange markets			

amplitude of the payments that are likely to go through its accounts during each day, and be ready to take compensating measures.

In some countries, such as Canada, the central bank is in a position to know with perfect certainty the size of the net payments that go through its accounts. As a result, and because there are no reserve requirements in Canada, in normal circumstances the central bank is able to bring back to zero its position at the clearing house, so that the net position of the private banking system is also brought back to zero at the end of each day. In that case, the central bank can achieve nearly perfectly its target overnight rate. In other countries, such as the USA or the UK, things are more complicated, as the central banks do not have perfect information about the drains on reserves that must be compensated for. They also lack information about the demand for free reserves and the extent of discount window borrowing. As a consequence, these central banks cannot perfectly equate the supply of and the demand for reserves at the target interest rate, so that the actual overnight rate can wander away from its target.

As indicated in the last column of Table 4.20, central banks can take various compensatory measures to absorb fluctuations in reserves caused by daily transactions between the central bank and the banking system. Mainstream textbooks emphasize the open-market operations mentioned in the previous sections. Open-market operations are usually understood as outright purchases or sales of government securities by the central bank. As we saw earlier, however, a large proportion of open-market operations are conducted on different terms, on the basis of reverse operations. Repurchase agreements add flexibility and security to the conduct of monetary policy. With repurchase agreements, central banks can add to or subtract liquidities from the financial system, but on a temporary basis, say for one week. So, if at the end of the week central bankers desire to call back the previously injected liquidities, they need do nothing; the adjustment will be automatic. 'Under the RP agreement, the Desk buys securities from dealers who agree to repurchase them at a specified price on a specified date. The added reserves are extinguished automatically when the RPs mature' (Akhtar, 1997, p. 37). Liquidity-absorbing repos do the opposite: they allow the central bank to reduce the liquidities that would be temporarily in excess amount, and to reverse this operation automatically at the end of the purchase and resale agreement.

Repos operations should be understood for what they are: both the central bank and the banks are perfectly willing to engage in these special kinds of open-market operations. In the case of liquidity-absorbing repos, for instance, the central bank wants to drain excess reserves in order to maintain the overnight rate at the level of its choice. The banks, on the other hand, want to get rid of excess funds that would draw either no interest rate (as was the case in the USA until 2008) or a rate of interest below market rates (as in Canada) if it were to be deposited at the central bank. Repos operations have replaced outright open-market operations. We shall discuss repos further when tackling the shadow banking system.

There are, however, other means, besides open-market operations, repos or direct advances, at the disposal of the central bank to add liquidities to or subtract them from financial markets. Some central banks, such as the Bank of England, have issued their own bills, which they sell to banks or dealers. Several central banks – in particular the Bank of Canada – now use transfers of government deposits between the books of the central bank and those of commercial banks as the main tool to control the amount of settlement balances in the system. To increase liquidities in the banking system, government deposits are transferred from the central bank to private banks; to subtract liquidities, government deposits are transferred the other way. These government deposit transfers are often used to compensate the flows originating from government expenditures and collected taxes. We have already seen, with the help of Table 4.17, that government expenditures financed by cheques drawn on the central bank automatically lead to the creation of reserves. Reciprocally, taxes collected from private agents and deposited as government deposits in the accounts of the central bank withdraw reserves from the banking system. An obvious way to neutralize these effects is to transfer government deposits the other way. For instance, if, near the deadline for income tax collection, payments made by households to the order of the tax-collecting government agency have been settled at the clearing house, the liquidities of the banks may be replenished by the central bank, simply by moving government deposits back to the accounts at the commercial banks.

Table 4.21 Transfers of government deposits

Central bank		Commercial banks	
Assets	Liabilities	Assets	Liabilities
	Government deposits +100 Deposits of banks -100	Reserves -100	Household deposits -100
	Government deposits 0 Deposits of banks 0	Reserves 0	Household deposits -100 Government deposits +100

This is shown with the help of Table 4.21. Suppose households have paid their taxes with cheques worth 100 units. Commercial banks then lose the equivalent of 100 units of reserves or clearing balances, as shown in the first row of Table 4.21. But the loss can be made good by the autonomous transfer of government deposits from the books of the central bank to those of the private banks, as shown in the second row of Table 4.21. Reserves will then come back to their zero level.

4.3.15 The Defensive Role of Central Banks

It is now time to summarize what we have learned so far. Causality is probably the most crucial aspect of economics: this is how, in many instances, theories can be distinguished from one another. The T-accounts that have been presented must be comprehended with the causal story appended. This causal story underlines the importance of the autonomy of credit creation from the previous existence of bank deposits. The coherence of the described financial system also relies on the law of reflux, the existence of credit lines and the acquisition of own funds.

Our study of the banking system has shown that lending and borrowing between banks is an inevitable feature of financial systems. The banking system is a system that relies entirely on trust. Banks must make loans and purchase financial assets to make profits. Nothing limits the creation of credit and money other than the prudence of the bankers and the self-restraint of borrowers. When confidence falters, financial institutions will decline to lend to each other, and financial markets, notably the interbank market, will freeze. In times of crisis, there is thus a need for a special operator that can intervene as a lender of last resort – the central bank.

In normal times, the role of central banks in their day-to-day operations is also a defensive one. The central bank must neutralize, or sterilize, the net inflows or outflows of the clearing and settlement system. Otherwise there would be wild gyrations in the amount of outstanding reserves and in the overnight interest rate, as was pointed out more than 100 years ago by Walter Bagehot (1873).

Day-to-day monetary policy implementation means maintaining some level of short term interest rates by permanently adjusting quantities with regard to high frequency, partially transitory shocks, which are not or only marginally related to macroeconomic developments . . . Any attempt to control in the short run the monetary base leads to extreme volatility of interest rates since the market will, due to stochastic and seasonal fluctuations in the demand for base money, permanently be short or long of reserves, as already observed by Bagehot (1873). (Bindseil, 2004b, p. 18)

Why have central bankers focused on interest-rate targeting instead of money targeting throughout history, except for the brief stint with monetarism? Two justifications have been proposed. The most famous one is that of Poole (1970). It is a macroeconomic justification. It says that if the LM curve is more unstable than the IS curve, meaning that the demand for money is volatile, interest-rate targeting will produce less output variability than monetary targeting. The other explanation is the microeconomic one proposed in the quotation above; it is associated with the post-Keynesian view. It is tied to the inner functioning of the clearing and settlement system, and linked to the day-to-day operations of the central bank, as explained in the previous section, and as accounted for by central bank economists (Bindseil, 2004a at the ECB; Clinton, 1991 at the Bank of Canada; Holmes, 1969 at the Fed). It is this microeconomic explanation that is the relevant one. As Fullwiler (2013, p. 178) puts it, ‘when analysing monetary policy, the better starting place is the modern central bank’s role in the national payment system’.

Thanks to the works of MMT scholars – Mosler (1997–98), Wray (1998) and Fullwiler (2003; 2006) – post-Keynesians now have a much better understanding of the defensive operations conducted by central banks and the technicalities involved in the interaction between the Treasury and the central bank when taxes are collected and when public expenditures arise. It needs to be said, however, that the first post-Keynesian to provide an in-depth understanding of the defensive role of central banks was Alfred Eichner.

Through his empirical work, Eichner discovered that there was no correlation whatsoever between changes in the holdings of government securities by the Federal Reserve and the reserves of the commercial banks.

No matter what additional variables were included in the estimated equation, or how the equation was specified (e.g., first differences, growth rates, etc.), it proved impossible to obtain an R^2 greater than zero when regressing the change in the commercial banking system’s nonborrowed reserves against the change in the Federal Reserve System’s holdings of government securities. (Eichner, 1986a, p. 111)

This lack of correlation led Eichner to look for an alternative explanation for the changes in the balance sheets of central banks – an explanation that he found in the work of Lombra and Torto (1973), whose article covered the defensive operations of the Fed, in addition to the endogenous money and reverse causation arguments put forward by post-Keynesian authors. Their explanation was fully endorsed by Eichner (1987, p. 849), who later claimed that ‘the Fed purchases and sales of government securities are intended primarily to offset the flows in and out of the domestic monetary–financial system and thereby hold bank reserves constant’, just as we have argued in this section.

As pointed out earlier, the implementation procedures of central banks have changed and are now much more transparent, reflecting the actual behaviour of central bankers. Those who have tracked these changes are fully aware that central banks do not attempt to exert quantity constraints on bank reserves, as is still claimed by some post-Keynesian structuralists. Still, this does not mean that central banks fully accommodate the private sector. Accommodation, as already explained by Eichner (1987, p. 847) and Rochon (1999, p. 164), should take on a meaning unrelated to quantitative adjustments. It should refer to whether or not the central bank is pegging the target interest rate, whatever the

economic conditions. When central banks are not accommodating, they are pursuing what Chick (1977, p.89) calls 'dynamic' operations, increasing or reducing the target interest rate. This is a topic for the next section.

4.4 NEW DEVELOPMENTS IN MONETARY POLICY IMPLEMENTATION

4.4.1 Interest-rate Targeting

As a consequence of recognizing the total failure of monetarism from the mid-1970s to the early 1980s, central bankers have thought hard about how monetary policy strategy and monetary policy implementation ought to be changed. Perhaps the best-known change is the explicit or implicit adoption of inflation targeting as a new monetary policy strategy by a number of central banks. Just as important, however, have been the changes in monetary policy implementation, in particular the explicit recognition of interest-rate targeting. Although the new procedures are not that much different from previous ones, it is easier to pierce through their veil. In the case of the USA, for instance, as demonstrated by the description of the central bank reaction function by J.B. Taylor (1993), which became known as the Taylor rule, the behaviour of the Federal Reserve can be interpreted as one of interest-rate setting, long before it announced official targets. As Mosler (2002, p.419) puts it, 'the federal Open Market Committee's target rate has been the focus of activity under previous Fed policies as well, and the difference is that prior to 1994 the target rate was known only to the Fed, whereas currently it is disclosed to the general public'. The procedures of some central banks are more transparent than those of others, making the horizontalist story more obvious, as is the case of Canada, New Zealand or Sweden. In other cases, such as the American system before 2008, some further investigation is required, but it can soon be demonstrated that the operational logic is identical to that of the more transparent central banks.

The Cambridge hare that was evoked by Robertson is very much present: today's principles of central banking are no different from what they were before the advent of monetarism. 'Today's views and practice on monetary policy implementation and in particular on the choice of the operational target have returned to what economists considered adequate 100 years ago, namely to target short-term interest rates' (Bindseil, 2004a, p. 10). These principles, following Fullwiler (2017), can be summarized in the following five statements:

1. The operating target is necessarily an interest-rate target.
2. Daily operations of central banks are about the payment system, not reserve requirements.
3. The operations of central banks accommodate the banks' demand for reserves by off-setting changes to its balance sheet that are inconsistent with such accommodation.
4. Compulsory reserves are means to smooth the demand for reserves; they have nothing to do with controlling monetary aggregates.
5. The implementation of a new overnight rate target does not require any change in the supply of reserves.

The fifth principle will be explained next. We can add a sixth principle: the ability of the central bank to fix interest rates is not so much related to its monopoly over the issue of high-powered money; rather, it has to do with the fact that the central bank is part of the clearing-house arrangement, and that ultimately banks and other direct clearers are required by law to settle their clearing accounts on the books of the central bank (Lavoie, 2003d, p. 538; Rochon and Rossi, 2011, p. 111). The central bank has the power to corner the market, forcing all banks to borrow from the central bank at the interest rate of its choice. Thus, in contrast to what was claimed by a number of economists, even heterodox ones, the advent of electronic money, the abandonment of compulsory reserve requirements and the reintroduction of private currency cannot undermine the power of central banks to control the overnight rate.

Four different kinds of operating system can be designed to steer interest rates with interest-rate targeting: a symmetric corridor system, a no-interest-on-reserves system, a floor system and a ceiling system. Before the advent of the Global Financial Crisis, the symmetric corridor system appeared to become the standard for monetary policy implementation. But with the 2008 financial crisis, several central banks have moved towards a floor system, and many more made the move towards this system when central banks responded to the financial and economic consequences of the COVID-19 crisis.

4.4.2 The Symmetric Corridor System

In the symmetric corridor system, the central bank announces a target overnight rate, with a symmetric corridor around it. The corridor system is also called the channel or the tunnel system. The top of the band is the lending facility rate – the old discount rate – at which commercial banks are free to borrow overnight as much as they wish, provided they have adequate collateral. Banks are also free to take daylight advances when payments go through the clearing house. The bottom of the band is the interest rate paid on the positive clearing balances (reserves) that commercial banks keep overnight at the central bank; this is the deposit facility rate, which Americans call interest on reserves (IOR). The symmetric corridor system is illustrated in Figure 4.1. Here it is assumed that there is no reserve averaging. The curve in bold illustrates what happens to the overnight rate as a function of the discrepancy between the amount of non-borrowed reserves supplied and the demand for reserves (these are called settlement balances in Canada and clearing balances in many other countries). In the particular case of Canada, where there are no reserve requirements since 1994 and where a corridor system has been in place, except for one year, between 1999 and 2020 when the COVID-19 crisis started, banks do not wish to hold any reserves under normal circumstances, so that the daily demand for reserves is zero. As a consequence, the supply of reserves is normally set at zero as well.

But what if it turns out that at the end of the day the Bank of Canada has not completely neutralized the effect of its transactions with the private sector and the commercial banks as a whole are left in a negative position at the clearing house, lacking enough clearing balances to settle payments, so that we are on the left of the zero point on the horizontal axis? In this case, some banks will be forced to borrow funds from the Bank of Canada. The banks that are in a negative position in the clearing house, as illustrated for instance in Table 4.11, will compete with each other to borrow clearing balances from the

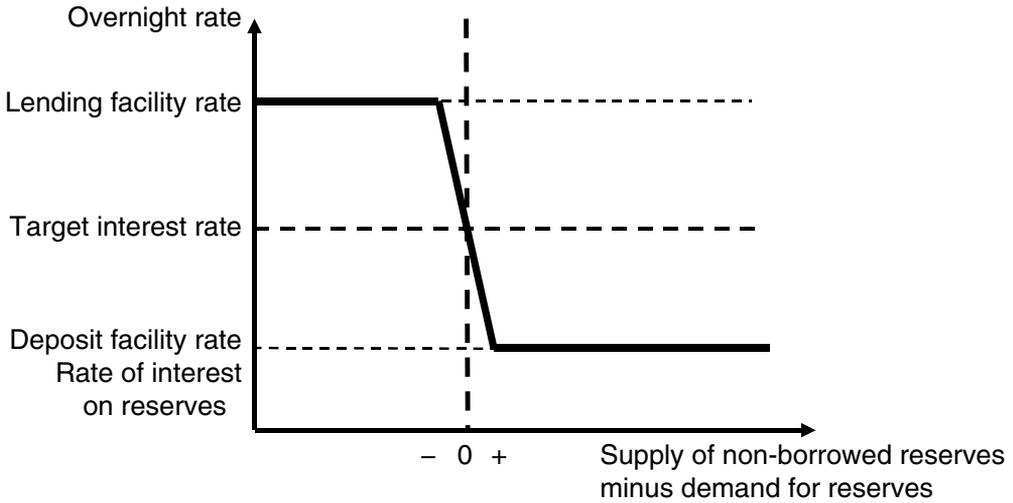


Figure 4.1 *The symmetric corridor system*

banks that have excess clearing balances. Not all such banks will be successful, however, and some will end up being forced to borrow from the central bank at the lending facility rate. The successful banks will manage to borrow funds on the overnight market, but at an interest rate that is likely to be somewhere between the target interest rate and the lending facility rate. If the deficiency in reserves is large, the overnight rate will reach the lending facility rate, illustrated in Figure 4.1.

Symmetrical effects will occur if the private banking system has been left with a positive amount of clearing balances at the end of the day, so that the banking system lies to the right of the zero point. The greater this excess amount, the more likely it is that the overnight interest rate will fall towards the deposit facility rate, as more of the excess clearing balances will find no borrower and will be deposited at the central bank for the night. Note the causality involved here. It is not that higher overnight rates relative to the lending facility rate induce banks to make more use of the borrowing facility or the discount window at the central bank, which is the usual interpretation of American scholars, an interpretation that Bindseil (2004a, p. 138) calls the reverse causation fallacy; rather, it is the lack of available clearing balances that forces banks to borrow funds from the central bank, thus pushing up the overnight rate in the process, as banks compete to obtain clearing balances.

The slope of the curve shown in Figure 4.1 depends on technical features, such as whether a bank has some market power – for instance if a single bank holds excess reserves while a large number of banks are in a deficit position in the clearing house. It will also depend on the strength with which clearing participants hold the convention regarding the target interest rate. For instance, if banks believe that the central bank has the ability and the information required to keep the overnight rate on target, the slope is likely to be flatter: the discrepancy between the supply of (non-borrowed) reserves and the demand for them will have to be large to raise or bring down the actual overnight rate towards the ceiling or the floor of the corridor.

Under normal circumstances, the overnight rate will be positioned in the middle of the corridor, a few basis points or less away from the target overnight rate. This is because overnight market participants know that they can rely on the deposit rate of the standing facilities at the bottom of the corridor if they have positive clearing balances and that, conversely, they can rely on the lending rate of the standing facilities at the top of the corridor if they have negative clearing balances. With enough competition and confidence in the participants to the clearing house, the overnight market rate ought to gravitate around the middle of the band, that is, around the target set by the central bank, as it does in Canada, within a couple of basis points. The standard differential between the top and the bottom of the corridor is 50 basis points.

As noted by Fullwiler (2013, p.182), even with an inadequate knowledge of money-market conditions, a ‘central bank could very precisely achieve its target rate simply by narrowing the corridor system between its lending rate and the remuneration rate’, reducing the differential to ten basis points, for instance. Central bankers are, however, reluctant to do so, because they fear that banks might stay away from the money markets, using instead the standing facilities of the central bank, as they believe that money markets provide information about the health of their participants. Mosler (2010, p. 103), by contrast, argues that ‘the interbank market serves no purpose’, so that we might as well have banks depositing all their surplus clearing balances at the central bank and borrowing all their missing funds from the central bank, at a single target rate – a kind of degenerated corridor system.

Bindseil (2004a, p.252) claims that ‘the simplest system for controlling short-term interest rates is that with a symmetric standing facilities corridor around the target rate’. The reason for this assessment is that within a symmetric corridor system ‘changes to the monetary policy stance, that is, of the target short-term interest rate, may be made without any change in reserve market conditions by simply moving the standing facilities corridor in parallel with the target rate’ (ibid.). In other words, as confirmed by Borio and Disyatat (2010, p. 56), ‘Crucially, the interest rate can be set quite *independently* of the amount of bank reserves in the system. *The same amount of bank reserves can coexist with very different levels of interest rates; conversely, the same interest rate can coexist with different amounts of reserves.*’ This is what they call the ‘decoupling principle’. We shall come back to this principle when we discuss the extraordinary measures that have been taken by a number of central banks since the beginning of the Global Financial Crisis and because of the COVID-19 crisis. In the meantime we can note that, with the symmetric corridor system, to change short-term interest rates, the central bank needs only to announce a new target overnight rate. The actual overnight rate will immediately adjust to the new target.

4.4.3 The US System until 2008

Things are not so simple in systems where reserves are not remunerated. The American procedures followed before 2008 are the typical example of such an asymmetric setup, as illustrated in Figure 4.2. This system is characterized by the availability of a discount window, where banks can borrow at the lending facility rate, and by the feature that bank reserves or bank deposits at the central bank get no remuneration. If banks are truly free to borrow from the central bank when the overnight rate reaches the lending facility rate,

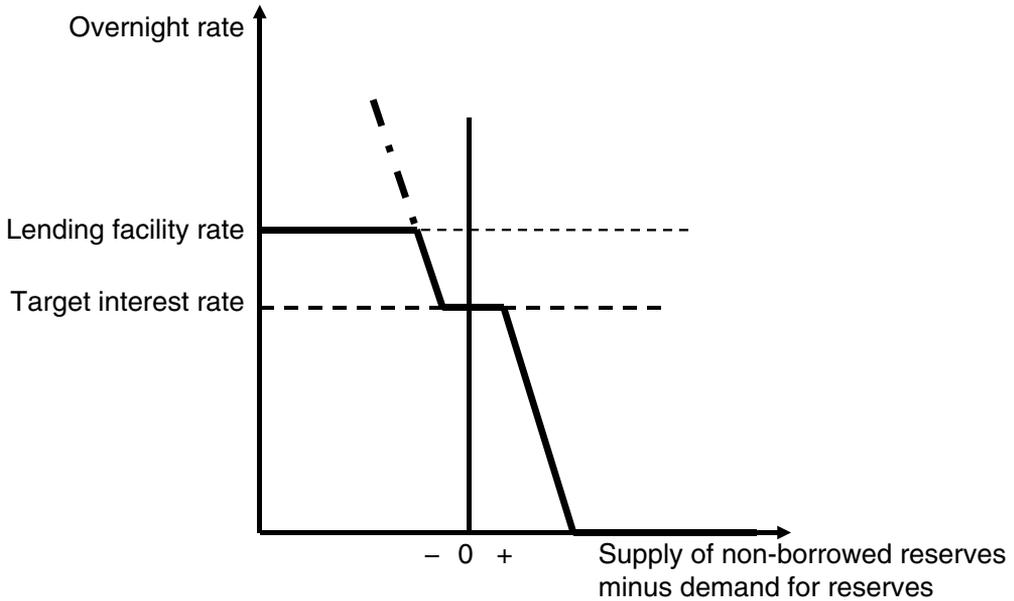


Figure 4.2 *The no-interest-on-reserves system*

then the overnight market interest rate cannot exceed the discount rate. This became (nearly) the case in the USA starting in 2002, when the discount rate was relabelled the 'primary credit rate'. Before that, because borrowing at the discount window was frowned upon, the actual overnight rate could exceed the discount rate, as shown by the dotted line, the spread between the two rates being assessed as the value of 'frown costs'. This also reflected in part the fact that the Federal funds rate is a rate without collateral (in contrast to the repo rate), whereas banks borrowing at the discount window have to show proper collateral.

The curve shown in Figure 4.2 reflects again the impact of the differential between the supply of non-borrowed reserves and the demand for reserves on the overnight rate. The curve has three flat segments. Besides the flat portions at the zero rate and at the lending facility rate, with averaged reserve requirements the curve should exhibit a flat portion around the target overnight rate, or rather at the expected overnight rate. On the last day of the averaging period, however, the demand for reserves should become inelastic and this flat portion should disappear (Whitesell, 2006; Ennis and Keister, 2008; Fullwiler, 2013). In any case, what Figure 4.2 aims to show is that interest-rate targeting is not an easy task with such a system because, unless the flat middle portion is sufficiently large, the central bank needs to supply an adequate amount of reserves to the system for the target interest rate to be achieved.

The range of possible overnight rates is wide, the more so if there are frown costs attached to the discount window. An additional difficulty is that the target overnight rate is usually set near the interest rate on lending facilities, so that there is asymmetry, as the target is not half-way between the two extreme possible values that can be taken by the actual overnight rate. This approach 'is much more challenging for the central bank and

less transparent, especially if the central bank's target rate is not published' (Bindseil, 2004a, p.87), which was the case in the USA before 1994. It should also be noted that, for changes in the target interest rate to be achieved, the central bank needs to modify the supply of reserves. It is this absence of the decoupling principle in the US monetary system, plus the lasting influence of the reserve-position doctrine – advocated by Keynes, American scholars and even American central bankers, as reflected in procedures such as 'borrowed reserve targeting' in place in the USA between 1979 and 1982 – that may have led some structuralist post-Keynesians to claim that central banks could exert quantity constraints on bank reserves.

4.4.4 The Floor and the Ceiling Systems

The subprime financial crisis and the delirium caused by the failure of the Lehman Brothers investment bank forced the Fed to modify its implementation procedures. The crisis provoked large fluctuations in the fed funds rate. In an effort to regain proper control over short-term interest rates, the Fed adopted the corridor system in October 2008. But just a month later, it moved to a new system, the 'floor system', which is still in force at the time of writing (2022). This eventually became the zero-interest policy (ZIRP), as the floor was set at zero (within a range of 0 to 0.25 per cent!).

With the floor system, the central bank sets the target overnight rate of interest at the level of the rate of interest on deposits at the central bank. In fact, the floor system is nothing new: it was in place in the USA between 1933 and 1951, when short-term interest rates were driven down to zero. Japan, attempting to come out of its balance-sheet recession (Koo, 2009), has long been on a floor system. Moreover, the floor system was adopted by New Zealand and Norway before the financial crisis. Six major central banks, including the ECB and the Bank of England, were on the floor system in 2016; due to the COVID-19 crisis, more central banks, including the Bank of Canada, have recently joined that group.

In Figure 4.3, the supply of clearing balances is set so that it largely exceeds the demand for reserves (to the right of the zero vertical line in the figure). Thus the actual overnight rate is always equal, or nearly so, to the interest rate on deposit facilities at the central bank, now also the target interest rate. There are two advantages to this procedure. First, central banks do not need to worry about being precise in their actions to neutralize autonomous flows of payments, in contrast to what is needed within the corridor system. Provided it is large enough, the supply of reserves can be of any amount. Second, central banks can let their balance sheet balloon and force banks into using their deposit facilities without worrying about the impact on overnight interest rates. This is particularly useful during a financial crisis when central banks wish to use all possible tools.

Otherwise, the floor system retains some of the advantages of the symmetric corridor system. It retains the decoupling principle. The target interest rate can be raised or lowered, with actual overnight rates responding fully to this change without having to tamper with the supply of reserves. It should be noted that the floor system was advocated before it was implemented in the USA, both by post-Keynesian authors (Fullwiler, 2005) and by central bankers (Keister et al., 2008). The only drawback of the floor system is that it creates worries among ignorant investors and mainstream economists, who believe that large amounts of excess reserves will generate price inflation.

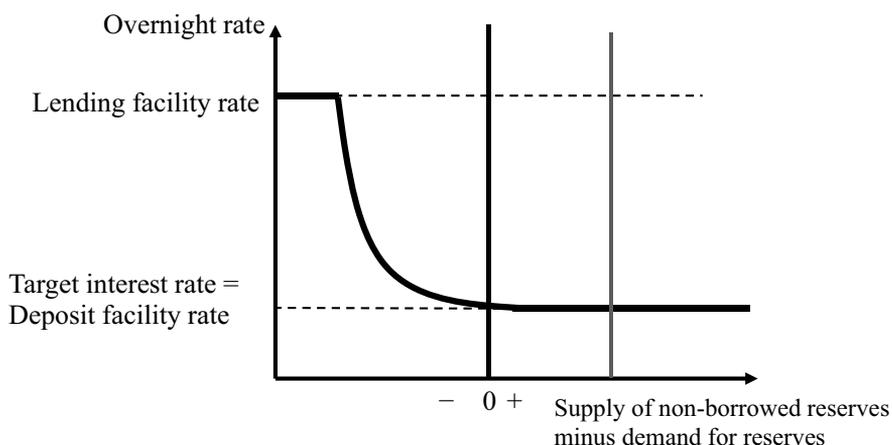


Figure 4.3 *The floor system*

A symmetric alternative to the floor system is the ‘ceiling system’. In this interest-setting system, the central bank compels commercial banks in the aggregate to use the lending facilities. This is done by supplying at all times an insufficient amount of clearing balances, forcing the commercial banks to borrow them from the central bank (the system is always on the left of the zero vertical line in Figure 4.4). This was indeed a typical feature of overdraft economies, within which as we saw when discussing the anti-chartalist story, central banks do not acquire any government securities, forcing banks to take advances from the central bank, and hence, as was said, ‘to be in the Bank’. Most continental European countries were in this situation until asset-based systems became the fashion because overdraft economies were perceived to be more inflation-prone. Different rates could be charged to the borrowing banks, depending on how much they had already borrowed.

Figure 4.4 illustrates the ceiling system, with the supply of reserves set at such a low level that banks taken overall are forced to come to the central bank. With the target interest rate equal to the lending facility rate, this ensures that the actual overnight rate approximates the ceiling rate. Although the ECB is usually depicted as operating within a corridor system, I would argue instead that the ECB was operating under a modified ceiling system until its sovereign debt crisis. This is because the target interest rate of the ECB was the minimum bid rate, which is the rate at which banks could engage in repo operations with the ECB, under the main refinancing operations. As was pointed out earlier, repurchase agreements are equivalent to loans. Thus, overall, eurozone banks were structurally indebted to the ECB, with the actual supply of reserves, not counting the borrowed ones, equivalent to the supply of reserves shown in Figure 4.4. With the eurozone crisis, the monetary market became fragmented, with Southern eurozone banks still borrowing from the central bank, while Northern eurozone banks were flooded with reserves (Febrero et al., 2018).

The ceiling system, just like the floor system or the symmetric corridor system, displays the decoupling principle. The target interest rate can be modified without any change in the outstanding amount of reserves. While Minsky (1986a, pp.282, 361) advocated a return to the ceiling system, most observers today believe that a floor system is preferable

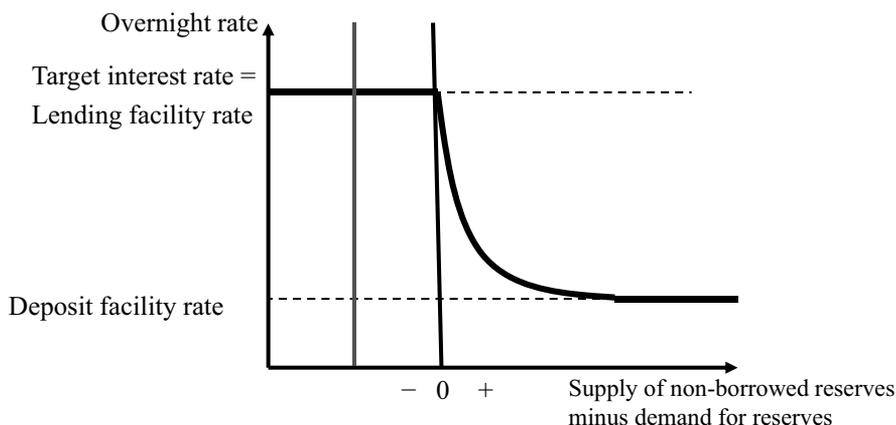


Figure 4.4 The ceiling system

to a ceiling system because in the former case commercial banks carry central bank assets rather than central bank liabilities.

The new developments in monetary policy implementation over the last 25 years and as a consequence of the financial crisis have certainly vindicated the post-Keynesian theory of endogenous money. The simplest representation of this theory is to draw a horizontal supply of high-powered money, set at the target overnight rate. This representation can also be found in a few orthodox textbooks (Cecchetti, 2006, p. 463). This is meant to represent the claim that, between meetings of the interest-setting committee of the central bank (the FOMC in the USA, for instance), the monetary authorities will supply any amount of reserves corresponding to the target overnight interest rate that they wish to achieve. The shape of this supply curve was the subject of an intense debate among post-Keynesian economists, as many argued that the central bank reaction function ought to make it upward-sloping when several periods came to be considered (Fontana, 2009, p. 115). By contrast, one could take the view here that, from a daily perspective, the supply of clearing balances is a vertical line, as the central bank has substantial control (as in the USA or Europe) or perfect control (as in Canada) over the supply of non-borrowed reserves.

In past writings of post-Keynesians, mine included, this distinction was not always made clear. Along with others, I used to argue that central banks have little control over the overall supply of reserves (the sum of borrowed and non-borrowed reserves). For instance, when sitting on the left of the zero line in Figure 4.2, banks access the discount window and get the reserves they need. This has been verified time and time again, in particular when the Fed attempted to reduce the overall amount of reserves by cutting into the amount of non-borrowed reserves, only to find no change in the overall amount as banks made use of the lending facilities. If too many (non-borrowed) reserves were provided, banks that had previously taken advances at the central bank would use these extra reserves to reduce their indebtedness *vis-à-vis* the central bank – the case of overdraft economies with a ceiling system – thus getting the supply of reserves equal to the demand for reserves through a kind of Kaldorian reflux mechanism. But with the floor system, the central bank is able to achieve its target interest rate while simultaneously setting an amount of excess reserves of its choosing.

4.4.5 Credit Easing and Quantitative Easing

A look at the strategies pursued by the Fed

What remains intriguing is the case of the floor system when the central bank is flooding the banking system with excess reserves in an asset-based economy. In this case, one cannot say that the supply of reserves is demand-led. Banks do not necessarily wish to hold these huge amounts of reserves. In that sense, the supply of reserves is not necessarily equal to the demand for reserves. There is nothing that banks can do in the aggregate to remove excess reserves as long as they are not indebted to the central bank or as long as there are no compulsory reserves. Thus our monetary principles need to be slightly modified: the supply of reserves can be said to be demand-led, but only in so far as the target interest rate is not set at the bottom of the corridor.

The floor system has been accompanied by quantitative easing policies, which have generated many questions among economists and the general public. What can post-Keynesian theory say about these? The first obvious question relates to the purpose and efficiency of quantitative easing policies. Ironically, these policies, based on monetary aggregates, as were monetarist policies, are now called non-conventional policies, because they do not rely on interest-rate targeting. To some extent, they seem to be a remnant of the academic belief in the quantity theory of money.

Some authors distinguish between credit easing and quantitative easing. In the case of credit easing, the central bank proceeds to a swap, borrowing or purchasing illiquid private assets or long-term government securities that are less liquid in exchange for highly liquid Treasury bills. This helps those banks and other financial institutions that lack liquidity, as they can now use the liquid Treasury bills to borrow funds from more liquid banks through repurchase operations. The swap corresponds to situation A in Table 4.22: Treasury bills and private assets are exchanged between the central bank and the commercial banks. This is what the Federal Reserve did for about a year, between August 2007 and September 2008. The swaps have no repercussions on reserves, because the purchases of private assets by the central bank are ‘sterilized’ by the sale of Treasury bills to the commercial banks. This, however, cannot go on for ever, because at some point the central bank will run out of Treasury bills.

Table 4.22 Accounting for balance-sheet changes during the Global Financial Crisis

Central bank (the Fed)		Commercial banks	
Assets	Liabilities	Assets	Liabilities
(A) T-bills: -100	(B) Govt deposits: +100	(A, B) T-bills: +100	Deposits Own funds
(A, B, D) Private assets or long-term securities: +100	(C, D) Reserves: +100	(C, D) Reserves: +100	(C) Fed credit: +100
(C) Advances: +100		(A, B, D) Private assets or long-term securities: -100	

This is why the Fed adopted a new credit-easing scheme. While it continued to purchase (or borrow) private assets from financial markets, it asked for the collaboration of the US Treasury, requesting it to auction equivalent sums of Treasury bills that were sold to financial institutions. This was the so-called Treasury Supplementary Financing Program (SFP). The proceeds of these Treasury-bill sales were then repatriated in the form of government deposits at the Fed, thus draining the reserves created by the use of the various credit facilities of the central bank. Other central banks, such as the Bank of Canada, organized identical transactions. This scheme, which was also pursued in the Spring of 2020, at the beginning of the COVID-19 pandemic in North America, produces a swap that is no different from the previous one: the central bank takes in illiquid private financial assets while the private sector acquires liquid short-term government securities. But this time, in contrast to what had occurred previously, the size of the Fed's balance sheet rose precipitously: the rise in its holdings of private assets was not compensated by a decline in its holdings of government securities; instead it was accompanied by a rise in its liabilities – the increased deposits of the federal government at the central bank. This corresponds to situation B in Table 4.22.

This new scheme lasted only a few weeks, however, as American authorities started to worry about the approaching US government debt ceiling. Just before adopting the floor system, the Fed gave up on sterilization, as banks and other financial institutions started to borrow huge amounts from the Fed, thus creating reserves and boosting the size of its balance sheet. This is shown by case C in Table 4.22. As the use of credit programmes declined, the Fed switched to quantitative easing in 2009, by purchasing large amounts of private assets, thus creating reserves, still without attempting to sterilize them. This is case D in Table 4.22. Further episodes of quantitative easing then followed whenever the Fed felt that the US economy was moving away from recovery.

The new normal: a floor and a sub-floor at the Fed

Fed officials had some trouble in explaining the new monetary policy operational framework and its consequences to people trained in mainstream economics. Fed officials now distinguish between *limited-reserves* regimes and *ample-reserves* regimes, with the former corresponding to Figure 4.2 while the latter corresponds to Figure 4.3 and what we called the floor system. Indeed, the Fed has now decided that despite planning to gradually diminish the size of its balance sheet after the COVID-19 shock, 'the level of reserves will be such that the Fed continues to implement policy with an ample level of reserves' (Ihrig and Wolla, 2020, p. 22). The new normal at the Fed is thus a floor system.

In fact, this floor system had to be slightly modified, as Fed operators kept facing a puzzle. While theory would tell us that the federal funds rate cannot fall any lower than the interest rate on reserves (the IOR), the reality was that with the IOR at 0.25 per cent the federal funds rate oscillated between 0.05 and 0.20 per cent. The reason, it was found, is that government-sponsored enterprises (such as Fanny Mae or Freddy Mac) could not receive interest payments on their reserve balances at the Fed, so that, lacking bargaining power, they were offering them to other financial institutions at a rate between zero and the IOR rate.

When, in 2014, the Fed started considering raising the target interest rate, which happened in 2016, it became concerned with this slack control over the fed funds rate as it

feared that raising the floor rate might not succeed in achieving its new targets. The Fed thus set up a sub-floor, by providing reverse repos to financial institutions who were denied, by law, access to interest-paying reserve accounts at the Fed, such as the said government-sponsored enterprises and money market funds. These reverse repos were called overnight reverse purchase agreements (ON RRP). The rate paid on ON RRP transactions is set at 25 basis points below the interest rate on reserves. In terms of Figure 4.3, the target federal funds rate would be in between two horizontal lines at the low end of the graph. The ON RRP thus acts as a supplementary tool for the successful operation of the US floor system, as the effective federal funds rate gravitates in between the floor and the sub-floor (Ihrig and Wolla, 2020; Grossman-Wirth, 2019). Another puzzle occurred in 2019 and 2020 when both the federal funds rate and one-day repo rates went above the top of this floor corridor, thus inducing the Fed in 2021 to institute a standing repo facility, which allows banks to convert government bonds to reserves on demand at an administered rate.

Purpose and puzzles of quantitative easing

While credit-easing operations had the obvious goal of supporting the prices of faltering private assets and providing liquidity to the financial sector, thus avoiding a complete meltdown, the justification for quantitative-easing policies when conditions have stabilized is not so clear. Once there is a large amount of reserves in the private banking system, ensuring that the overnight rate stays at the floor of the corridor, there is no obvious justification for handing out reserves exceeding the required amount by a factor of 20 or 30.

Once the financial crisis has been contained, there are three broad transmission mechanisms that might justify quantitative easing operations (Fiebigler and Lavoie, 2020; 2021). The first set of mechanisms can be associated with post-Keynesian theory. The second mechanism relies on some sort of signalling. As to the third mechanism, it relies on Friedman's monetarism.

We start with the first transmission mechanism. The purchases of government bonds or of private financial assets support the prices of these assets, thus reducing long-term yields and allowing companies to issue bonds and shares at a lower cost to finance their investments. The purchases may also generate capital gains for rentiers, or help them avoid capital losses, thus inducing them to increase their consumption. Keynes (1930b, pp. 370–1) himself mentions these two effects when he argued that in conditions of acute slump central banks could conduct open-market operations *à outrance*. This in addition would bring short-term rates to zero. From a post-Keynesian perspective, there is nothing wrong with these two claims, which are similar to those of credit easing, although the ultimate effects on aggregate demand are likely to be modest. In an open economy, the low interest yields may also generate a depreciation of the currency, as domestic agents may use their newly-acquired money balances to rebalance their portfolio by purchasing foreign financial assets.

The second transmission mechanism is based on various signalling effects. It is said that quantitative easing will engender a rise in expected present or future inflation, and thus an increase in actual inflation and a decrease in expected real interest rates. It is also thought that by boosting asset prices, it will increase the collateral values of borrowers and the equity of lenders. These latter effects are associated with the bank lending channel

of New Keynesians. Post-Keynesians have little confidence in these. As to the third transmission mechanism, it relies on the standard monetarist story and the intellectual influence of Milton Friedman. It is said that the sellers of the private assets will deposit the proceeds in banks, which will now have more funds and more reserves at the central bank to make loans. Such proponents of quantitative easing then rely on the standard money multiplier story, itself based on the myth of the fractional reserve system, claiming that the new deposits will permit a multiple expansion of credit and money aggregates, and hence increases in nominal GDP, in total contradiction to post-Keynesian theory.

Two phenomena associated with quantitative easing have been puzzling for mainstream observers of monetary aggregates. First, as bank reserves of ever greater size accumulate at the central bank, observers complain that commercial banks are refusing to collaborate or are somehow constrained and unable to use them to provide new loans. This was notably the case in the USA. By contrast, within the eurozone, which was an overdraft system, quantitative easing was not necessarily associated with rising reserves, as several European banks used instead their newly acquired clearing balances to reduce their debt *vis-à-vis* other banks, thus allowing the latter to reduce their overdrafts at the ECB. There is no puzzle from the post-Keynesian point of view: holding extra reserves at the central bank will not induce banks to find more creditworthy borrowers.

A second puzzling feature, noted particularly in the UK, is that increases in the official definitions of the money supply, however measured, seem to be much smaller than increases in bank reserves. This is puzzling to most observers, because they would argue that 'surely the central bank remains able to *increase* the supply of money and credit to the economy, for instance by autonomously deciding to step up its purchases of private sector assets (such as bills, commercial paper, bonds, equity, real estate, and so on)' (Werner, 2005, p. 59). But this second puzzle is easy to resolve if one is aware of the reflux principle put forth by the Banking School and Kaldor. This is shown in Table 4.23, where it is assumed that non-bank agents are the sellers of the financial assets purchased by the central bank, here assumed to be gilts (government long-term securities) or asset-based securities (ABS). Indeed, in reality, non-bank agents rather than banks often were the main counterparty to the quantitative-easing operations of central banks. This was a conscious decision of central bankers, as they thought that in previous quantitative easing operations the banks had declined to grant more loans despite the presence of excess reserves. The idea was thus to directly provide firms and households with additional money balances.

Table 4.23 *Balance sheet of commercial banks and sellers of illiquid assets following a quantitative-easing operation*

Commercial banks		Sellers of financial assets	
Assets	Liabilities	Assets	Liabilities
Reserves +100	Deposits +100	Gilts and ABS -100	
		Deposits +100	
Reserves +100	Deposits +10	Gilts and ABS -100	Loans -90
Loans -90		Deposits +10	

When the central bank purchases financial assets, the sellers deposit the proceeds of the sales at their commercial bank, and banks acquire reserves through the clearing house. This is non-controversial and shown in the first row of Table 4.23. But why do private agents accept to sell their financial assets? It may be because they are only too happy to get rid of illiquid assets at what they think is a good price. It may also be because they wish to deleverage and reduce their debt. The effect of such a decision is shown in the second row of Table 4.23. While banks are still holding 100 units in reserves, it may be that 90 per cent of their clients have decided to use their newly acquired deposits to reduce the amount of bank loans that they had previously taken. In this case, a quantitative-easing operation of £100 billion sterling will only lead to an increase of £10 billion sterling in the money supply. Richard Koo (2009), an iconoclast economist working for a financial firm, explains that this is likely to happen in a balance-sheet recession during which agents do their best to pay down their debt, as is the case in Japan, the USA and the UK. Thus there is no mystery here.

Quantitative easing appears as a desperate attempt by monetary authorities and some economists still adhering to monetarism to demonstrate that monetary policy is always effective. Quantitative easing is often justified by the belief that loan officers check the reserves position of their bank before making new loans. Until 2001, Japanese central bankers were reluctant, and rightly so, to pursue quantitative-easing operations, because the 'Bank of Japan argued vigorously that such measures would be meaningless' (Koo, 2009, p. 73). The reason, as Koo (*ibid.*, p. 75) explained, is that 'borrowers – not lenders, as argued by academic economists – were the primary bottleneck in Japan's Great Recession'. This is tied to the sayings that 'you can't push on a string' or that 'you can bring a horse to water, but you can't force it to drink'. This explains the first puzzle noted above. Now, as Fullwiler (2013, p. 184) remarks, 'whereas this might have been a time for economists to consider that perhaps the money multiplier model and the Quantity Theory of Money have causation wrong, instead the lack of impact from increased quantities of excess reserves is often blamed on IOR [interest on reserves]'.

Alternative versions of quantitative easing

While quantitative easing, as presented here, has been applied in several countries, a number of economists and activists have been advocating variants of quantitative easing. Table 4.24 sketches some of these alternatives (Fiebiger, 2016a; Lavoie and Fiebiger, 2018). In all cases it is assumed that households wind up with the proceeds of the fiscal or monetary intervention, and hence improve their net worth position. Case 1 corresponds to a standard fiscal operation, where the government spends by issuing bonds, here held by banks. Case 2 corresponds to the standard case of quantitative easing, where the issued bonds are purchased back by the central bank. Case 3 describes a proposal made by a number of European economists (Couppey-Soubeyran, 2020), in an attempt to circumvent Maastricht rules and provide more fiscal space to eurozone countries: the central bank cancels the government debt that it holds on its balance sheet. As shown, this leads to a drop in its net worth, while the debt of the government disappears. Case 4, which in an accounting sense, is identical to case 3, corresponds to what has been called quantitative easing for the people or also helicopter money. Here, the central bank – instead of the government – sends funds to households, thus creating again a hole in its net worth.

Table 4.24 Fiscal operation, quantitative easing and alternatives

Case	Government		Central bank		Commercial banks	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1		Bonds $B + 10$ Net worth -10			Bonds $B + 10$	Household deposits $+10$
2		Bonds $B + 10$ Net worth -10	Bonds $B + 10$	Reserves $H + 10$	Reserves $H + 10$	Household deposits $+10$
3, 4				Reserves $H + 10$ Net worth -10	Reserves $H + 10$	Household deposits $+10$

What are the flow consequences of these various actions? In case 1, the government has to make interest payments equal to $i_{LT}B$, where B is the amount of the debt and i_{LT} is the long-term rate of interest on bonds. In case 2, the government needs to make the same interest payment as in case 1. The central bank receives a flow of $i_{LT}B$, but it has an interest outflow equal to $i_{TR}H$, where H is the amount of bank reserves (with $H = B$) and i_{TR} is the target rate of interest, that is, the rate of interest on reserves. The central bank thus makes a profit of $(i_{LT} - i_{TR})B$, which can be transferred to the account of the government. The net cost of case 2 for the government is thus $i_{TR}B$. As to cases 3 and 4, the government makes no interest payment, but the central bank needs to make interest payments equal to $i_{TR}H$. These payments will diminish the profits of the central bank by the same amount, thus reducing the dividends that it can send back to the government. The net cost of cases 3 and 4 for the government is thus $i_{TR}B$ – the same cost as in case 2. In a world where central banks pay interest on reserves, cancelling the government debt held by the central bank or going for helicopter money thus makes no difference whatsoever in terms of flow payments compared to a standard quantitative easing operation. And it makes very little difference compared to a standard fiscal operation. What really counts for flow payments is the actual level of the interest rate. Somehow, the debt is always there and cannot go away.

4.5 LIQUIDITY PREFERENCE

The efficient-market hypothesis sees agents in financial markets acting essentially as arbitragists who intervene when they see small deviations in the prices of financial assets. They are actors who make sure that financial resources are *allocated* efficiently – note that nothing is created here. The arbitragists ensure that prices move smoothly, in a continuous way. In his excellent critique of rational financial markets, Justin Fox (2009) recalls how Lawrence Summers (1985, p. 634) depicts the rational-market hypothesis. Its adherents ‘have shown that two quart bottles of ketchup invariably sell for twice as much as one quart bottles of ketchup except for deviations traceable to transaction costs . . . They ignore what seems to many to be the more important question of what determines the overall level of asset prices.’

There is another element that seems to be entirely absent from the efficient-market hypothesis: liquidity. It is all very well, as mainstream authors do, to claim that markets

find the correct price at all times. Let the prices fall and someone will be willing to buy, they say. But this occurs only if there is enough liquidity (Davidson, 2009). If all orders go one way, as everyone wants to sell assets, the market will freeze unless there is a market-maker out there who is willing to swim against the tide. The importance of liquidity can be recalled through the following story, told by Paul Jorion (2008). Take a bull livestock breeder. Suppose he owes \$10 000 to his banker. Now comes the time to pay. He doesn't have the cash to pay back the loan so he proposes to his banker that he will raise the money by auctioning off one of his bulls. The auction market is full of livestock breeders. Nobody, however, is willing to buy at \$10 000, nor at \$9000, and nor at \$8000. The banker tells the breeder to forget it, and to leave the auction. The livestock breeder insists that pretty soon someone will buy his bull at a fair price. Then the banker says. 'Don't you understand? All the livestock breeders here owe me money!' As Keynes (1936, p. 155) says, the objective of most participants in financial markets is 'to pass the bad, or depreciating, half-crown to the other fellow', as was pretty obvious during the subprime financial crisis.

So far most of our attention has been devoted to the links between the banking system and the central bank. Not much has been said about the behaviour of commercial banks or about the determination of interest rates – the target rate of the central bank and other interest rates. And hardly any space has been devoted to the concept of liquidity preference. We remedy to this in the present section.

4.5.1 Some Controversies over Liquidity Preference

The rate of interest, but which one?

Many of the controversies between structuralist and horizontalist post-Keynesians arose, in my view, from two related misunderstandings. There has been confusion about the meaning of 'the' rate of interest; and there has been confusion about the meaning of 'liquidity preference'. Broadly speaking, I think it is fair to say that horizontalist post-Keynesians have the target rate of interest of the central bank in mind when they talk of 'the' rate of interest. This for them is the base rate – the benchmark rate. On the other hand, structuralist post-Keynesians view the long-term rate of interest as 'the' rate of interest. Obviously, to talk of 'the' rate of interest is a simplification, and this possible confusion illustrates only too well that economists should be more careful and differentiate clearly between the two rates, and possibly even more rates, as emphasized by James Tobin (1982).

The confusion over 'the' rate of interest may perhaps be attributed to Keynes (1936, p. 167), who found it convenient 'to mean by the rate of interest the complex of the various rates of interest'. Thus, when Keynes himself says that liquidity preference determines 'the' rate of interest, it is not quite clear whether he means the short-term rate, the long-term rate, or the spread between the short- and the long-term rates. I tend to favour the last of these three possibilities, since although Keynes (*ibid.*) says that he will assume that 'money is co-extensive with bank deposits', he also recognizes that he could just as well include instruments such as Treasury bills in the definition of money.

For instance, Keynes's concept of the liquidity trap cannot apply to the short-term rate, which we know can be brought down to zero; rather it must apply to the long-term rate or to the spread between the short- and the long-term rates. The proper view of the liquidity trap is thus a situation where the central bank gets the base rate down, but without much

impact on other interest rates. The cause of the liquidity trap, as explained by Keynes, is twofold: when interest rates are low, even a small change in the yield on bonds will produce a large capital loss (Godley and Lavoie, 2007a, pp. 167–8); furthermore, even if the central bank provides advances at zero per cent while deposits at banks are remunerated at zero per cent, banks still face fixed costs that need to be covered through interest rates on loans or other charges. Still, amazingly, many eurozone governments had negative interest rates on their 10-year bonds, and in 2021 only German bonds of a 30-year duration carried a positive interest rate, thus questioning Keynes's liquidity trap!

Keynes's view of the liquidity trap is thus different from its orthodox interpretation, according to which the liquidity trap is associated with the 'zero lower bound' – the fact that short-term nominal interest rates cannot fall below zero (Krugman, 1998). Post-Keynesians like Moore (1988, p. 264) had long recognized that, because the 'zero nominal floor on deposits imposes a positive nominal floor on bank lending rates', expansionary monetary policy was rendered ineffective 'during periods when prices are falling'.

When horizontalist post-Keynesians claim that 'the' interest rate is exogenous, they refer to the target interest rate set by the central bank, as well as the short-term interest rates that should gravitate around it. This makes structuralist post-Keynesians uneasy because for them the key rate of interest is the long-term rate, and that rate is obviously neither exogenous nor under the direct control of the central bank. Similarly, when horizontalist post-Keynesians like Kaldor (1982, p. 26) affirm that 'liquidity preference turns out to have been a bit of a red herring', leading some post-Keynesians to complain that horizontalists leave 'no room for liquidity preference in the determination of interest rates' (Wray, 1989, p. 1187), what the horizontalists have in mind is the target interest rate set by the central bank. This rate is not determined by the liquidity preference of the public. Dow and Dow (1989, p. 148) understood this well when they wrote that, if liquidity preference is understood in a narrow way, as the determinant of the rate of interest on short-term assets, then it is 'not of much interest' and 'the liquidity preference theory of interest becomes irrelevant' since 'the monetary authorities can, and normally do, set short-term rates in any case'.

To understand the role of liquidity preference, its meaning has to be enlarged to all agents of the economy and to all financial instruments, as Keynes had it in his *Treatise on Money*, rather than being limited to households and money, as in the *General Theory* or the standard IS/LM model. Such a generalization was proposed early on under various guises by several post-Keynesian authors (Kregel, 1984–85; Mott, 1985–86; Le Héron, 1986; Wray, 1992a; Brown, 2003–04). Besides the impact on the lending activity of banks and their lending rates, which will be discussed in a later section, several authors have suggested, following Hicks (1974, p. 45), that, broadly speaking, liquidity preference is measured by the differential between short and long interest rates: 'liquidity preference determines not the long rate of interest but the spread between the short and the long rates' (Wells, 1983, p. 533). As Dow and Dow (1989, p. 148) again rightly put it, 'liquidity preference, then, in practice determines the difference between the interest rate on liquid deposits and on less liquid substitutes. The monetary authorities set the rate at the short-term end of the spectrum; liquidity preference (along with other considerations) determines the mark-up to long-term rates.'

Thus it would seem that there was not much to quarrel about on this front. Still, a controversy erupted about the determination of the base rate itself. Whereas it could be

recognized that the monetary authorities set a target rate of interest, it could also be contended that ‘market opinion’, or the financial lobby, determines the conventional rate that the central bank wishes to establish. Indeed, for a number of years, especially in the 1980s, when they imposed high nominal and real interest rates, central bankers in many countries were setting the discount rate equal to the Treasury-bill rate plus a certain number of basis points, pretending then that they were in no way responsible for the high Treasury-bill rates. Indeed, Paul Tucker (2004, p. 359), from the Bank of England, recognized that the monetarist episode ‘had the unfortunate effect ... of clouding the Bank’s thinking about the feasible role of open market operations in the framework for setting interest rates’. This led Moore (1991, p. 412) to retort that ‘central banks must be held accountable for the ruling level of short-term interest rates, no matter how much they disavow’.

What rate of interest causes the others?

This lack of transparency, plus the misunderstandings just evoked, led to a controversy over causality. Both Moore (1988) and Pollin (1991; 1996; 2008) ran Granger–Sims causality tests that showed bidirectional causality between long-term market rates and the discount rate or the federal funds rate. Pollin (1991) also ran Granger–Sims causality tests between the federal funds rate and the discount rate, which at the time could be said to play the role of the target rate. Pollin (*ibid.*, p. 390) then argued that ‘the predominant causality runs from the federal funds rate to the discount rate’, and from his further tests he concluded that there was support for interactive causality, ‘with primary influence running from the market to the Fed’ (Pollin, 1996, p. 510). Pollin’s (2008) newer tests again showed two-way causality between long-term interest rates and the federal funds rate, with the stronger effect often going from the former to the latter.

Using a different method, Atesoglu (2005) concludes instead that there is unidirectional causality from the federal funds rate to the AAA bond yield and the 30-year Treasury-bond yield, a result similar to the unidirectional causality going from fed fund rates to mortgage rates found by James Payne (2006–07), with a complete pass-through (Cook, 2008). Finally, both Pollin (2008) and Atesoglu (2003–04) agree that the prime rate (the rate at which banks lend funds to good customers) has been clearly determined by the federal funds rate since the mid-1990s and during the 2000s. This is not surprising because, since the early 1990s, banks in the USA set the prime rate simply by adding 300 basis points to the target federal funds rate.

There has been a revival of interest for this causality issue. As one would expect from these delicate time-series tests, results have been disparate despite involving only US data. Rahimi et al. (2016), by combining linear and non-linear Granger causality tests, conclude the federal funds rate causes the 10-year Treasury-bond rate, the more so during recent business cycles. Rahimi et al. (2017), with another econometric strategy, find instead more evidence of bi-directional causality. Levrero and Deleidi (2021), using yet another method, find bi-directional causality when the long-term interest rate is the 10-year sovereign bond, but find instead unidirectional causality going from the federal funds rate to the 10-year AAA corporate bond. As to Kim (2019), looking at the 2004–2008 period only but using several different long-term rates, he finds a long-run bi-directional causality between the federal funds rate and long-term rates, while in the short run there would be a unidirectional causality going from the federal funds rate to long-term rates.

Moore's (1991) and Palley's (1991) interpretation of Pollin's (1991) results was that long-term rates partly reflect future expected short rates, as in the Hicksian expectations theory of long-term interest rates, and this is why they carry information about future federal funds rates. This, as pointed out by Pollin (2008) and Kim (2019), seems a rather weak explanation for the two-way causality linking short and long rates. Anyway, recent evidence shows that investors are totally unable to correctly anticipate future long-term interest rates just one semester or one year ahead (J.P. Morgan, 2016). If we accept that this two-way causality exists, a possible alternative explanation is that both rates respond to a third variable, for instance inflation. An even better explanation, as was clear during the subprime financial crisis and also in March 2020 at the beginning of the COVID-19 crisis, is that, when the economy is faltering, long-term interest rates and yields on private assets tend to increase, due to rising uncertainty and rising default risks. This induces the central bank to reduce the target federal funds rate in an attempt to bring down long-term rates and yields on private assets. The fact that long rates 'cause' short rates does not necessarily mean that rising long-term rates induce the central bank to raise targets on short-term rates; it may mean instead that rising long-term rates trigger a reaction from the central bank, which attempts to stop the increase in most of the spectrum of interest rates by lowering short-term rates.

The argument built around expectations makes more sense when we try to understand why, sometime in the past, federal funds rates appeared to 'cause' the discount rate. As Mosler (2002, p.420) points out, 'markets anticipate Fed action'. This, however, as he continues, 'is in sharp contrast to the notion often supported by the media that market rates, rather than anticipating Fed action, contain information as to where the Fed should target the federal funds rate', the target being roughly the discount rate at the time that Pollin (1991) was writing. Pollin (*ibid.*, p.389) seemed somehow to accept the media view – that is, the view that market opinion pushes the central bank around, when he made the surprising claim that 'although the Fed autonomously sets the discount rate, it does so not simply to influence the market but to keep apace with it'.

In his latest review of the topic, Pollin (2008) seems to acknowledge that, despite his evidence of two-way causality, the federal funds rate can be considered as an exogenous variable, under the control of the Fed. But he adds that this rate has little influence on long-term rates, especially those on private assets, which are essentially determined by liquidity preference features, such as risk of default assessment. Thus, once again, when all is said and done, there seems to be little to disagree about. Horizontalists are likely to claim that, if monetary authorities are sufficiently insistent and consistent, a shift in interest-rate differentials can only be temporary. Summoning Keynes's authority, it could be said that the convention established in the prevailing spread 'will not be always unduly resistant to a modest measure of persistence and consistency of purpose by the monetary authority' (1936, p.204). In the case where long-term rates were high relative to short rates, financial operators would come to realize that substantial profits could be made by borrowing short and lending long. Eventually, unless 'the authorities' nerves are shaken by the ferocious growls with which the bears have been deafening them all this time', the convention enshrined in the base rate should prevail, and the liquidity premium should be back to its normal level (Robinson, 1952, p.30).

Fullwiler (2013, p.192) suggests that the monetary transmission towards long rates would be improved if central banks were to announce long-term interest-rate targets.

Indeed, Keynes (1936, p.206) made the same suggestion when he wrote that ‘perhaps 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 practical improvement which can be made in the technique of monetary management’. At this stage, Pollin’s (2008) and Cömert’s (2013) argument is that deregulation and financialization have weakened the ability of the monetary authorities to influence long-term rates as described above by Keynes and Robinson, and hence that rules and regulations should be introduced to help the central bank steer long-term rates towards values that are compatible with the targeted short-term rates. Indeed, extraordinary measures have been taken by central banks, and this may explain why, contrary to predictions, they have retained influence over long-term rates.

4.5.2 The Reaction Function of the Central Bank

Let us now assume that indeed the central bank has the power to set the target overnight interest rate at the level of its choice. In the previous section we examined the implementation procedures that the central bank can put in place to ensure that the actual overnight interest rate is equal to the target interest rate set by the central bank. But we have not discussed what this target rate of interest ought to be. The fact that the base rate of interest is considered to be exogenous in post-Keynesian theory does not mean that it should remain forever pegged at a constant level. Ever since the appearance of J.B. Taylor’s (1993) paper on the Fed’s reaction function, much attention has been devoted to how central banks set their target rate of interest. In post-Keynesian terms, within the generalized liquidity preference approach suggested by the post-Keynesian structuralists, the central bank reaction function can be interpreted as a formalization of the shape taken by the liquidity preference of the central bank. Rochon and Setterfield (2008) have proposed an excellent survey of the various ways in which the target interest rate could be set by the central bank. The present subsection is very much inspired by their paper.

The countercyclical approach

At the most global level, central banks can follow either a countercyclical policy or an income-distribution approach, which Rochon and Setterfield, not quite appropriately, call the ‘activist’ and the ‘parking-it’ approaches. A countercyclical policy means that the central bank attempts to fine-tune some economic aggregate by changing the target nominal interest rate, possibly changing the real interest rate as well. In the case of the parking-it approach, the central bank sets the nominal or the real rate of interest without attempting to fine-tune macroeconomic conditions. The countercyclical policy corresponds to the current behaviour of central banks. The income-distribution approach is more akin to a post-Keynesian view that sees fiscal policy, instead of monetary policy or interest-rate policy, as the main tool for fine-tuning the economy.

Currently, most central banks seem to pursue an activist policy based on inflation targeting. This is so, even though, as is the case of the ECB and the Bank of Canada, the statutes of these central banks do not specify a single mandate. On the other hand, the Federal Reserve in the USA has had an explicit dual mandate since 1977, based on price stability and maximum employment, and it can be argued that this is reflected in Taylor’s (1993)

assessment of the Federal Reserve reaction function. Giovannoni (2008) even shows empirically that output forecasts, rather than inflation forecasts, have driven changes in federal funds rates. As readers should expect by now, a number of post-Keynesians have argued that central banks should be concerned with unemployment rates and rates of capacity utilization, rather than with inflation. It has also been suggested that target interest rates take into account, or should take into account, the evolution of exchange rates and the rate of asset inflation, in particular on the stock market. The Global Financial Crisis has also underlined the weaknesses of a monetary-policy regime unconcerned with financial stability. In any event, there is little evidence that inflation targeting per se has really been responsible for contemporary low inflation rates (Tatliyer, 2017), while it may have slowed down GDP growth in emerging economies (Khan, 2022).

Rochon and Setterfield point out that Basil Moore, who is often depicted as assuming a constant target interest rate, can be seen as being part of the activist policy camp. They claim that Moore (1989, p.487) argued in favour of ‘a central bank reaction with short-term interest rates as the dependent variable’, which would depend on ‘the future state of the domestic economy (demand factors)’, as well as on the goals of the monetary authorities including ‘full employment, price stability, growth, balance of payments, terms of trade, exchange rates, the distribution of income’. Moore (1988, p.264) mentions in addition foreign exchange reserves and foreign exchange rates as determinants of a central bank reaction function. Indeed, Moore (1988) was quite confident that the central bank could achieve fine-tuning, although he was somewhat less confident 20 years later (Moore, 2006).

If a central bank pursues countercyclical policies, should it target a real interest rate or a nominal rate? In the New Consensus analysis of the central bank reaction function, the target is a real interest rate set around the natural rate of interest, with the target set above or below the natural rate depending on whether the central bank wishes to diminish or increase the rate of price inflation. Alan Haight (2007–08) argues instead that the fight against overly high inflation rates does not require increases in real interest rates; it is enough to raise nominal interest rates by less than the rise in inflation. The rise in nominal interest rates will be enough to slow down the economy.

Why is this so? As we will note in Chapter 6, in North America the cyclical behaviour of the economy is driven to a large extent by the housing market. In that market, prudent borrowers rely on monthly payment to monthly income ratios, which depend on nominal interest rates and not on real rates. In addition, at least until the advent of subprime mortgages and NINJA loans (no income, no job and assets), it is standard practice in the industry to determine the maximum possible mortgage on the basis of this monthly payment to monthly income ratio. As a consequence, ‘higher nominal rates lead (as any realtor knows) to smaller mortgages and smaller new homes’ (Haight, 2007–08, p.262). Because individual nominal incomes normally rise through time, most of the risk of default occurs during the early years of the mortgage, so that taking into account the nominal rate, rather than the real rate, is rational, as debt burdens will be front-loaded. The housing market is thus driven by these mortgage payments to income norms and by nominal interest rates. Indeed, in an empirical analysis of the euro area, Deleidi (2018) shows that higher interest rates do slow down loans provided to the housing market while there is no significant effect on loans granted to firms. Still, Fazzari et al. (2008, p.560) argue that nominal interest rates, at constant real rates, also slow down business

investment because firms rely on internal cash flows to finance their investments, and higher nominal interest rates on debt do reduce cash flows.

Two points can thus be made. First, the central bank reaction function does not have to be as steep as portrayed by the New Consensus. Second, low inflation targets with equivalently low interest rates generate conditions conducive to predatory lending, since they are likely to lead to overly large mortgages.

The income-distribution approach

The current view of post-Keynesians seems to trend towards the income-distribution approach. There are three reasons for this. The first reason, which has been developed by Arestis and Sawyer (2006) in several papers, is that monetary policy has only an indirect effect on inflation rates, so that inflation control can be achieved only at the cost of large losses in economic activity. The second reason is that the countercyclical approach has some ties with the natural rate of interest, a concept that the post-Keynesians reject, and that in any case has been proven to be impossible to measure with sufficient precision for actual use in central bank decisions. The third reason is that the interest rate is viewed by post-Keynesians as a distribution variable, as pointed out at the beginning of the chapter. Eichner (1987, p. 860) is quite clear about this when he says that the base rate of interest 'is a politically determined distributional variable rather than a market determined price'. Changes in real interest rates caused by inflation targeting produce changes in income distribution, affecting in particular the share of income going to rentiers. Monetary policy should thus focus on achieving rates of interest that are consistent with an appropriate income distribution. 'Monetary policy should not be designed so much to control the level of activity, but rather to find the level of interest rates that will be proper for the economy from a distribution point of view' (Lavoie, 1996e, p. 537). As Rochon and Setterfield (2008) indicate, at least three rules have been suggested in the past.

The first rule, which they call the Kansas City rule, since its proponents came from the University of Missouri in Kansas City (UMKC), simply says that the overnight interest rate ought to be zero. In its original version, the rule was to target 'a low and stable base rate' (Wray, 1997, p. 569). Forstater and Mosler (2005), supported by other neo-chartalists such as Éric Tymoigne (2009, p. 114), argue, making use of the neo-chartalist story developed with the help of Table 4.17, that if the government runs a deficit financed by the central bank, overnight interest rates will tend towards zero, and hence that this ought to be the normal case. Other rates of interest would adjust to this zero base rate. There are three problems with this proposal. First, as follows from our discussion of the previous section, it is not clear that the normal overnight rate is zero, even if the government is running a deficit. Second, the proposal, in a world where inflation is not nil, is likely to bring the euthanasia of the rentier. While Keynes (1936, p. 376) thought that this was likely to happen, for other reasons, however, the low interest rates that a number of economies have been experiencing since the Global Financial Crisis have clearly created sustainability problems for workers' pension funds. Third, in normal times, with some price inflation, near-zero interest rates will induce agents to purchase and hold real physical durable goods (real estate, precious metals), instead of holding financial assets as hedges against inflation, which is likely to lead to real asset inflation and possibly to financial instability (Epstein, 2019, ch. 6).

The next two rules ensure that real interest rates remain positive. Rochon and Setterfield (2008) call the first of these two rules the ‘Smithin rule’. This is because John Smithin (1994, p. 199) has been arguing for more than 25 years that ‘the most sensible advice for central banks to follow would be to stabilize real interest rates at low but still positive levels’. This advice has been reasserted by Smithin (2020) more recently. He argues that when monetary authorities act on countercyclical lines they tend to produce excessive fluctuations in real interest rates, moving from overly low real rates, as was the case in the 1970s, to overly high real rates, as was the case in the 1980s and 1990s (the break occurring around 1979–80). We may add that the Global Financial Crisis has led many central banks to adopt zero-interest-rate policies or quasi zero-interest-rate policies, with negative real interest rates. Smithin believes that these regime switches harm the economy, in particular when a regime of high real rates of interest is imposed on society. This, he says, is the result of years of zero or negative real rates of return on bonds, which correspond to the euthanasia of the rentiers. This, however, cannot but be followed by years of ‘revenge of the rentiers’, when the regime of high real interest rates prevails.

The solution for Smithin (1996, p. 86) is thus for the central bank to abandon inflation targets and instead

to stabilize after-tax real interest rates at low but still positive levels (say no more than 1 per cent or 2 per cent) . . . Under such a regime, the already-rich may not become very much richer, but, as long as real returns remain positive, at least they would not suffer the erosion of capital which was the original source of the anti-inflationary agitation.

Sawyer (2011, p. 286) has proposed an alternative, with the real interest rate set equal to or just below the trend growth rate of real output, so as to provide some room for fiscal policy without fearing explosive debt ratios. Hein (2012, p. 138) has made a proposal that sits somewhere between the second and third rule, by advocating ‘a slightly positive real rate of interest, below the long-run rate of productivity growth’.

The last of three distributional rules is the rule based on a *fair* interest rate. This rule was proposed by Lavoie and Seccareccia (1999; 2019). The concept of a fair rate of interest has resurfaced under the (unfortunate) appellation of the ‘natural’ rate of interest in the works of Luigi Pasinetti (1981, ch. 8; 1993, ch. 6), who has given it a very precise meaning. According to Pasinetti (1981, p. 174), the fair rate of interest ‘stems from the principle that all individuals, when they engage in debt/credit relations, should obtain, at any time, an amount of purchasing power that is constant in terms of labour (a labour theory of income distribution)’. The fair rate of interest, or what Pasinetti (2002) himself later called the *just* rate of interest, thus maintains the purchasing power, in terms of command over labour hours, of funds that are borrowed or lent, and preserves the intertemporal distribution of income between borrowers and lenders. Thus the fair rate of interest ‘means a zero rate of interest in terms of labour’ (Pasinetti, 1993, p. 92). The fair rate of interest, in real terms, should be equal to the rate of increase in real wages. In an economy where the share of profit remains roughly constant, this growth rate would simply equal the growth rate of labour productivity. With price inflation, the fair rate of interest, in nominal terms, would be equal to the average rate of wage inflation, that is, the growth rate of labour productivity plus the rate of price inflation. Formally, we have

$$i_{fair} = \hat{y} + \hat{p} \quad (4.1)$$

where, as in Chapter 1, \hat{y} is the growth rate of labour productivity, and \hat{p} is the rate of price inflation. Thus, in a world with no technical progress and no inflation, the nominal interest rate ought to be zero, as was argued by the Church at the time when these conditions were roughly fulfilled.

The constancy of the profit share is not a stylized fact anymore. Several OECD countries since the late 1980s have seen a slow secular rise in the profit share. Under these circumstances, nominal wages will grow at a rate which is smaller than that determined by equation (4.1). Lavoie and Seccareccia (2019) show that the fair nominal rate of interest then ought to be:

$$i_{fair} = \hat{w} = \hat{y} + \hat{p} - \hat{\pi}\pi / (1 - \pi) \quad (4.1A)$$

where \hat{w} is nominal wage inflation and where π is the profit share, with $\hat{\pi}$ the growth rate of the profit share.

A numerical example may help to clarify the notion of the fair rate of interest. Take an economy with a 5 per cent rate of price inflation. Suppose that the average wage is initially \$10 per hour. Suppose furthermore that a borrower contracts a \$10 000 loan. This person has thus borrowed the equivalent of 1000 hours of labour time. Suppose now that labour productivity has risen by 2 per cent during the year, and that the average real purchasing power of wages has also risen by 2 per cent. Nominal wages have thus risen by 7 per cent, reaching \$10.70 per hour a year later. If the rate of interest charged to the borrower is also 7 per cent, that is, if it is equal to the growth rate in nominal wages, the borrower will have to reimburse an amount of \$10 700 at the end of the year. However, since the average nominal wage rate has now risen to \$10.70 per hour, the amount given back by the borrower is still equivalent to 1000 hours of labour time. As long as the actual rate of interest is equal to the fair rate of interest, as defined above, the purchasing power temporarily exchanged between the borrower and the lender remains constant in labour time.

In practical terms, the Smithin rule and the fair-rate rule may not be very different. Labour productivity has been growing at around 1 or 2 per cent per year in most industrialized countries, which is the premium suggested by Smithin. It seems that Keynes (1936, p. 221) had an even simpler rule, with a real interest rate equal to zero, when he says that ‘a man would still be free to accumulate his earned income with a view to spending it at a later date. But his accumulation would not grow.’ The other practical difficulty is to identify the interest rate that we are talking about. Both Keynes and Pasinetti seem to refer to the ‘pure’ interest rate, one devoid of default risk. This would imply the interest rate on government securities, either short-term or long-term ones. One could also argue that the most relevant variable is the mortgage rate, since mortgages are generally the most important personal liability.

4.5.3 The Liquidity Preference of the Public

While risk assessment, or rather underestimated risk, has been in the news during the Global Financial Crisis, post-Keynesians have associated the concept of liquidity

preference with an environment of fundamental uncertainty. Within that environment, economic agents do not know probability distributions, they are aware that past sequences of events may not reoccur, they hope to learn more in the near future, and they may therefore delay decisions. This sort of liquidity preference is inversely related to the degree of confidence, to the reliability of beliefs, to Keynes's weight of an argument and animal spirits, as we explained in Chapter 2, and as it has been pointed out in particular by Jochen Runde (1994). A generalized theory of liquidity preference is thus intertwined with a generalized theory of rational behaviour in an environment dominated by uncertainty. But how do we express this concretely and in a simple manner? How will an increase in liquidity preference express itself? How will the rush towards liquid assets be absorbed? Here we consider the case of the liquidity preference of the public, looking at the management of households' (or their representatives') portfolio of assets.

A pricing adjustment mechanism of portfolio choices

Broadly speaking, we can say that there are two possible answers to these questions. An increase in the liquidity preference of the public, that is, an increase in the demand for safe and liquid assets, can be handled either through a change in the prices of the assets, and thus in the yields of these assets; or it can be handled through a change in the supply of these assets. Household portfolios will contain real assets, equities, bonds, Treasury bills, money-market instruments, bank deposits and cash. The case of cash, meaning here banknotes issued by the central bank, is peculiar. We know that cash is provided on demand by banks, in exchange for bank deposits. We know further that banks can procure cash to their clients either by getting advances from the central bank or by selling Treasury bills to the central bank (or through the repo market), which will necessarily be done unless the central bank wishes to create chaos in the economy. Thus an increase in the demand for cash will be absorbed by an increase in the supply of cash.

How about the other elements of the public's portfolio? There are two ways to see the necessary portfolio adjustments required by a change in liquidity preference. The first framework, the simplest one, is provided by Carlo Panico (1985). It is inspired by Keynes's chapter 17 of the *General Theory*. The idea is that the rates of return of all assets, modulated to take into account both their risk and their illiquidity (including the risk of a capital loss), need to be equalized. Each asset thus carries a risk and illiquidity discount, which we will call σ . In the case of money deposits, the risk and illiquidity discount may in fact be a positive liquidity premium. Once the portfolio has been modified to reflect the risk and illiquidity discounts, from the point of view of individual investors, the following equality should thus be realized:

$$i_{CB} = i_D - \sigma_D = i_{ST} - \sigma_{ST} = i_{LT} - \sigma_{LT} = i_{TSR} - \sigma_{TSR} \quad (4.2)$$

where i_{CB} is the target overnight rate of the central bank, i_D and σ_D are the interest rate and illiquidity discount on bank deposits, i_{ST} and i_{LT} are the yields on short-term and long-term securities and i_{TSR} is the rate of return on stock market shares, while σ_{ST} , σ_{LT} and σ_{TSR} are their respective illiquidity discounts.

Equation (4.2) can be interpreted as follows. The central bank sets i_{CB} , which is the base rate of the system. The other rates adjust to it, the independent variables being the risk and illiquidity discounts σ_i associated with each kind of asset. An increase in the

illiquidity discount associated with long-term assets will lead to an increase in i_{LT} – the yield required to hold on to long-term assets. Thus a reduction in the base rate of the system, i_{CB} , will succeed in reducing long-term rates if the illiquidity discount does not change; but it could also lead to an increase in long-term rates if the decrease in the base rate induces an even bigger increase in the illiquidity discount on long-term assets (for instance because of fears of inflation). The low long-term rates that were observed in the US in the 2000s, before the financial crisis, can be explained by this phenomenon: low and reduced illiquidity discounts on long-term bonds.

We may thus concur with Amstad and Martin (2011, p. 2) when they say that if ‘credit-risk premia . . . and the term structure of interest rates are sufficiently stable, intervention in the overnight market . . . will allow the central bank to steer longer-term rates’. By contrast, ‘in times of market stress . . . stabilizing the overnight rate will not prevent longer and more economically relevant interest rates from fluctuating with credit-risk or liquidity premia’.

An alternative adjustment mechanism of portfolio choices

There is a second way in which the portfolio choice and portfolio adjustment can be conceived. This second framework was first proposed by Tobin (1969) and slightly modified by Godley (1996). It also bears some resemblance to the asset price model proposed by Wray (1992a). The implicit assumption of their models is that assets are imperfect substitutes, and that people will not put all their eggs in the same basket, even when there are still differentials in rates of return net of the illiquidity discounts. Here, for simplicity, we provide an example with only three assets: bank deposits, short-term assets, D and A_{ST} , as well as a number of long-term securities, A_{LT} , the price of which is p_{LT} , which is inversely related to its yield. In matrix form, we have the following representation of the demand for assets:

$$\begin{bmatrix} D \\ A_{ST} \\ p_{LT}A_{LT} \end{bmatrix} = \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \lambda_{30} \end{bmatrix} V + \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} i_D \\ i_{ST} \\ i_{LT} \end{bmatrix} V + \begin{bmatrix} \lambda_{14} \\ \lambda_{24} \\ \lambda_{34} \end{bmatrix} Y \quad (4.3)$$

The various λ_{ij} parameters act as indicators of liquidity preference for the various assets. For those not at ease with matrices, the first row simply says that the demand for bank deposits D is such that

$$D = (\lambda_{10} + \lambda_{11}i_D + \lambda_{12}i_{ST} + \lambda_{13}i_{LT})V + \lambda_{14}Y \quad (4.4)$$

This means that the demand for deposits is a proportion of wealth V and of income Y (the transaction demand for money), and that the proportion of wealth is given by the parameter λ_{10} , but that this proportion is modulated by the rates of interest on deposits, short-term assets and long-term assets. The λ_{11} parameter is positive, as a higher rate of interest on deposits should induce portfolio-holders to hold more deposits, while the λ_{12} and λ_{13} parameters are negative since higher interest rates on other assets should induce investors to keep less of their wealth in the form of deposits. The λ_{14} parameter is also positive, as a higher income is likely to require transactions of a higher value and hence larger balances in bank accounts. The other two rows can be interpreted in a similar way.

There are compulsory restrictions on the various λ_{ij} parameters – the so-called adding-up conditions. These can be split into vertical conditions, as described by Tobin:

$$\begin{aligned}\lambda_{10} + \lambda_{20} + \lambda_{30} &= 1 \\ \lambda_{11} + \lambda_{21} + \lambda_{31} &= 0 \\ \lambda_{12} + \lambda_{22} + \lambda_{32} &= 0 \\ \lambda_{13} + \lambda_{23} + \lambda_{33} &= 0 \\ \lambda_{14} + \lambda_{24} + \lambda_{34} &= 0\end{aligned}$$

and into horizontal conditions, as described by Godley (1996, p. 18):

$$\begin{aligned}\lambda_{11} &= -(\lambda_{12} + \lambda_{13}) \\ \lambda_{22} &= -(\lambda_{21} + \lambda_{23}) \\ \lambda_{33} &= -(\lambda_{31} + \lambda_{32})\end{aligned}$$

The first vertical adding-up condition simply says that the total of the desired proportions of each asset must sum to unity. In other words, the sum of the demands must be equal to wealth V , so that $D + A_{ST} + p_{LT}A_{LT} = V$. You can have more of one thing by having less of another. The other vertical conditions arise from the same constraint. The horizontal adding-up conditions mean that the effect on an asset of an increase in its own rate of interest, with all other rates remaining constant, should not be any different from that of a fall, of the same size, in all the other rates, with the own rate staying the same. The horizontal constraints can be replaced by symmetry constraints, in which case the horizontal conditions will hold. Symmetry means that, in the main matrix, we have $\lambda_{ij} = \lambda_{ji}$ for all $i \neq j$, as proposed by Karacaoglu (1984). For instance, the equality $\lambda_{12} = \lambda_{21}$ implies that an increase in the interest rate on short-term assets will generate a drop in the demand for deposits of the same amplitude as the drop in the demand for short-term assets generated by an identical increase in the rate of interest on deposits. All these adding-up conditions ensure coherence of the analysis.

In the Godley and Lavoie (2007a) models, the demand for assets, given by equation (4.3), is confronted with the supply of assets. In contrast to equation (4.2), where the adjustment is presumed to occur only through prices, that is, through the values taken by the interest rates, here adjustments to portfolio preferences can occur either through the supply of the assets or through prices. In the case of equation (4.3), this implies changes in interest rates, but also changes in the price of long-term assets. For instance, looking at equation (4.4), if there is a rise in liquidity preference that compels households to wish to hold more bank deposits at the current rates of interest i_D , i_{ST} and i_{LT} , while the other agents in the economy do not accommodate by modifying the net supply of these corresponding assets, then the demand for bank deposits will have to adjust itself to the existing stock of bank deposits, through a decrease in the interest rate on deposits and an increase in the interest rate of the other two assets. Wray (1992a, p. 79) believes that this is the most likely effect of an increase in liquidity preference ‘because private institutions are not likely to issue more money in response to rising liquidity preference’.

Naturally, it should be emphasized that here we are looking at instantaneous changes in portfolio considerations. We are not concerned with the case where households acquire bank deposits through their pay cheques and their saving over the period, or where

households reduce their money balances by purchasing goods, while firms reduce their money balances by paying down their outstanding debt with the proceeds of these sales – a case associated with the Kaldorian reflux mechanism.

In Godley and Lavoie (2007a), it is assumed, except in the case of the market for stock market shares, that the adjustment to changes in liquidity preference occurs mainly through changes in asset supplies. Besides the case of stock market prices, adjustments through prices can also occur, but only in a subsidiary manner. There are two reasons for this. First, their models assumed that banks administer the rate of interest on deposits and on loans, while the central bank targets the interest rate on (short-term) Treasury bills. Second, the models purported to show that the monetary authorities, with the collaboration of the government, were able to set both the short-term and the long-term interest rates. Thus banks and the monetary authorities will modify interest rates only in so far as the actual values of the target ratios that they have assigned to themselves depart from an acceptable range. This can happen because of changes in liquidity preference (the λ_{ij} parameters of equation (4.3), in particular the λ_{i0} parameters) but also because of changes in the real economy.

Examples of changes in the liquidity preference of the public

Let us examine what can happen within such a framework if households modify their liquidity preference parameters. Let us assume that households wish to move the contents of their portfolio towards money (cash and bank deposits) and away from short-term and long-term assets, which we shall call here bills and bonds. Let us further assume, for good measure, that there is a 10 per cent compulsory reserve ratio on bank deposits, calling HPM (high-powered money) the sum of cash and reserves. Thus, let us assume that, at the current yields on assets, households as a whole wish to acquire 10 additional units of cash and 90 units of bank deposits, while getting rid of 20 units of bills and 80 units of bonds. This is shown in the first column of Table 4.25 (items [1] to [4]). Note that an increase in the holdings of an asset is represented by a plus sign, while the decrease of an asset is represented by a minus sign. It should further be noted that the increase in a liability is represented by a minus sign, while the decrease in a liability carries a positive sign. To help readers, liabilities are shown in parentheses.

The matrix accounts for changes in the stocks of assets that are made through actual transactions outside each sector. It does not include capital gains, so that both the columns and the rows of the matrix must sum to zero. In other words, if only other households purchase the bonds that households wish to discard, households as a whole will manage to reduce the proportion of bonds in their portfolio, but only through a capital loss on bonds, the prices of which will fall. In other words, the reduction in the value of bonds will imply a reduction in the wealth of households, and will not allow households to hold more of the other assets. How important this restriction is will become obvious as we go through the examples.

Here in Table 4.25, we assume that the central bank takes the necessary measures to absorb whatever change in the demand for bonds [5] and bills, at the current rates of interest, by being the counterparty to the sales, which allows the households to acquire the cash and the deposits they long for. The banks thus wind up with 90 units of extra deposit liabilities [6], forcing them to acquire an extra 9 units in reserves [7]. Banks will also purchase 81 units of bills [8]. How is that? This can be understood from the

Table 4.25 *An accommodated increase in the liquidity preference of households*

Assets	Households	Government	Central bank	Banks	Σ
HPM	+10 [1]		(-19) [9]	+9 [7]	0
Deposits	+90 [2]			(-90) [6]	0
Bills	-20 [3]		-61 [10]	+81 [8]	0
Bonds	-80 [4]		+80 [5]		0
Σ	0	0	0	0	0

clearing-house equations. As households initially sell 100 units of bonds and bills that are purchased by the central bank, banks acquire 100 units of clearing balances at the clearing house, as was shown earlier. However, banks must deposit 9 units at the central bank, as reserves, and they must acquire 10 units of cash to give to their depositors, so they are left with 81 units of clearing balances, which they use to purchase the bills. As to the central bank, it must increase its reserves and banknote liabilities by 19 units [9]. Having bought 80 units of bonds, its balance-sheet adding-up conditions thus imply that it must have net sales of 61 units of bills [10]. These correspond to the 20 units they bought from households and the 81 units they sold to the banks. Thus, as Bibow (2009, p. 122) puts it, the bearishness of the investors and the increased liquidity preference will ‘impact on the rate of interest only to the extent that the banks’, and here the monetary authorities, ‘do *not* offset such a move by enlarging the pool of liquidity’.

Table 4.25 could be taken as a sketchy description of the activities of the Federal Reserve following the 9/11 terrorist attack on New York City, which have demonstrated that a central bank can avoid large price adjustments ‘by reducing the outstanding supply of securities available for sale to the general public. The public can then satisfy its increased bearish tendencies by increasing its money holdings without depressing the market price for financial assets in a disorderly manner’ (Davidson, 2009, p. 91).

The construction of Table 4.25 raises several questions. What if banks decline to buy the Treasury bills, keeping their 81 units of clearing balances as excess reserves? Then the overnight rate would quickly drop down, making this a poor decision. What if the central bank starts to run out of Treasury bills, having sold all its stock? Then we need the collaboration of the government. The government needs to restructure its portfolio of liabilities by buying back the bonds that households wish to sell and by issuing the equivalent value of bills (which are bought by the banks or their dealers). This corresponds to the description given by Robinson (1952, p. 29): in attempting to keep long rates down, the authorities ‘issue bills in order to buy bonds, the quantity of money being adjusted to whatever level is required to keep the bill rate at its bottom stop’. All this is shown in Table 4.26. In this case, the central bank acquires 19 units of bills [11], instead of being constrained to sell them, as in the previous case.

What if banks consider that they are acquiring an overly large amount of safe assets (Treasury bills) relative to bank loans, a possibility examined by Godley (1999a)? Then they might decide to lower the interest rate on deposits so as to encourage households to hold on to their bills, thus making an adjustment through an indirect price mechanism. Similarly, the government in Table 4.26 may feel that its bills-to-bonds liability ratio

Table 4.26 *An accommodated increase in the liquidity preference of households, but with the central bank running out of Treasury bills*

Assets	Households	Government	Central bank	Banks	Σ
HPM	+10 [1]		(-19) [10]	+9 [8]	0
Deposits	+90 [2]			(-90) [7]	0
Bills	-20 [3]	(-80) [6]	+19 [11]	+81 [9]	0
Bonds	-80 [4]	(+80) [5]			0
Σ	0	0	0	0	0

becomes overly tilted towards short-term liabilities. It may then decide to stop supporting the price of long-term bonds, letting long interest rates rise. The changes in interest rates, through equation (4.3), will thus induce households to weaken their desire to move towards the most liquid assets.

Another situation of this sort can be envisaged. What if the economy is of the overdraft type, with the central bank not holding any government securities? Then the banks would have to buy the securities, and get advances from the central bank. From this we can see once more why a sovereign debt problem can arise within the eurozone, where the ECB and its national central banks by convention did not purchase any government securities. In normal times, banks purchase the bills and bonds that investors no longer want to hold, getting the advances that they need from the central bank, as shown in Table 4.27. But if banks refuse to purchase the securities, for instance because they fear a sovereign default, then all the adjustment to changes in liquidity preference will fall on asset prices and interest rates. The adjustments in prices and yields will be such that households will be induced to hold their assets in the same proportions as before their change of preferences.

There is a final case worth considering, that of a diversity of opinion within the same sector. Let us split households into bears and bulls, the bears being willing to divest themselves of bonds, and the bulls being keen to purchase them. This situation corresponds to the case discussed by Bibow (2009, pp. 103–4), based on Keynes's *Treatise on Money*, when Bibow says that 'the banking system . . . may facilitate a changing degree of diversity of opinion within the general public . . . by providing advances ("financial loans") to the "bulls" who therewith buy out the "bears", the latter being content, for the time being, with holding more saving deposits at rising security prices'. The situation is described with the help of Table 4.28. Banks agree to make advances to the bulls, who will now have the funds necessary to purchase bonds. Thus, relying again on Bibow (2009, p. 122), we can say that 'the banking system plays a key part in determining financial conditions'.

Indeed, one could argue that Table 4.28 describes the standard case, if we reinterpret the bulls as the specialist traders, because, as argued by Christopher Brown (2003–04, p. 334),

a sell-off of any significant magnitude . . . will force specialist traders to access 'contingent capital', which most often takes the form of pre-negotiated lines of credit with commercial banks. Thus the money supply is likely to expand endogenously in consequence of the intense selling pressures as specialists draw on overdraft privileges to make the market.

Table 4.27 *An accommodated increase in the liquidity preference of households within an overdraft system*

Assets	Households	Government	Central bank	Banks	Σ
HPM	+10 [1]		(-19) [9]	+9 [8]	0
Deposits	+90 [2]			(-90) [5]	0
Bills	-20 [3]			+20 [6]	0
Bonds	-80 [4]			+80 [7]	0
Advances			+19 [10]	(-19) [11]	0
Σ	0	0	0	0	0

Table 4.28 *Bulls get bank advances to purchase the bonds that bears no longer wish to hold*

Assets	Bears	Bulls	Central bank	Banks	Σ
HPM			(-10) [9]	+10 [7]	0
Deposits	+100 [5]			(-100) [6]	0
Bills			+10 [10]	-10 [8]	0
Bonds	-100 [4]	+100 [3]			0
Loans		(-100) [2]		+100 [1]	0
Σ	0	0	0	0	0

Davidson (2009, ch. 6) makes the point that it is precisely the breakdown of these liquidity arrangements that led to chaotic financial markets during the Global Financial Crisis, as market makers were either unable or unwilling to absorb sales. Daniel Neilson (2019, p. 89), an advocate of Perry Mehrling's (2017) *Money View* makes the same point: market liquidity – the relationship between dealers and their customers – requires funding liquidity, that is, the willingness of banks always to abide by their line-of-credit agreements with dealers.

We can also interpret Table 4.28 in the opposite way. It could be that bulls are now wishing to purchase bonds, getting the funds to do so through bank loans. In that case, the transactions described by Table 4.28 will occur only if the prices of bonds rise, as the bears will agree to divest themselves of their bonds only if they make a capital gain, since their portfolio coefficients have not changed. In that situation, the value of bonds and the amount of money in the economy will both rise. This case illustrates clearly how asset price inflation can occur when financial investors have easy access to bank credit (Toporowski, 2000).

4.5.4 The Liquidity Preference of Banks

During our discussion of the public's liquidity preference, we referred to the behaviour of banks. The liquidity preference of banks can be and has been understood in two ways. The first way is analogous to the liquidity preference of households. Le Héron and Mouakil (2008) thus put forward a set of equations similar to those of equation

(4.3), where banks choose between making loans and purchasing corporate paper, bills or bonds. Liquidity preference may also be assessed by some ratios. This was mentioned when pointing out that banks might wish to achieve a range of safe assets to deposits ratio. When discussing Table 4.26, we pointed out that banks may wish to achieve or raise this ratio by raising the interest rate on deposits, thus purchasing bills from investors lulled into bank deposits by their higher yields. The size of the banks' balance sheets would thus rise in this case. An alternative would be for banks to securitize their loans, to get them off their balance sheet, in which case the size of their balance sheet would be reduced.

The liquidity preference of banks can also be interpreted in a second sense, defined by the state of banking entrepreneurs' animal spirits. This is tied to the issue of credit rationing, which Keynes (1930b, p.364) described by saying that 'there is normally a fringe of unsatisfied borrowers' who believe that they should be able to get more loans from banks. Banks with high liquidity preference are reluctant to increase loans or to take on new customers. Hence we can use the concept of liquidity preference in the banking industry to measure the willingness of banks to extend credit to potential borrowers. It represents the confidence banks have about an uncertain future. By this definition, a fall in banks' liquidity preference implies that banks are more willing to grant loans, or are willing to grant them on easier conditions. This second meaning is the focus of the current subsection.

Horizontalism was always associated with creditworthiness and credit restraints

The consideration of banks' liquidity preference and of credit restraints has given rise to another controversy between horizontalist and structuralist post-Keynesian economists, but in my view an unnecessary one. Some post-Keynesian authors have depicted horizontalist banks as 'passive' (Dow, 1996, p.497), or 'mere ciphers' (Cottrell, 1994, p.599). There is now a large body of evidence to show that this never was the case (Rochon, 1999, pp.169–73 and ch. 8). Horizontalists such as Moore and Kaldor have emphasized the elasticity of the supply of credit by banks and their accommodating behaviour ('non-discretionary'), through overdraft arrangements, for instance, rather than the credit-rationing aspects. But this is to be expected since these authors were initially trying to convince their readers that money was endogenous. However, both Kaldor and Moore underlined the possibility of credit restraints and the importance of finding creditworthy borrowers. Indeed, creditworthiness is a crucial feature of money endogeneity, and it ties in with the property and collateral requirements of bank lending as emphasized by Heinsohn and Steiger (2000). Indeed, as Minsky (1986a, p.256) explains, 'the fundamental banking activity is accepting, that is, guaranteeing that some party is creditworthy'. The following quotes demonstrate quite clearly that earlier horizontalist authors paid attention to the issue of credit restraints and creditworthiness:

The banking system fixes a rate (or a set of rates) for the money market and then lends however much borrowers ask for, provided that they can offer satisfactory collaterals. (Le Bourva, 1992, p.449)

Changes in the stock of loans and money will be governed solely by the demand for loans and the *credit-worthiness* of would-be borrowers. (Godley and Cripps, 1983, p.77)

At any one time the volume of bank lending or its rate of expansion is limited only by the availability of credit-worthy borrowers. When trade prospects are good or when the money value of borrowers' assets (collateral) rises as a result of a rise of prices, the demand for bank credit rises but by the same token the credit-worthiness of potential borrowers also improves, so that the demand for and the supply of credit move simultaneously in the same direction. (Kaldor, 1981, p. 15)

Banks are not inclined to approve bank loan applications just because they have excess reserves. They will, in fact, be willing to grant loans only to those who can demonstrate that they are 'credit-worthy', and once this demand for loans has been satisfied, no additional credit is likely to be extended. (Eichner, 1987, p. 854)

Commercial bank loan officers must ensure that loan requests meet the bank's income and asset collateral requirements. They must in general satisfy themselves as to the credit-worthiness of the project and the character of the borrower. It is precisely for these reasons that banks develop client relationships with their borrowers. (Moore, 1988, p. 24)

Banks . . . attempt to meet all customer loan requests up to some prearranged credit ceiling, provided that the latter meet the bank's minimum collateral and risk requirements. (*Ibid.*, p. 57)

Banks establish individual borrowers' credit limits according to their estimates of borrowers' credit worthiness, and then are price setters and quantity takers in their retail loan and deposit markets. Bank borrowers ordinarily do not fully exhaust their credit commitments . . . This is not to deny that many small borrowers are effectively credit-constrained. New businesses and poor households in particular do not possess the income, assets, and credit records criteria that banks require in order to make profitable and financially sound loans (the banks' three C's: credit, collateral and character). (Moore, 1994, p. 123)

How best can we represent the concept of credit rationing? Before we propose what I think is the most enlightening representation, a few preliminary remarks are in order. Nowadays, most businesses hold a credit line with a bank, and financial institutions have credit lines with each other. Many individuals also have automatic access to advances, through their credit cards, their bank account, or with their house as collateral. Lines of credit, which are a contract between a bank and a borrower, play an important role in the initial financing of production, since they provide flexible access to finance. By agreeing to a credit line, firms enter into a contractual agreement with the bank that specifies the maximum amount that can be borrowed when needed, the conditions under which access to the line is given, as well as the rate of interest on the amount being drawn down (Wolfson, 1996).

The rate may be fixed, in which case the borrower pays some fee to guarantee the fixed rate; or the rate may vary, in which case its level will be set as a mark-up over some market rate, say the Treasury-bill rate, or even the target rate of the central bank. This mark-up is a risk premium imposed by banks to cover default risks as well as administrative costs. As a result, the interest rate on credit lines tends to follow the general evolution of market rates, as formalized by Lavoie and Reissl (2019), but of course, each category of borrowers will be facing a different interest rate. Thus a small firm with a low debt to equity ratio is likely to face higher interest rates than a large corporation with a high debt to equity ratio. With otherwise identical firms, banks will necessarily impose a higher interest rate on those firms burdened by a heavier debt load, due to the perceived higher risk. The same kind of rating will occur when firms issue commercial paper or corporate bonds on

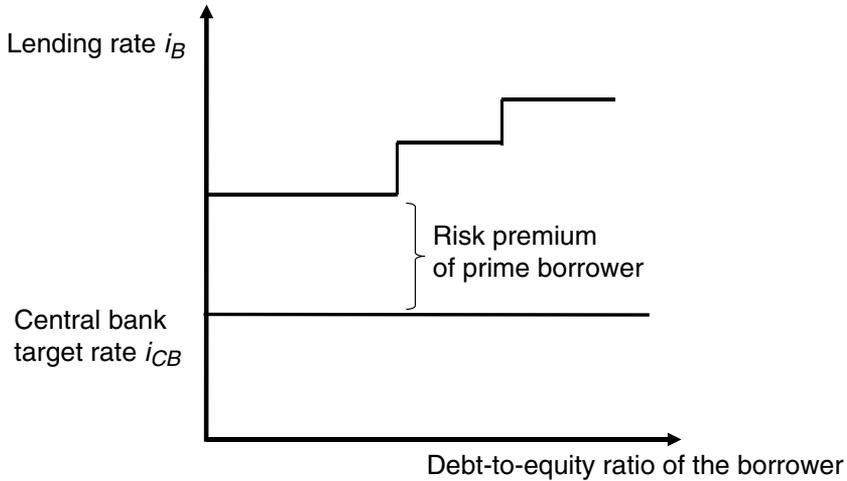


Figure 4.5 Kalecki's principle of increasing risk, understood at the microeconomic level

the financial markets. This instance of Kalecki's principle of increasing risk is illustrated in Figure 4.5. This must be understood as a cross-section of firms, not as an evolution through time.

Credit rationing with notional and effective demand for credit

The main characteristics of credit-line contracts are that they specify the interest rate charged and the maximum amount that can be borrowed, as was argued in a different format with equation (3.3) when discussing the finance frontier of firms. At some point, the banks or the markets will simply refuse to lend any longer. How can this be represented graphically? Some authors suggest drawing an upward-sloping supply curve of credit; others propose a truncated supply curve, or one that suddenly becomes vertical.

I have argued in the past that credit constraints have nothing to do with the slope of the credit supply curve. This was pointed out in Lavoie (1985c, p. 845), where the conditions imposed on borrowers ('profitability, collateral and the like') were considered as 'shift parameters'. Arestis and Eichner (1988, p. 1010) also consider that credit rationing acts as a shift variable in regression analysis. But the appraisal of what a good borrower is must be partly subjective. Borrowers and lenders may have a quite different opinion about this and about how much can be lent to a good borrower. In Lavoie (1992b, pp. 177–8), I refer to the 'effective' demand for credit curve, based on the 'existing collateral and risk requirements for borrowing. When these requirements are modified, say relaxed, they shift upwards the effective demand curve for credit.' In another article, I refer to 'solvent' demand, claiming that 'more exacting norms will shift the solvent demand for credit curve to the left' (Lavoie, 1996d, p. 287). Wolfson (2012) refers to '*creditworthy* demand' within the same context. One thus needs to make a distinction between the demand curve for loans, as perceived by the borrowers, and the demand curve for loans that is judged to be effective, solvent or creditworthy in the eyes of the lenders.

Borrowers who do not meet the banks' criteria simply do not receive credit. As long as borrowers are credible – that is, perceived as being able to reimburse their debt – banks

will agree to lend to them. Hence, as Robinson (1952, p.29) puts it, ‘the amount of advances that banks can make is limited by the demand from good borrowers’. The obvious question then is whether banks can tell who is creditworthy. Banks have developed a number of sophisticated ways to do this. For instance, they will rate borrowers into various risk categories, based on the borrowers’ history, their past relationship with the bank, the value of their collateral, the kind of project to be financed, a number of debt and liquidity ratios, including the borrowers’ cash flow relative to the estimated interest burden. Those borrowers who do not meet the criteria or who decline to fulfil the collateral requirements will be turned down (or rather should be turned down, which was not the case during the years leading to the subprime financial crisis!). Individuals or firms meeting all conditions will be given access to a line of credit that will cover their normal financial needs.

A good example of all this occurred right after the default of the investment bank Lehman Brothers. Firms were suddenly unable to access some of their funds deposited in the shadow banking system and could not finance their expenditures by selling corporate paper, as nobody dared to purchase what was suddenly considered unsafe assets. For a few weeks, several firms suffered from a credit crunch, as banks were reluctant to lend to new customers or to increase lending to customers who had not made adequate credit-line arrangements.

Martin Wolfson (1996), who once worked for the Federal Reserve, has written what I consider to be the best post-Keynesian article on the issue of credit rationing. As pointed out in Wolfson (2012, p. 117), his purpose was ‘to develop a framework to analyse credit rationing that incorporates a horizontal endogenous money supply curve’. For a borrower of a given risk class, Wolfson (1996) draws a horizontal credit-supply curve, at the rate of interest relevant to this risk class. There are then two credit-demand curves, in analogy with the notional and effective labour-demand curves that will be described in Chapter 5. There is a ‘notional’ demand curve, which corresponds to the demand for loans by entrepreneurs, according to their own expectations. There is then another demand curve, the ‘effective’ demand curve, which takes into account only the demand that responds to the conditions and expectations of the bankers. The terms ‘notional’ and ‘effective’ are also endorsed by Gillian Hewitson (1997, p. 132). The horizontal distance, at the existing lending rate, between the notional and the effective demand curves is a measure of the extent of credit rationing. In other words, from the perspective of the bankers, the effective demand curve for credit is the demand curve arising from creditworthy borrowers. Credit supply is demand-led, but subject to the assessment of creditworthiness, which depends on both objective and subjective criteria. We may say that it depends on the liquidity preference of bankers or their animal spirits.

Figure 4.6 illustrates the two types of demand for credit. The total demand for credit, which includes the demand from those who are creditworthy and those who are not, can be called the ‘notional demand’ for credit. Yet, since banks consider only those who meet their creditworthiness criteria, and since they impose limits on the amounts that can be borrowed even by creditworthy borrowers, for all practical purposes the only relevant demand is the ‘effective demand’ for credit. It should be noted that, as the interest rate on loans increases, a widening gap arises between the notional and the effective demand curve. This may happen because of an adverse selection problem – only those who have no intention of reimbursing the loan will agree to pay high interest rates – as New

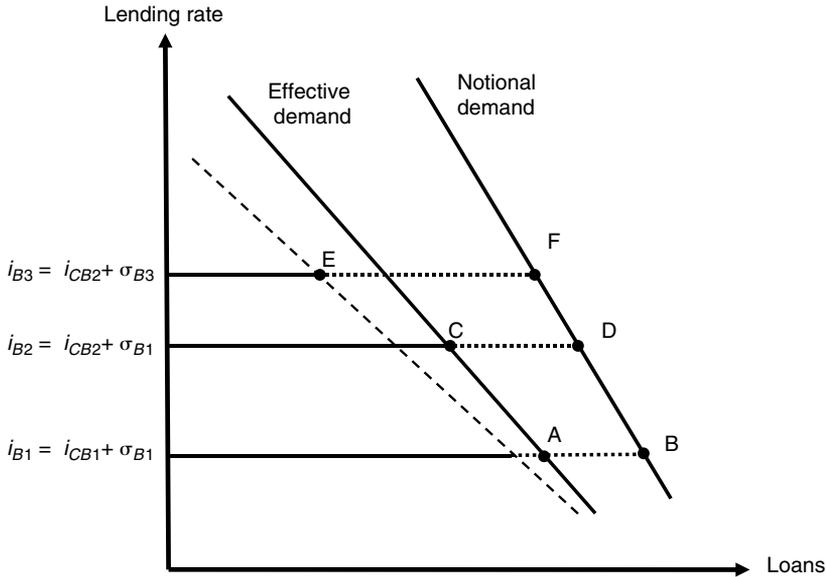


Figure 4.6 Credit rationing with notional and effective demand curves for credit

Keynesians have pointed out; or it may happen simply because the higher interest rate makes it more difficult for the borrower to fulfil his or her debt obligations, a point made earlier by Kalecki (1971, p. 105) when he noted that ‘the higher rate in itself might raise misgivings with regard to the future solvency of the firm’.

As to the rate of interest on loans, we can divide it into two components. As was the case with equation (4.2), there is a component i_{CB} reflecting the benchmark rate of the central bank and hence the prevailing money-market rates (such as the Treasury-bill rate or the overnight rate, or the rate at which banks can borrow on wholesale markets), while the second component reflects the illiquidity and risk premium, σ_B . Hence we can write the following:

$$i_B = i_{CB} + \sigma_B \tag{4.5}$$

We thus see that interest rates on loans can increase from two sources. First, since lending rates are essentially based on the benchmark interest rate set by the central bank, any increase in the benchmark will lead to an increase in lending rates. This is illustrated in Figure 4.6, with the move of the lending rate from i_{B1} to i_{B2} . In this case, fewer firms and fewer households will want to borrow, given the higher costs of borrowing. This is reflected by the negatively sloped notional demand curve. Yet, at the same time, a greater number of borrowers will not meet the banks’ criteria. In other words, fewer borrowers will be creditworthy, and, as a result, a greater number of borrowers will be turned down. Thus credit rationing, measured by the horizontal distance between the two demand curves, will increase from AB to CD.

A different case arises when banks see their confidence falter, as they decide to increase the average spread between the lending and the benchmark rates, thus raising the risk

and illiquidity premium σ_B . This is illustrated in Figure 4.6 by the increase in the lending rates from i_{B2} to i_{B3} . The increase in the risk premium and in the lending rate is then accompanied by a leftward shift of the effective demand curve, which is now the dotted line, as banks are in effect strengthening the criteria needed to get a loan. Credit rationing is now measured by the distance EF. As Wolfson (1996, p.461) puts it, ‘interest-rate spreads, nonprice standards, and credit rationing increase together’. Hence, for every different level of the illiquidity and risk premium, there exists a specific effective demand curve for credit.

In mainstream economics, credit rationing occurs because of asymmetric information. While this may indeed play a role, particularly in interbank lending, post-Keynesians focus instead on the role played by divergent expectations between banks and the consumers or entrepreneurs who wish to obtain loans to finance their activities. When credit markets freeze, it is not because there is a sudden rise in asymmetric information; rather it is because the degree of uncertainty has been rising, generating a loss of confidence and an increase in the liquidity preference of bankers, whose expectations regarding the future are not as rosy as they were. Banks and financial markets will enforce existing criteria or will impose more stringent criteria for lending. As a consequence, a number of borrowers will now be denied a loan: they simply are no longer considered creditworthy. Also, since banks are less confident about the future and expect more defaults on loans, they will want to protect their own rate of return by increasing the illiquidity and risk premium to compensate for the perceived increased risk, which is consistent with Wray’s (1992b, p. 1161) claim that the interest spread depends on the perception of lender’s risk. Indeed, risk premia are lower, not higher, during expansions (Robinson, 1956, p.230).

Of course, we can also look at the other side of the credit market, when risk and liquidity premia decrease and when nearly all the notional demand for credit is fulfilled and is judged to be creditworthy. This corresponds to the case where liquidity seems abundant, meaning that banks and other financial institutions are quite willing to provide credit, at low spreads, and may even encourage households and firms to borrow funds. Still, no one is forced to borrow, but marketing and bank offers may change the preferences of agents, as was argued in Chapter 2. This situation has been well described by Minsky, as will be discussed in the next subsection.

The complete model

How are we to summarize graphically all that we have said so far about interest rates, reserves, bank deposits and bank loans? Structuralists used to favour a continuous upward-sloping supply curve in the reserves and interest plane (Palley, 1994; Pollin, 1996, p. 505), as well as a rising credit supply curve. Horizontalists, as their name indicates, have instead supported a horizontal supply curve of reserves, at the target rate of interest, or, as Kaldor (1983a, p.22) has put it, as a ‘set of horizontal lines, representing different stances of monetary policy’. Fontana (2003; 2009, p.108), as a compromise, proposes a step function that rises with an increasing amount of reserves, meaning that, as more reserves need to be provided, the target interest rate set by the central bank rises in a discontinuous way. This aims to represent the countercyclical behaviour of the central bank and its reaction function. But then it should be clear that this can be the result only of a particular kind of feedback rule. Furthermore, how does that make any sense in countries

where banks do not hold reserves and where banknotes are supplied on demand? It is a bit disconcerting to see that a number of orthodox authors have adopted the horizontal supply of reserves at the target rate of interest as a pedagogical tool while some post-Keynesians are still hesitant to embrace it.

If there now exists a consensus among post-Keynesians, the apparatus shown in Figure 4.7 provides it. This is a four-quadrant figure initially developed by Palley (1996b, p. 588). Nearly the same figure can be found in more recent post-Keynesian works, for example those of Hein (2008, p.45), Fontana (2009, p.114), Fontana and Setterfield (2009b, p.134), Howells (2009, p.178) and Deleidi (2020, p.158). To read the graph, start from the target interest rate set by the central bank, i_{CB} . Through equation (4.5), this determines the interest rate that banks charge on their loans, i_B , once we know what the illiquidity and risk premium σ_B is equal to. Given the effective demand curve for loans, we then know the actual amount of loans in the economy. This amount then helps to determine the stock of bank deposits, depending, however, on the liquidity preference of the public as measured by the slope of the BD line. This slope may also be influenced by a variety of interest rates, but what is at issue is the spread between deposit rates and market rates. Assuming that banks hold reserves, either because they are compulsory or for prudential reasons, we can then derive the amount of reserves in the system.

Howells (2009, p.183) adds an additional quarter in the north-east corner that depicts the relationship between the lending rate and economic activity. One would think, however, that it is the amount of loans, or rather the amount of new loans, that would best predict the level of economic activity or its change. Indeed, forecasting models that

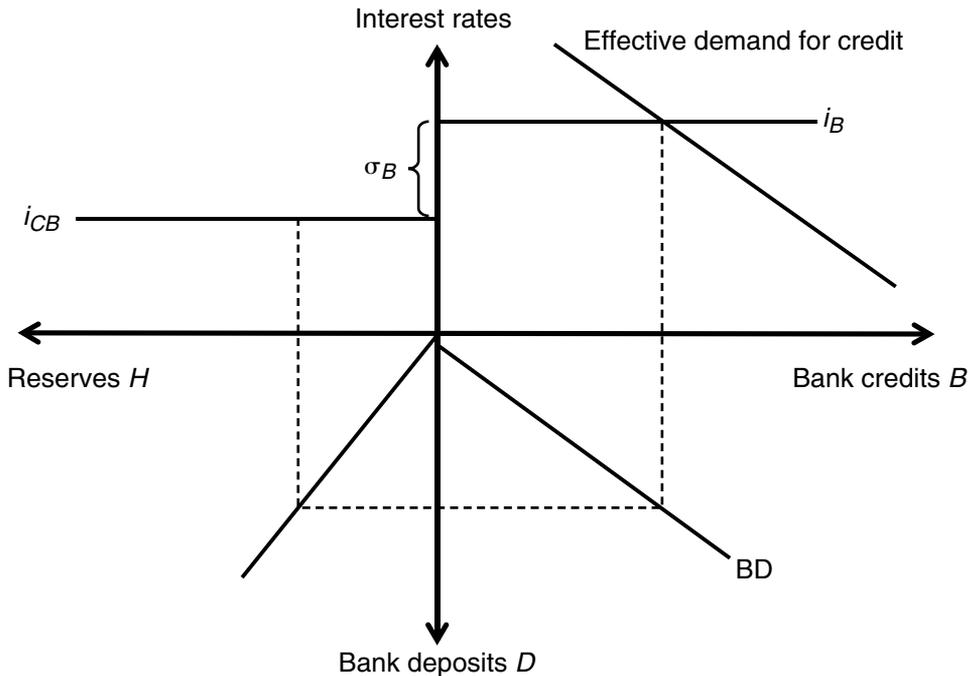


Figure 4.7 The complete model, with loans, deposits and reserves

incorporate the change in household debt and in business debt do well in foretelling economic activity (Godley et al., 2008; Werner, 2012). It should be kept in mind that the effective demand for credit may shift inward not only because banks are less confident, but also because non-financial firms have lost their own animal spirits or have become reluctant to borrow and engage in new projects.

4.5.5 Financial Instability

The introduction of the liquidity preference of banks thus enriches considerably the theory of endogenous money. First, it shows that the focus of our attention should be on the credit market, rather than on the money market. This important point has been made by a number of post-Keynesians, notably Wojnilower (1980) and more recently Peter Howells (2009; 2010) and Steve Keen (2017), because, except when the central bank intervenes when the liquidity preference of the public increases, it is the expansion of bank loans that generates increases in the supply of money. As Eichner (1987, p. 158) puts it, ‘it is the demand for credit rather than the demand for money that is the necessary starting point for analyzing the role played by monetary factors in determining the level of real economic activity’. Second, as pointed out by Dow and Dow (1989, p. 158), ‘an interest rate set by the monetary authorities is consistent with varying rates on bank loans, depending on the state of liquidity preference’. Finally, it is clear that, for a given stance from the central bank and a given interest rate on loans, several levels of loans and credit aggregates are possible, depending on the liquidity preference of banks broadly defined.

Minsky’s thesis

Indeed, financial instability may be attributed to the large shifts in the liquidity preference of banks (and those of other economic agents). We discussed this in some detail in Chapter 1, when introducing the paradoxes of tranquillity and liquidity. These paradoxes and the concept of the liquidity preference of banks are closely related to Minsky’s financial instability hypothesis, also sometimes referred to as the financial fragility hypothesis. Minsky had clearly established the main points of his thesis in the 1960s, before the post-Keynesian resurgence:

The broadest hypothesis is that the behaviour of an economic system with respect to the real variables is not independent of the financial structure of the economy . . . [The second hypothesis] is that the likelihood of a financial crisis occurring is not independent of the financial structure of the economy and the financial structure reflects the ‘past’ of the economy. The third hypothesis . . . is that the financial changes that take place during a sustained boom . . . are such that the domain within which the financial structure is stable is decreased as the boom lengthens . . . If in addition it is assumed that a sustained boom will not be broken by . . . any deficiency in demand, then it follows [the fourth hypothesis] that if a sustained boom is to be broken it must be broken by a financial crisis. (Minsky, 1964, p. 175)

The first hypothesis, that real and financial variables are interdependent, will be taken up in the last section of this chapter. The second hypothesis exemplifies the importance of time in post-Keynesian economics. The third hypothesis is split into two elements. The first element – in my view the crucial contribution of Minsky – restates the paradox of tranquillity, and says that any economy is likely to experience structural changes in its

financial ratios and its balance sheets as time goes by, as a result of financial innovations and changes in the behaviour and standards of agents when faced with a string of good years. This is the essence of the financial fragility hypothesis, and it is the common feature leading to all financial crises.

As the expansion proceeds, a series of successful experiences validates the expectations of the most optimistic actors. ‘Success breeds a disregard of the possibilities of failure’ (Minsky, 1986a, p. 237). Disbelief is suspended as those who warn about impending disasters and plead for more caution are proved wrong time and time again. ‘Stability – even in an expansion – is destabilizing in that more adventuresome financing of investment pays off to the leaders, and others follow’ (Minsky, 1975, p. 127). This is particularly the case within financial institutions where ‘lenders/money managers are “rewarded” for riding speculative waves and indeed are compelled to engage in these activities in order to cement their institutional position’ (Grabel, 1996, p. 22). In an environment of fundamental uncertainty, validation of optimistic expectations leads to a disregard of contrary opinion, as many market participants believe that a new era has arrived and that the old rules no longer apply. This is facilitated by the fact that people, and bankers in particular, tend to forget events long past – a key feature of Minsky’s fragility hypothesis. Financial agents suffer from myopic hindsight.

The second element of the third hypothesis is the financial instability hypothesis as such: it says that balance sheets and financial ratios will deteriorate. This will be modelled in Chapter 6, where it will be seen that the debt ratios of overconfident households are more likely to deteriorate in a boom, whereas the aggregate financial ratios of firms may or may not deteriorate with faster growth, a feature that Minsky (1986a, p. 237) anticipated when he conceded that, at least for a time, ‘internal financing through retained earnings is greater than anticipated, and the push towards a greater use of short-term debt in liability structures is frustrated’.

As to the fourth hypothesis – the advent of a financial crisis – Minsky (1982a, pp. 7, 83, 107) seems to blame the upper turning point on the inevitable tendency of interest rates to rise towards the end of the boom, either as a consequence of the rise in debt ratios (which may or may not happen) or because of the monetary restrictions imposed by the central bank, either as a preventive strike against inflation or as a consequence of actual demand inflation. But then, as pointed out by Toporowski (2005, p. 146), ‘it requires an explicit *policy* decision to raise interest rates’, and we can no longer say that the financial crisis is self-generating or that the turning point is fully endogenous. In any case, once the crisis is there, we may have what the financial press has called ‘a Minsky moment’. The financial crisis becomes a debt-deflation vicious circle of falling asset prices and forced sales, caused by margin calls resulting from the falling value of collateral, unless the government acts with determination.

There are some complications that cannot be easily handled in a simple macro-economic model. In the real world, financial fragility is associated with shorter-duration finance, relaxed banking standards, less intense credit evaluations, more reliance on collateral than on cash flows, the creation of financial innovations to avoid regulation, and a decrease in perceived risk. Aggregates usually deal with net ratios. It should be clear, however, that the robustness and resilience of a financial system are not independent of the gross liabilities or savings of each agent or each sector of the economy. Through leveraged buyouts, firms may accumulate debt in an effort to buy each other, with no increase

in output and profits. As another example, it is quite possible for the net flow of new consumer credit to double or triple, although net household saving and national income stay constant. For instance, it may be that rich households are saving more, while poor ones are getting more into debt. This can be seen by making use of the following two identities:

$$\text{Consumption} + \text{Gross saving} = \text{Household income} + \text{Net additions to consumer's debt}$$

$$\text{Net saving} = \text{Gross saving} - \text{Net additions to consumer's debt}$$

The fragility of the financial system with high consumer credit is much higher than one with low consumer credit. In a period of tranquillity followed by a period of financial euphoria, it is quite possible for gross liability and gross saving to outrun the growth of income and profits, thus leading to a multi-level financial structure – a financial bubble – with a large number of intertwined cash-flow requirements. Minsky has identified the fragile consequences of such layering, in particular the domino effect that it may lead to, as counterparties are unable to deliver funds when they are asked to do so. ‘Liabilities (debts) are issued to finance – or pay for – positions in owned assets . . . In a layered financial structure, the unit acquiring liabilities may have liabilities of its own, and its ability to fulfil its obligations depends upon the cash flow it receives from its assets, i.e., other units’ liabilities’ (Minsky, 1975, p. 87).

Another simple example of the fallacy of net measures is the case of credit default swaps (CDS) – a derivative that emerged in the mid-1990s. CDS were originally designed to hedge against risk: for a small premium, the CDS buyer was purchasing protection against possible losses due to a particular bond default. But in the 2000s, ‘naked’ CDS became fashionable, meaning that investors who did not hold bonds would purchase CDS, gambling that they would default, while CDS sellers – the insurers – were betting that there would be no bond default. Since this is a zero-sum game, it could be said that net risk is zero. But obviously it is not: if there are multiple defaults and the insurer goes bankrupt, the CDS buyer will not get compensation and its problems may spread to other counterparties. Thus, as pointed out in Chapter 1, the existence of protection may actually increase risk rather than diminish it.

The third hypothesis relies crucially on an understanding of financial agents. There are no limits to the amount of credit that can be granted by the banking system. Credit creation depends on the liquidity preference of banks and the confidence of borrowers. Banking is based on trust and confidence, so that if a bank believes that it can be profitable to lend more, it can always do so, as long as it maintains the trust of other banks. The business of banking is essentially a trade-off between the appeal of profits and the fear of losses. Profits are made by lending more, and by innovating in providing new financial products. The fear of losses will gradually vanish in good times, so that banks will encourage potential borrowers to take on loans by offering low interest rates based on low spreads because they assess low default risks. As the boom proceeds, banks will dare to target higher leverage ratios on their equity, and in addition their equity will be growing quickly thanks to the rising profits that are retained.

Hersh Shefrin (2016, p. 133) argues that ‘the global financial crisis provided a dramatic out-of-sample test of Hyman Minsky’s ideas, which he developed based on events he observed between 1960 and 1990’. Shefrin shows that there are many similitudes between

Minsky's insights and the explanations that have been provided by the 2011 report of the Financial Crisis Inquiry Commission which investigated the American 2008 financial crisis.

Alternatives and implications

It should be said, however, that there are alternative post-Keynesian views about the financial behaviour of (non-financial) entrepreneurs. Myron Gordon (1987) argues instead that when entrepreneurs have gone through a series of successful years, they tend to become more prudent, saving more and borrowing less, in order to protect the wealth that they have accumulated during these good years. It is the switch towards this more prudent behaviour, rather than continuing reckless behaviour, that will be at the origin of the downturn. Similarly, Toporowski (2005, p.146) believes that the boom ends as firms become more prudent and start hoarding financial assets, using them as a buffer against a possible future inability to service their debt, and thus diminishing their tangible investments in the process. This kind of behaviour is supported by the prospect theory of new behavioural economics, which presumes that 'an individual's degree of loss aversion increases as her forecast of the size of the potential loss increases' (Frydman and Goldberg, 2014, p. 16).

In the case of the subprime financial crisis, Minsky was certainly right with regard to financial institutions and households, although he himself tended to emphasize the indebtedness of firms, which was not really at stake during that crisis, since much of the borrowing in the private sector was done by households, who ended up as the Ponzi sector that could not service its debt (Brown, 2007; McCulley, 2009). We can also say that the fourth hypothesis came true – interest rates set by the Federal Reserve did rise in the mid-2000s, helping to cause the fall in real-estate prices that precipitated the subprime financial crisis and then the Global Financial Crisis.

Minsky's thesis regarding the transformation of financial structures is very close to that of Keynes. Both described a Wall Street or City view of financial markets, peopled by speculators. Keynes compared the financial markets, and especially the stock market, to a giant casino, where all agents try to outguess each other, and where market evaluations depend on the state of confidence, waves of pessimism and optimism, fads and conventions. Keynes provides the analogy of the beauty contest to show that financial markets are reflexive and prone to mimetism, as participants try to guess the average opinion to assess the overall level of asset prices, since the so-called fundamental determinants are themselves only partly objective (Fung, 2006). Orthodox dissidents also recognize that financial markets exhibit bubbles, fads, overshooting and possibly financial crises when corrections occur, but they blame these on information asymmetry whereas post-Keynesians attribute them to the normal behaviour of financial actors and financial markets.

An obvious conclusion from this is that banks and financial markets need to be closely regulated, as was argued by Minsky (1986a), even though the imposed rules and regulations may not succeed in avoiding bubbles and even crises. While this may restrain the entrepreneurial spirit of bankers, it may slow down the creation of innovations of 'massive destruction', as they have been called. Another reflection is that if lenders are so hard to restrain, then the restrictions ought to be imposed on the borrowers (Minsky, 1975, p.168; Pollin, 1996, p.503; Toporowski, 2000, p.123). This would weaken the ability and the desire of financial institutions to modify their own liquidity preference.

For instance, strict regulations on downpayments required to purchase residential units and be eligible for mortgages have proven to be quite effective in slowing down housing speculation and construction. With respect to financial speculation, tight regulations on margin requirements should also help in avoiding instability.

A further proposal in this direction is that monetary authorities ought to bring back credit controls, which have been known in the past under various names, such as credit ceilings, lending ceilings, the corset or window guidance (Werner, 2005, p. 269). There is a long tradition in post-Keynesian writings in favour of credit controls (Lavoie, 1996e, pp. 540–41). Credit controls have been advocated by Kahn, Le Bourva, Kaldor, Rousseas and Minsky. As argued by Wojnilower (1980, p. 307), credit controls can have ‘startling potency’. Werner (2005, p. 280) also believes that credit controls are ‘the most effective and also the most important tool of monetary policy’. There is a danger, however, that of evasion through disintermediation, moving financial activity from the regulated banking system towards the deregulated financial sector. We discuss this in the next subsection, in the context of the rise of credit aggregates.

4.5.6 The Shadow Banking System

Minsky and shadow banking

Before the advent of the subprime financial crisis, it seemed that financial institutions, particularly in the USA, were ‘awash with liquidity’. This led to what we called the paradox of liquidity in Chapter 1, with liquidity quickly evaporating just when it is most needed, when confidence collapses. This is tightly linked to the views of Minsky, who emphasized the development of near monies in his explanation of the endogeneity of the money supply and of the increase in the fragility of the financial system. These near monies are liabilities issued by what has come to be known as the shadow banking system. Minsky (1975, ch. 4) associates the rise in financial asset prices with the rise in money aggregates and in credit instruments. When banks and the shadow banking system become more optimistic and less wary of possible losses, there is an increased layering of the financial system, more liquidity is created, more loans are granted, and financial asset prices rise. This gives rise to a virtuous circle. The purpose of this short subsection is to explain how this virtuous circle can arise in the modern financial world.

Minsky (1975, p. 90) explains that an increase in the price of the stock market shares of a firm ‘decreases the ratio of cash-payment commitments . . . to the market valuation of the firm. To bankers and other financiers, such increased market valuation implies that the firm can issue more debt – undertake additional commitments.’ The virtuous circle is also helped by the fact that the financial assets, the prices of which have risen, will also generate more collateral, thus helping to purchase further assets and push prices even higher.

This phenomenon also applies to financial institutions, in particular non-bank financial institutions that hold a large fraction of their assets in the form of shares or securities that can be traded on exchanges. When the value of the assets that they hold goes up, the net value of these institutions goes up as well, and they are encouraged to expand their balance sheets. Defining the leverage ratio as the ratio of the value of assets to the difference between assets and debts (excluding equity), as shown for instance in Table 4.8), we can say that, *ceteris paribus*, when financial asset prices rise, the leverage ratio of these financial firms diminishes. Thus, under circumstances of rising financial asset prices,

financial institutions will attempt to increase the size of their balance sheet by looking for new borrowers, and in the case of non-banks, by searching for more funds to finance the acquisition of new liabilities, so as to bring leverage ratios to their previous historical values.

Adrian and Shin (2010) show, however, that financial institutions, and in particular investment banks, go beyond that. They show that leverage ratios tend to rise when their balance sheet expands, and that leverage ratios diminish when balance sheets contract. This means that financial institutions and investment banks in particular take discretionary decisions in an attempt to increase their leverage ratio when the prices of financial assets are booming. The targeted leverage ratios of these financial institutions are thus procyclical, a feature also noted by the BIS for its new Basel III regulations. Minsky (1986a, p. 117) had made exactly this point when, just like Adrian and Shin, he tied the instability of the financial sector to the possibility of upward-sloping demand curves, writing that ‘a rise in the relative prices of some set of financial instruments or capital assets may very well increase the quantity demanded of such financial or capital assets’. Adrian and Shin (2010, p. 428) also point out that the ‘adjustment in the fluctuations of balance sheets is through repos’ – the very short-term collateralized loans that we discussed earlier. But how can the shadow banking system end up creating more liquidity, when we have asserted earlier that it is only official banks that have been granted the power to create money deposits?

Banks versus non-bank financial institutions

Before we can answer this question, let us look in more detail at the distinction between banks and non-bank financial institutions. Bouguelli (2020) points out that the mainstream literature on the shadow banking system identifies three distinctive attributes relative to banks: shadow banks are not regulated, they do not have access to the lending facilities of the central bank, and their liabilities are not insured by the government. Bouguelli contends that these are non-essential distinctions, based on the assumption that banks, like shadow banks, are only financial intermediaries, and that the shadow banking system – now often called the market-based credit system – has supplanted the bank-based credit system, an assumption which can also be found in the *Money View* advocated by Perry Mehrling (2017, p. 4), according to whom ‘today the quintessential form of banking is so-called “shadow banking” – a money market funding of capital market lending’. In what follows, by contrast, I wish to identify the essential distinctive traits of banks.

First, in contrast to other financial institutions, banks can provide new credit without having earlier collected funds or without having to borrow from some other agent. As central bankers McLeay et al. (2014, p. 15) point out, ‘rather than banks receiving deposits when households save and then lending them out, bank lending creates deposits... Indeed, viewing banks simply as intermediaries ignores the fact that, in reality in the modern economy, commercial banks are the creators of deposit money’. Similarly, Jakab and Kumhoff (2015, p. 3) note that ‘in the real world, the key function of banks is the provision of financing, or the creation of new monetary purchasing power through loans... The bank therefore creates its own funding, deposits, in the act of lending’.

The second specific feature of banks is that their deposits are part of the payment system. They are the means through which debts are irrevocably discharged. Non-banks ultimately have to transfer funds to some bank account for the final payment to go

through. Payment is final, or settlement occurs, once the bank payment goes through the books of the central bank, or in some countries through the clearing house run by a bankers' association, as in Canada. As Michell (2017, p. 363) states, this is not the case for non-banks, as they 'fund themselves by issuing liabilities which cannot be used for settlement purposes'. This second feature was underlined by Davidson (1972, p. 157), when he wrote that 'the difference between the liabilities of non-bank financial intermediaries and commercial bank liabilities is that clearing institutions exist for the latter which permit them to be a perfect substitute for legal money'. By contrast, the liabilities of non-banks 'cannot be used in settlement of an obligation' and cannot 'be generally used to discharge a contract'.

It follows that 'as long as banks create credit at the same rate as other banks, and as long as customers are similarly distributed, the mutual claims of banks on each other will be netted out and may well, on balance, cancel each other out. Then, banks can increase credit creation without limit and without "losing any money"' (Werner, 2016, p. 373). This was recognized long ago by Keynes (1930a, p. 26), and then Le Bourva (1992, p. 461), who argued that with the compensation occurring at the clearing house, there would be no limit to the amount of loans that banks could create, provided that these banks 'are all moving in step, with no one bank getting ahead or lagging behind'. For Unger (2016, p. 5), from the Bundesbank, this implies that 'credit expansion in the traditional banking system is not subject to the laws of supply and demand to the same extent as it is for other parts of the financial system'.

This is not the case for non-banks, since, as we shall see, if they wish to create more credit, they must first get either more deposits from the public or new loans from banks. Tobin (1963), when claiming that banks were no different from other financial institutions, argued that banks face the same constraints as other financial institutions, because bank deposits are also subjected to the portfolio decisions of households. In particular, when depositors decide to transfer their bank deposits at accounts in other banks or in non-banks, then the originating bank in its turn will need to borrow funds. The difference, however, is that non-banks need *financing* to start with; when banks grant loans, they create their own finance; if the created deposits leak out, then financing will be needed, but this will occur *afterwards*. This is the distinction made by post-Keynesian circuitists (Botta et al., 2015), inspired in particular by Graziani (2003). Credit creation by banks is associated with *initial finance*; the funds that might be needed to be recovered afterwards are *final finance* or *funding*.

Because of their emphasis on the special role of banks, post-Keynesians are sometimes accused of not realizing that non-banks can also provide credit, so that in a modern monetary economy, bank credit is in competition with other sources of financing. Some past authors have been quite clear about the limitations of non-bank financial institutions. Eichner (1987, p. 825), for instance, has written that: 'It is only through a bank loan that the amount of funds circulating as checkable deposits can increase. If, instead, funds are borrowed from a nonbank financial intermediary, the latter will need to draw down its cash balance at some bank.... This is why the existence of nonbank financial intermediaries can usually be ignored and the flow of funds model simplified by eliminating the nonbank financial sector'.

In view of the rise of the shadow banking system, this quote might be perceived as an overly extreme statement of the post-Keynesian view on this topic. There have always

been post-Keynesian authors taking a more balanced position. Edwin Le Héron (1986), has noted that non-banks could create credit. Palley (1996a, p. 128) made a similar critique, under the guise of the *endogenous theory of finance*, when he contended that both versions of endogenous theory, the horizontalist and the structuralist ones, ‘are flawed because of their exclusive attention to the banking system’. Palley (1996a, p. 133) argued that insofar as ‘the activity of direct capital markets are pro-cyclical’, non-bank financial institutions will also influence economic activity.

These critiques have recently been resurrected by Nersisyan and Dantas (2017, p. 281), who argue that non-banks, just like banks, have the power to change the level of economic activity. Their view is based on an extension of the idea of a hierarchy of money, which can be attributed to Minsky. They claim that banks and non-banks provide liquidity transformation, accepting to convert into liquid debts the less liquid debts of those institutions sitting at the bottom of the monetary hierarchy. They contend that non-bank financial institutions are ‘liquidity creators’, the activities of which ‘affect the real economy’, and thus ‘they can be a source of instability’. Thus, for these two authors, the claimed specificity of banks relative to non-bank financial institutions ‘does not capture the elasticity of finance, or the financial fragility that may arise due to the activities of these institutions The ability of the financial sector to create liquidity is much more elastic than the endogenous money theory allows for’ (Nersisyan and Dantas, 2017, pp. 297–8). Similar arguments are made by advocates of *critical macro-finance*, who argue that the definition of money must be extended to include repos, and possibly the shares of money market funds and asset-backed commercial paper (Gabor, 2020; Murau and Pforr, 2020).

While one can readily admit that non-bank financial institutions have the power to affect the stability of the economy, it still seems appropriate to consider banks as very special institutions, different from other financial institutions. Bouguelli (2018, p. 653), in a comment on Nersisyan and Dantas (2017), argues that ‘making a sharp distinction between commercial banks and other financial institutions’ provides a ‘framework that has the advantage of clarity’. Bouguelli (2020) also provides a table where he identifies how mainstream and critical macro-finance advocates on one hand, and post-Keynesian authors on the other hand, differ in their understanding of shadow banking. Bonizzi and Kaltenbrunner (2020) concur with Bouguelli, arguing that the special role of banks in determining credit conditions is being underplayed by advocates of critical macro-finance and of Mehrling’s *Money View*, who believe that shadow banking is now the centrally important channel of credit. In the next subsection, I try to illustrate the usefulness of this distinction with the use of several examples, once more with the help of T-accounts. The main objective is to understand how non-bank financial institutions can originate credit of their own accord, while still being dependent on the traditional banking system.

Examples of credit creation and securitization

In the first example, a non-bank financial institution (NBFI), in this case an investment bank, desires to lengthen its balance sheet. How can it do it? The simplest solution is to draw on its credit line with a commercial bank, as shown on line 1 of Table 4.29. This then allows the investment bank to provide credit to a non-financial agent, as shown on line 2, by purchasing newly-issued securities. There have been two sources of credit, one from the bank loans and the other because of the bonds acquired by the investment

bank. Note that the commercial bank could have just as well purchased the security, avoiding an additional layer of debt. Note also the investment bank needed first to obtain a loan from the commercial bank.

In the second case, we assume that a non-bank issues money-market funds (MMF) deposits that are desired by some wealth holders. Start with a bank that makes a loan to a non-financial agent who desires to acquire some liquid assets (bank deposits) for future use. This is the first line of Table 4.30. Assume now that the depositor transfers 20 units to a non-bank financial intermediary by purchasing shares in a MMF (we could have assumed just as well that the non-bank is a finance company that issues commercial paper bought by the non-financial agent). This implies that the non-bank now has 20 units of deposits at the commercial bank, which it may decide to transform into a certificate of deposit (CD). This occurs in line 2 of the table. The non-bank is now in a position to provide credit to some other non-financial agent, for instance by purchasing securities worth 15 units, thus reducing its certificate of deposit at the bank by 15. Line 3 shows the consequences of such a move. Whereas we started with bank loans of 100 units, we now have 115 units of credit: 100 units of bank loans and 15 units of sold securities. As to the non-financial sector, it now holds 95 units in bank deposits and 20 units as money-market shares for a total of 115 units of liquid assets, while the non-bank financial sector is left with 5 units of CDs.

To sum up: the overall amount of credit in the economy has risen; the amount of bank deposits has not; and the amount of liquid assets held by the non-financial sector has also risen. Thus, the non-bank financial system has contributed to the creation of liquidity, and has provided new credit, just as Le Héron, Palley or Nersysian and Dantas would have it. Note that, once more, the additional credit could have been provided just as well

Table 4.29 Credit creation by an investment bank

Case	Banks		NBFI (Investment bank)		Non-financial agents (NFA)	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1	Loan +100	Deposit of NBFI +100	Deposits +100	Loan +100		
2	Loan +100	Deposit of NFA +100	Securities +100	Loan +100	Deposit +100	Securities +100

Table 4.30 Credit creation by a money-market fund

	Banks		NBFI (money market fund)		Non-financial agents	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1	Loan +100	Deposit +100			Deposit +100	Loan +100
2	Loan +100	Deposit +80 CD +20	CD +20	MMF share +20	Deposit +80 MMF +20	Loan +100
3	Loan +100	Deposit +95 CD +5	CD +5 Security +15	MMF share +20	Deposit +95 MMF +20	Loan +100 Security +15

by the banking sector. Note further that the non-bank financial institutions would have been unable to provide any credit unless non-financial agents had previously transferred some of their bank deposits to the non-banks. Still, it is clear that in this second case the additional credit has originated from the non-bank financial institution, and not from a bank. A counterpoint, however, is that the funds acquired by the non-banks have arisen initially from some previous loan granted by a bank.

We move on to a third case, tied to securitization, as described by Unger (2016). Here, start with a bank that grants a mortgage to some household; this is the first row of Table 4.31. In row 2, it is assumed that the real-estate builder is paid and acquires the deposits that were initially in the hands of the purchaser of the residence. A public institution or semi-public institution, such as a government-sponsored enterprise (GSE) in the United States or the Canadian Mortgage and Housing Corporation, issues commercial paper (CP) which is bought by real estate builders who swap part (say 70 per cent) of their bank deposits for commercial paper. In row 3, with the proceeds, the GSE is now in a position to purchase the mortgage loans and ease the maturity position of the commercial banks. Finally, ‘the loans are bundled together to form an asset-based security (ABS) which are then retained on the GSE’s balance sheet’ (Unger, 2016, p. 7). In Table 4.31, these are described as mortgage-based securities (MBS).

In the case described by Table 4.31, loan origination (initial finance) is performed by banks, not by non-banks. However, funding (final finance), as it will appear at the end of the process, is now partly in the realm of the *market financial system*. It seems that banks are losing out to non-banks: this is true from the standpoint of *stocks*, but it is not when considering *flows*: the flow of credit does indeed originate from banks.

We can look at another case of securitization, as described by Table 4.32. Start again with mortgage loans being granted by a bank to some households. The deposits so created end up in the bank account of the real-estate constructor. In the second step, shown in row 2, 70 per cent of the mortgages are securitized and sold to a non-bank financial institution, this time a bank-sponsored conduit – a Structured Investment Vehicle (SIV). The SIV purchases the mortgage-based securities (MBS) by getting a temporary loan from its sponsoring bank. In addition, in the last step, shown in row 3, the conduit manages to sell 60 units of asset-backed commercial paper (ABCP) to real-estate builders, whose bank deposits then fall down to 40 units.

In this case of securitization, it is clear once again that the loan originates from the banks; furthermore, the conduit needs to have access to bank loans to handle

Table 4.31 *Securitization with government-sponsored enterprises (GSE)*

	Banks		NBFI (GSE)		Non-financial agents (NFA)	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1	Mortgage +100	Deposit +100			Deposit +100	Mortgage +100
2	Mortgage +100	NFA Deposit +30 GSE Deposit +70	Deposit +70	CP +70	Deposit +30 CP +70	Mortgage +100
3	Mortgage +30	NFA Deposit +30	MBS +70	CP +70	Deposit +30 CP +70	Mortgage +100

Table 4.32 *Securitization with a conduit*

	Banks		NBFI (SIV)		Non-financial agents (NFA)	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1	Mortgage +100	Deposit +100			Deposit +100	Mortgage +100
2	Mortgage +30 Loan to SIV +70	NFA Deposit +100	MBS +70	Loan +70	Deposit +100	Mortgage +100
3	Mortgage +30 Loan to SIV +10	NFA Deposit +40	MBS +70	ABCP +60 Loan +10	Deposit +40 ABCP +60	Mortgage +100

discrepancies between its purchases of mortgage-based securities and its sales of asset-backed commercial paper. For Unger (2016, p.14), in agreement with what seems to be the consensus post-Keynesian position on this issue, ‘the largest part of the shadow banking system enters the credit intermediation process only after the loans to the ultimate borrowers and the means of payment to finance them have already been created’. Thus, for Sissoko (2017), the expansion of securitization and the quasi-monetary liquidity of repos and corporate paper is only made possible through the credit enhancement and guarantees provided by commercial banks.

This is the *originate-and-distribute* model of banking, which makes the chain from the initial borrower to the ultimate fund holder longer, more complex and more opaque. It is also likely to make the system more fragile. While even some heterodox authors thought that securitization would diminish the overall risk of credit (Aglietta, 1996, p.572), Minsky (1986b, p. 14) by contrast feared that ‘fraud and incompetence will intrude in the long chain of securitization’, as risk is being passed along by the originator to agents who may lack a proper understanding of the characteristics of the financial asset, with all the risks of contagion that this entailed. Securitization enhanced the globalization of financial markets and thus contributed to the lack of appreciation of the risks involved. The new originate and distribute banking model allowed lenders to avoid, at least in the short run or even medium run, the destructive consequences of the decline in underwriting standards.

We conclude this subsection by coming back to the question that motivated its existence: is the shadow banking system conducive to more credit? When Palley (1996a, p. 133) highlighted endogenous finance and securitization, he argued that ‘raising finance in capital markets ... is more expansionary [than bank loans] because it by-passes the monetary constraint imposed by reserve requirements By taking transactions out of the banking system, this reduces the need for bank services and helps circumvent emerging liquidity shortage within the banking system’. Today, this regulatory arbitrage argument is based on capital adequacy requirements. When discussing the BIS rules earlier in the chapter, we already contended that the main purpose of securitization was instead to raise returns on bank equity, and not necessarily to provide more credit. Nonetheless, it can be conceded that insofar as ‘non-bank financial intermediaries are able to obtain additional bank credit which could not be made directly available to entrepreneurs’ through bank loans (Davidson, 1972, p.183), the presence of securitization and of the shadow banking system may indeed lead to increased credit. But this may be due to degrading lending standards more than anything else.

Repos and the shadow banking system

We discussed earlier, with the help of Table 4.29, how an investment bank could decide to lengthen its balance sheet. Besides obtaining a loan from a commercial bank, the investment bank could also engage itself in a repo operation. Suppose the investment bank feels that its leverage ratio is too low, either because the prices of its assets have gone up or because it is willing to take more risk. It will be on the lookout for both investors and borrowers. One way to attract funds is to offer reverse repos, collateralized by its holdings of government bonds, to a non-financial firm that holds deposits at a bank, perhaps by offering a higher rate of return than that on deposits. This is shown in the first line of Table 4.33. The non-financial agent, say General Motors (GM), now has less bank deposits but more of a highly liquid asset, the repo. It feels as liquid as before, while the investment bank now has cash – deposits held at the commercial bank. Additional liquidity has thus been created, without the involvement of commercial banks.

The next step is for the investment bank to decide what it will do with its new money balances. Row 2a shows one possibility. The investment bank could purchase a security that was issued by another non-financial agent, say IBM. Credit aggregates have thus risen, again without the apparent involvement of the banking sector. Row 2b shows another possibility: the investment bank could instead acquire mortgages initially granted by banks. In this case, there is no net credit creation: what has been gained by the investment bank has been lost by the commercial bank.

One is tempted to deduce from the above table that these repo operations make murkier the distinction between *bona fide* banks and non-bank financial intermediaries. We are left with two alternative views of the shadow banking system. The first view, which is the mainstream view, is that financial markets and economic growth would collapse without the presence of market-based lending based on repos, securitization and the originate and distribute model. The second view is that market-based lending is more dangerous than bank-based lending and depends on the latter, and hence that monetary authorities should promote a return to bank-based lending. We already outlined the fact that securitization tends to induce lenders to be less rigorous in their

Table 4.33 Extending credit through repos

	Banks		NBFI (Investment bank)		Non-financial agents (NFA)	
	Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
1		NFA Deposit –100	Deposit +100	Repo +100	Deposit –100 Repo +100	
2a		GM Deposit –100 IBM Deposit +100	Security +100	Repo +100	GM Repo +100	IBM Security +100
2b	Mortgage –100	NFA Deposit –100	MBS +100	Repo +100	Deposit –100 Repo +100	

analysis of creditworthiness. The use of repos has also several drawbacks. While repos offer safety to the repo lenders, as we pointed out earlier since the lender holds and owns the collateral for the duration of the repo, borrowers bear an enormous risk (Gabor, 2016; Sissoko, 2019).

The repo, as explained earlier, is like a time-limited loan, based on the value of a collateral, usually a security. Before the 2008 financial crisis, these securities could be private asset-based securities, thus creating a linkage between securitization and repos, but now they are mostly safe government-issued securities, often long-term ones. What we have not yet pointed out is that the loan is provided with a haircut, meaning for instance that, notwithstanding the interest rate charged, the borrower will only obtain a loan of 95 units in exchange for a security worth 100. Because the collateral is based on daily market-to-market values, this means that if the value of the security falls below 100, the borrower will be subjected to a collateral call and being asked to provide additional collateral. If the call is not met, the lender sells the security, settles the loan, and sends the remaining proceeds to the borrower. In this sense, a properly arranged repo contract is safer for the lender than an uninsured bank deposit.

The risk of changing circumstances, however, is entirely born by the borrowers, who are always subjected to the possibility of forced selling. In addition, the repo may not be rolled over, or it may be offered to be rolled over based on considerably larger haircuts. Such events lead to collapsing financial markets, as financial institutions, one after the other, are unable to face their settlement constraints. The collateral then needs to be sold, its price declines, leading to further increases in haircuts, with the repo market whirling into a death spiral. Sissoko (2017; 2019) argues that contemporary repos are similar to the margin-call bank loans that were popular in the 1920s and which precipitated the Great Depression crisis of the 1930s. Market-based lending based on repos is thus pro-cyclical. When trust and confidence vanish, holders of large financial assets move away from the shadow financial system. Just as easily as liquidity was created, liquidity vanishes, thus creating problems for the shadow banks as well as those institutions that relied on the shadow banking system to obtain funding, forcing them back into the fold of commercial banking or requiring the intervention of the central bank. Repo funding looks good from a micro-economic point of view, but it ends up being dangerous at the macroeconomic level.

4.6 A SYSTEMIC VIEW OF THE MONETARY ECONOMY

We pointed out towards the end of the previous section that Minsky emphasized that the real and the financial sides of the economy could not be treated separately if one wanted to have a full picture of how modern economies function. Besides Minsky, Davidson and Eichner are two other post-Keynesian economists who contributed to the development of an integration of the monetary and real sides of the economy. But Godley is probably the one who worked hardest at providing a formalized version of this integration of the real and financial sides. We already had a glimpse of this work when we examined the liquidity preference of the public. Here we look at the overall picture, in what we shall call the stock–flow consistent approach. But before we do so, we examine the aggregate financial balances that have been proposed by Godley and that have illuminated his forecasting work, in particular for the Levy Economics Institute.

4.6.1 Financial Balances: The Fundamental Identity

Around 1974, in his efforts to make better conditional forecasts in the context of what became known as the New Cambridge model, Godley put forward what he called a three-sector financial balances model, which is now known by many as the ‘fundamental identity’. These financial balances, although they are simply a flow-of-funds accounting identity, help to provide some rigour in what can and cannot be said. Once we know the financial position of the private sector, there is a constraint on what the external and government deficits can be. The framework, used extensively by Godley in his conditional forecasts of the US economy, has been adopted by several researchers in universities and in financial institutions, notably by Jan Hatzius at Goldman Sachs, since the late 1990s. The three sectoral balances, or the fundamental identity, are usually written as

$$(S - I) = (G - T) + CAB \quad (4.6)$$

where S stands for the saving of the private sector, I is the non-financial investment of the private sector into tangible fixed capital goods (including dwellings) and inventories, G is government expenditure, T are taxes, and CAB stands for the current account balance. All these variables are understood to be expressed in nominal terms (e.g. in dollars).

In the terminology used by Godley, $(S - I)$ is the net accumulation of financial assets of the domestic private sector ($NAFA$), $(G - T)$ is the government deficit or what used to be called the public sector borrowing requirement ($PSBR$), while CAB is often referred to as the external balance. Godley (1999c, p. 8) interprets equation (4.6) as saying that public deficits and current account surpluses create income and financial assets for the private sector, whereas budget surpluses and current account deficits do the opposite. Of course, we could also rewrite equation (4.6) in a more neutral, standard national accounting, form:

$$(S - I) + (T - G) - CAB = 0 \quad (4.7)$$

In words, using national accounting expressions, equation (4.7) says:

$$\text{Domestic private net lending} + \text{Domestic public net lending} + \text{Foreign net lending} = 0$$

What is saved but not used to purchase new tangible investment goods must have been saved in the form of financial assets. Thus $(S - I)$ can also be interpreted as the ‘net financial saving’ or ‘net financial investment’ of the domestic private sector. This is the amount that is lent to the other two sectors. When the amount is negative, it is the ‘net borrowing’ of the domestic private sector. The domestic private net lending is in fact made up of two components: the additions to financial assets, from which are subtracted the additions to financial liabilities. Thus, if an individual purchases \$100 worth of bonds by borrowing \$100 from the bank, the addition to financial assets is \$100, the addition to financial liabilities is \$100, while net lending and net financial investment ($NAFA$ in Godley’s terms) are zero.

In a similar way, the government surplus $(T - G)$ is the domestic public net lending. Finally, the negative of the current account balance CAB is the current account deficit.

This means that the country has an equivalent financial account surplus (which used to be called a capital account surplus), which is net foreign lending, that is, the net amount that the domestic economy has to borrow from abroad. In other words, if the private and the public domestic sectors cannot fund their own expenditures (both running a deficit), they must borrow funds from the foreign sector. Thus the expression ‘foreign net lending’ means net lending to the domestic economy by non-residents.

Equation (4.6) (or 4.7) is highly useful because it shows clearly the constraints that any economy faces. This is not a matter of opinion. The equation, or rather the fundamental identity, is derived from the national accounts identity. It is a matter of accounting, not economics. Using standard notations, and noting that M now stands for imports, and not machines as it did in earlier chapters, gross domestic product, GDP , is defined as

$$GDP = C + I + G + X - M$$

We move from the gross domestic product to gross national income, GNP , by taking into consideration net foreign income accruing to nationals, FY , so that GNP equals

$$GNP = C + I + G + X - M + FY$$

We subtract taxes T from both sides, and move private expenditures:

$$(GNP - C - T) - I = (G - T) + (X - M + FY)$$

$(GNP - C - T)$, income less consumption and taxes, is simply private domestic saving. The term $(X - M)$ is net exports, or the commercial balance, while $(X - M + FY)$ represents the current account balance. We have thus recovered equation (4.6).

We can further split the private sector into two components, the business sector and the household sector. Doing this from equation (4.7), we obtain a four-sector fundamental identity instead of a three-sector fundamental identity:

$$(S_f - I_f) + (S_h - I_h) + (T - G) - CAB = 0 \quad (4.8)$$

The f and h subscripts refer to firms and households, S_f thus being the gross retained earnings of the firms, while I_h is gross residential investment.

The fundamental identity in the form of equation (4.8) was proposed by Steindl (1982) when he tried to assess the impact of household financial saving. Parguez (1980) used an equation such as (4.8) to find the determinants of the retained earnings of firms. We shall also see in Chapter 5 how Kalecki started from the national accounts to identify the determinants of business profits.

One could also have a five-sector fundamental identity by distinguishing between the financial business sector and the non-financial business sector. This would be important when attempting to distinguish between the accumulation of retained earnings by production firms and banks, as we know that with financialization a large proportion of business profits have ended up in the coffers of financial institutions (Fiebiger, 2016b). We may thus insert S_b as the gross retained earnings of banks (more precisely, the overall

financial sector), while I_b stands for fixed capital investment of the banking sector. Assuming $(T - G) = CAB = 0$, in order to get to the heart of the matter, we have

$$(S_f - I_f) + (S_h - I_h) + (S_b - I_b) = 0 \quad (4.9)$$

It is then clear that if the private non-financial sector invests more than it saves, thus spending more than its income of the current period, it must be a net borrower; and net lending here can come only from the banking sector. The $(S_b - I_b)$ amount tells us whether the retained earnings of the financial sector exceed their expenditures in tangible capital. The financial saving of the financial sector could go into direct loans or into the purchase of financial assets issued by the non-financial business sector. This amount is only indirectly related to the new creation of credit and deposits. For instance, households could decide to put all their current financial saving into bank deposits, so that bank credit to the non-financial business sector would have to be much larger than $(S_b - I_b)$.

Two special cases can be drawn from equation (4.6). The first case is known as the ‘twin deficit’, and is at the heart of the so-called consolidation programmes of the International Monetary Fund (IMF) and at the heart of the old Washington Consensus. If net lending by the private sector is assumed to be nil, then equation (4.6) becomes

$$(G - T) = - CAB \quad (4.10)$$

The IMF asserts that, by restraining government expenditures and imposing austerity programmes, both the government deficit and the external deficit will be solved at once, thus killing two birds with one stone (and the economy as well, post-Keynesians would say!).

The second special case assumes away the external sector, and allows us to focus on the domestic economy. With $CAB = 0$, equation (4.6) can be reduced to

$$(S - I) = (G - T) \quad (4.11)$$

In this case, the government deficit is the mirror image of the private domestic surplus. When considering these balances as a percentage of GDP, if the current account balance is fairly stable, these two balances will move together, without being equal to each other, as happens in several countries, Canada in particular.

At the world level, where by definition the current account balance overall must be nil, this implies that any private financial surplus must be accompanied by a government deficit. MMT post-Keynesians take this as meaning that, if the private sector wishes to accumulate additional safe government assets, as should be the case in a growing economy, the government cannot but run a deficit. Arguing for balanced budgets over the business cycle thus makes no sense in a growing economy, since it means a diminishing proportion of safe assets in the portfolios of the private sector. It is important to understand, however, that households can accumulate new financial assets even if the government is not running a deficit. This can be seen by relying on the fundamental identity given by equation (4.8). Assuming that $(G - T) = CAB = 0$, and assuming in addition that

$I_h = 0$, so that S_h represents both the saving and the net lending of the household sector, we get another special case:

$$S_h = I_f - S_f \tag{4.12}$$

As one would expect, households can have a positive financial balance and accumulate new financial assets even though the government is running a balanced budget, provided business firms have investment expenditures. It should thus be clear that the household sector can accumulate financial assets without the government sector running a deficit. Equation (4.12) can also be read in reverse. It is then clear that, for a given amount of investment, an increase in the saving of the household sector will lead to a decrease in the saving of the business sector (their retained earnings). Thus, as Eichner (1987, p. 831) puts it, ‘should a nonfinancial sector reduce its current outlays as the means it has chosen to increase its net savings, the gross income of some other non-financial sector will necessarily fall by the same amount’. This will be noted again in Chapter 5.

These important clarifications being made, the implication of equation (4.11) in light of the subprime financial crisis is shown in Figure 4.8, inspired from Krugman (2009). Assume that private saving and taxes are positively linked to GDP, while investment and government expenditures are exogenous variables. The private financial balance is thus shown as the upward-sloping line, while the public deficit is represented by its mirror image, the downward-sloping curve. Figure 4.8 assumes that the world economy was at point E before the crisis, at the intersection of the two continuous lines, with a balanced government budget. As a consequence of the financial crisis, the private sector invests less and there is a sudden shift in thriftiness, so that the line representing the private financial

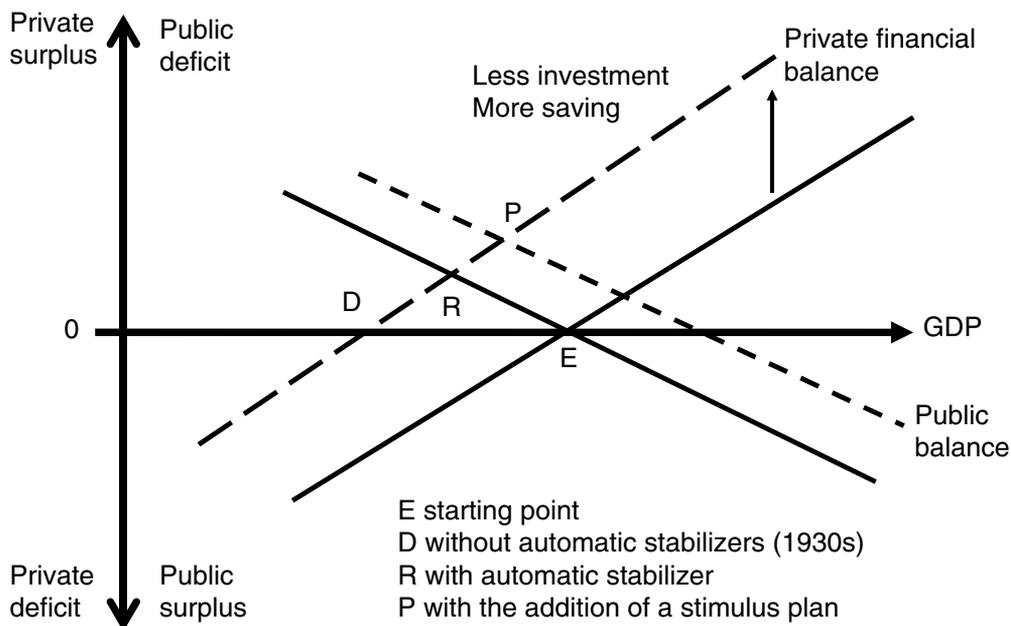


Figure 4.8 A two-sector financial balance model, when the crisis hits

balance shifts up to the dotted upward-sloping line. Letting the automatic stabilizers do their work leads the economy to point R. If deficit hawks had had their way, forbidding any public deficit, the economy would have been propelled to point D, in a state of deep depression, where the public budget is balanced, but with an economy so depressed that financial saving would have been brought back to zero, despite the greater degree of thriftiness – the famous Keynesian paradox of thrift. Finally, with a stimulus plan, that is, a discretionary increase in the government deficit, the economy ends up at point P, at the intersection of the two dotted lines, and in a mild recession. We could also draw a graph representing the fundamental identity in its three-dimensional case, but we shall wait until Chapter 7 to do so.

4.6.2 The Principles of a Systemic Monetary Analysis

While the financial balances presented above have proven to be useful, it is time to discuss monetary transactions from a less aggregated point of view. The French circuit writers have always advocated a mesoeconomic approach, that is, the study of structural – or macroeconomic – laws, independent of the behaviour of agents. Mesoeconomics is found halfway between the macroeconomic emphasis on effective demand and the microeconomic analysis of agents. With respect to technical conditions and relative prices, Leontief's and Sraffa's input–output models can be considered as examples of mesoeconomic analysis. Here we wish to study the interactions of sector-based balance sheets and financial flows.

Over the last 15 years, a number of post-Keynesians have adopted the use of matrices to better explore mesoeconomic relations. Inspired largely by the work of Wynne Godley (1999a) and Lance Taylor (2004), but also by that of Eichner (1987), I believe that such an approach constitutes an important new way of unifying all heterodox macroeconomists, helping to replace and dismiss the neoclassical concept of constrained maximization. This new approach, developed in Godley and Lavoie (2007a), also shares many points of interest with the macroeconomic work of James Tobin (1982), himself a neoclassical synthesis Keynesian. In fact, in his Nobel Prize acceptance speech, Tobin explained why his approach is different from that of the standard neoclassical model. According to Tobin, a proper macroeconomics must entertain four important characteristics, which can be summarized as follows:

1. Stocks and flows must be fully integrated into the analysis, and their accounting must be done in a fully coherent manner.
2. All models should include a multitude of sectors and of assets, each with its own rate of return.
3. It is important to incorporate all monetary and financial operations, and thus integrate the central bank and commercial banks.
4. There cannot be any 'black holes': all flows must inevitably have an origin and a destination; and all budget and portfolio adding-up constraints must be respected, both for behavioural relations and for the actual values of the variables.

Because Tobin was at Yale University, in New Haven, Connecticut, this approach is often referred to as the New Haven School. At the very same time, in the 1970s, a closely

related methodology and approach were independently developed and advocated by Godley and his colleagues in the Department of Applied Economics at the University of Cambridge, leading to the creation of the so-called New Cambridge approach (Godley and Cripps, 1983). What Godley's and Tobin's analyses emphasize is the need for a coherent macroeconomic framework that links together the flow dimension of macroeconomics with the stock dimension of real capital, financial assets and debts, with their corresponding rates of return, while also taking into account the interrelationships and correspondence between the various sectors of the economy. While neoclassical economists have rejected Tobin's approach and have fallen back on the unrealistic 'representative agent', where consumers and producers are one and the same, some post-Keynesians have embraced Tobin's approach, incorporating it, however, into a monetary production economy where the supply of money is endogenous and where behavioural equations respond to Kaleckian or Keynesian precepts rather than neoclassical ones.

This approach has become known as the stock–flow consistent approach, or the SFC approach. 'SFC models are necessarily based on social accounting frameworks that consistently "integrate" conventional product and income accounts with "flows of funds" accounts and a full set of balance sheets' (Dos Santos, 2006, p. 543). Thus, as Taylor (2004) shows, SFC models make use of the social accounting matrices (SAM) of national accounting, which also arose from the work being conducted at the Department of Applied Economics at Cambridge, under the initial leadership of Richard Stone. Indeed, the accounting of the SFC models is inspired by the System of National Accounts of the United Nations, as described in its 1968, 1993 and 2008 versions, whereas orthodox macroeconomics is still based on the outdated 1953 version. Some of the users of the SFC approach hope that SFC models will become the locus of some form of post-Keynesian consensus in macroeconomics, as it allows us to entertain both monetary and real issues within a single model. Nikiforos and Zezza (2017) provide a survey of this rapidly expanding field of research.

In fact, the SFC approach is perhaps a misnomer, as several other theories relate stocks and flows in a consistent way. The peculiarity of the post-Keynesian SFC approach is that its models truly integrate the real and the monetary sides. To get a gist of what this SFC approach truly implied for Godley and his collaborators, we should recall some of the other names that were initially suggested, such as the 'real stock flow monetary model' (Godley, 1993, p. 63), or names that Godley and I juggled with during the 2000s, such as the 'financial stock–flow coherent' approach or the 'sectoral stock–flow coherent' approach, to emphasize that stock–flow consistency is not limited to the trivial link between real investment and tangible capital, but involves most importantly the inter-related financial flows and stocks of assets and liabilities of the main economic sectors of the economy, including those of the all-important banking sector (Dos Santos, 2006, p. 543). These models also attempt to integrate both short-run and long-run dynamics (Macedo e Silva and Dos Santos, 2011).

In attempting to provide a useful model of the economy that deals with both the real and the monetary factors, economists face various difficulties. They have to identify the structural framework that they believe is relevant to the problem at hand. This means that they have to choose the number of sectors that they wish to consider: firms, banks with or without non-bank financial institutions, the government, a central bank, the external sector, households, split or not into two categories, say workers and capitalists. From there,

the modeller must choose what assets and liabilities ought to be included in the model, and whether an asset or a liability can be omitted without much damage from a particular sector. For instance, should banks be assumed to hold long-term bonds or be assumed to issue equities, and should firms be assumed to issue corporate paper or be assumed to hold financial assets such as Treasury bills? These choices lead to the construction of a balance-sheet matrix, which ensures that the model starts out in a coherent way. A proper balance-sheet matrix also helps in designing a proper transactions–flow matrix, which will take into consideration all the financial flows associated with the assumed stocks. The same transaction–flow matrix also ensures that each sector fulfils its budget constraint. To take into account capital gains, a third matrix, a re-evaluation matrix, is also needed.

Finally, appropriate behavioural equations must be inserted within the framework so defined. Different closures of the model will bring about a different choice of endogenous and exogenous variables. For instance, we are likely to suppose that the government deficit is an endogenous variable. But we could also suppose that the government wishes to achieve a given deficit and will do everything to achieve this target, in which case the deficit variable is an exogenous variable, implying that some other exogenous variable will need to become endogenous. This gives rise to two different scenarios in forecasting exercises.

4.6.3 Examples of Balance-sheet and Transactions–Flow Matrices

Besides the revaluation matrix, Godley’s systemic approach rests on two matrices: a balance-sheet matrix and a transactions–flow matrix. Balance sheets deal with stocks, both tangible and financial. Tangible stocks include machines and buildings (the fixed capital of firms), as well as the value of real estate held by households. Tangible stocks also include durable goods such as cars still in circulation. We may also wish to include the inventory stocks of firms (*IN*) – goods that are produced but have yet to be sold. These tangible stocks have no counterpart in the balance sheet.

Financial assets, on the other hand, do have a counterpart, which is debt, and which appears on the liability side of another agent’s or sector’s balance sheet (household, producing firms, banks, the government or the central bank). For instance, loans made by banks are an asset for the bank, but a liability for the borrower. Where do these stocks come from? The answer is that they are either the result of flows, which are added to the existing stocks; or they result from the re-evaluation of certain assets, an exercise that is excluded from the transactions–flow matrix. Each stock is then associated with a given flow through a dynamic equation, which links the past and the present. For instance, the value of all shares held by households at the end of a given year is equal by definition to the stock of shares held at a given time, s , and the price of the shares, p_s , as shown in Table 4.34. This value can be said to arise from three distinct sources: the value of the shares held at the end of the previous year (that is, at the beginning of the year at stake); the value of the new shares issued by firms and purchased by households at market prices during the current year; and the capital gain resulting from the increase in the price of the shares that were held at the beginning of the year during the course of that same year. Long-term bonds are also subject to capital gains.

Table 4.34 shows a simple balance sheet, where inventories have been added to the balance sheet used in the Lavoie and Godley (2001–02) growth model. Obviously, the

Table 4.34 Modified balance-sheet matrix of the Lavoie and Godley (2001–02) model

	Households	Firms	Banks	Σ
Fixed capital		$+K$		$+K$
Inventories		$+IN$		$+IN$
Deposits	$+D_h$		$-D$	0
Loans		$-B_f$	$+B$	0
Shares	$+s \cdot p_s$	$-s \cdot p_s$		0
Balance	$-V_h$	$-V_f$	0	$-(K+IN)$
Σ	0	0	0	0

matrix omits important features of present-day capitalism. As pointed out earlier, all assets appear with a positive sign in the matrix, while liabilities carry a negative sign. The horizontal sum of all assets and liabilities must be zero, except for tangible assets, as explained earlier. The vertical sum for each sector yields the net worth V of the individual sector, which, as was also explained earlier, carries a negative sign as it is part of the liabilities and net worth of the balance sheet. Note that the net worth of firms seen from the perspective of the national accounts is different from net worth seen from the business perspective, as the latter would not consider the value of shares as a liability.

We shall start from the balance sheet of Table 4.34 to show the transactions-flow matrix that corresponds to it. The transactions-flow matrix, here given by Table 4.35, is an interesting tool because it can link all the important aggregates of the National Income and Product Accounts (NIPA) to the financial flows that affect balance sheets. The matrix describes a vertically integrated production economy, dealing with value-added only, as in NIPA, abstracting from the maze of interdependencies associated with intermediate production. To simplify further, we also assume that banks do not make profits (we set the rate of interest on loans, i_B , equal to the rate of interest on bank deposits, i_D).

One of the advantages of using accounting matrices is that they guarantee that nothing gets lost in the discussion: as stated above, all flows must come from somewhere and must end up somewhere. This explains why each line and column adds to zero: the matrix is perfectly balanced. Horizontally, each flow has an equivalent counterpart. Overall, the rows describe the nominal amounts that are exchanged from one sector to another. The equality between the inflow and the outflow arises for one of the following three reasons. First, supply may always adjust itself to demand, either because production adjusts or through changes in inventory stocks; second, demand may be rationed (as in the case of credit rationing); third, market prices may provide for an instantaneous adjustment between supply and demand (as in financial markets). Vertically, each transaction must be financed. The columns sum to zero and represent the budget constraints that each of the sectors must respect.

Let us begin the examination of the transactions-flow matrix given by Table 4.35 by considering households, which face an obvious budget constraint: they receive interest payments ($i_D \cdot D_{(-1)}$), dividends (P_D) and wages (wL), with which they can consume (C), increase their bank deposits (ΔD), or purchase newly issued shares on the financial market ($p_s \Delta s$). In the flow matrix, all components with a negative sign represent a use of

Table 4.35 The transactions-flow matrix in a closed economy without government

Account	Households	Firms		Banks		Σ
		Current	Capital	Current	Capital	
Consumption	$-C$	$+C$				0
Investment		$+I$	$-I$			0
Δ inventory stocks		$+\Delta IN$	$-\Delta IN$			0
Wages	$+wL$	$-wL$				0
Net profits	$+P_D$	$-(P_{ND} + P_D)$	$+P_{ND}$			0
Interest on loans		$-i_B \cdot B_{(-1)}$		$+i_B \cdot B_{(-1)}$		0
Interest on deposits	$+i_D \cdot D_{(-1)}$			$i_D \cdot D_{(-1)}$		0
Δ in loans			$+\Delta B$		$-\Delta B$	0
Δ in deposits	$-\Delta D$				$+\Delta D$	0
Shares	$-p_s \Delta s$		$+p_s \Delta s$			0
Σ	0	0	0	0	0	0

funds, while a positive sign represents a source of funds. For instance, wages, given by wL , which is the product of nominal wages, w , and employment, L , is a source of funds for households. Yet they also represent a use of funds from the point of view of the production sector, and so carry a negative sign when entered in the column of the firms.

The situation of firms is slightly more complicated. While they sell consumption goods (C) to households, they also sell investment goods (I) to each other, as well as finished goods not yet sold to customers, ΔIN . The income generated from these sales, either realized or based on an accounting entry, must always equal the wage and interest payments, plus the net profits from the private sector. These profits can be further divided into two components: dividends P_D to households, and retained earnings, P_{ND} , which is a component of the final financing of fixed capital and inventory stocks, all of these three elements appearing in the capital account of firms.

The middle of the matrix shows the interest payments that need to be paid at the beginning of the period, based on the stocks outstanding at the end of the previous period. The bottom of the matrix represents changes in claims and liabilities. Inasmuch as households increase either their holdings of stock market shares or of bank deposits, this implies a use of funds and therefore a negative sign. But when a firm gets a new bank loan, ΔB , which increases its stock of debt, it becomes a source of funds for the production sector, meaning that it should carry a positive sign. This is usually somewhat confusing to students since the bank deposits of households carry a positive sign in the stock matrix while the stock of loans taken by firms carries a negative sign. Similarly, when firms issue new shares, the proceeds $p_s \Delta s$ appear with a positive sign in the transactions-flow matrix. When firms buy back shares, as many of them do to sustain share prices, a negative number appears in the capital account of firms.

Readers should note that the terminology ‘source and use of funds’ may lead to an additional confusion when discussing the role of banks. When a bank lends and, in doing so, increases the stock of outstanding loans, the additional loans are assigned a negative sign in the transactions-flow matrix. The corresponding deposits that are created as a

result carry a positive sign. In this sense, we can claim that from the point of view of the banking sector, deposits are a ‘source’ of funds whereas loans are the ‘use’ of funds. A note of caution is required here: this may give the impression that deposits are therefore required to make a loan, but this would be a misleading inference. Rather, as we argued earlier in this chapter, loans make deposits. In other words, while the increase in deposits may be seen as a source of funds from a financial perspective, the causal element remains nevertheless the loans initially granted by banks. These loans are created *ex nihilo* at the request of firms considered creditworthy by banks.

The SFC approach presented here is certainly consistent with the integration of the real and financial analysis advocated by some of the key Fundamentalist post-Keynesians. For Minsky (1975, p. 118), for instance, ‘an ultimate reality in a capitalist economy is the set of interrelated balance sheets among the various units. Items in the balance sheet set up cash flows.’ The quote below describes well what an SFC model could or should look like.

It will be the objective of the historical model developed below to provide a simple analysis of capital accumulation by blending the stock and flow elements in the demand and supply of (i) real capital, (ii) money, and (iii) securities . . . with the more familiar concepts . . . of effective demand developed in the *General Theory*. Within such a framework it is possible to provide more perspective on the interplay among organized security exchanges, corporate financing policy, investment underwriters and the banking system in channeling the financial funds necessary for capital accumulation. Regrettably this is an analysis which is virtually ignored in most ‘analytical’ post-keynesian models. (Davidson, 1972, p. 31)

4.6.4 Monetary Creation in a Monetized Production Economy

The monetary circuit and the finance motive

Table 4.1 claimed that a distinct feature of post-Keynesian monetary theory is that it sees money as arising from the production process, since post-Keynesians, as discussed in Chapter 1, are concerned with a monetized production economy. Still, so far, while we have emphasized the role of banks in creating credit and in responding to the liquidity desires of the public, we have not paid a great deal of attention to the relationship between production firms and the banking sector. Yet there is a key relationship between firms and banks, which fund the production activities of non-financial firms. The essential consideration here is time. The production of goods and services takes time, and firms must be in some way able to remunerate their employees and pay their suppliers in advance of sales receipt. While firms generally hold financial assets that allow them to meet their regular monetary obligations, several firms rely on bank loans to start production. This is called ‘initial’ finance. Firms, whether they produce consumption goods in advance of demand or whether they produce investment goods to order, must be able to rely on bank credit. This logic applies equally well to growing or stationary economies, since in the latter case the banks must also agree to roll over debt. The logic of the monetary circuit has been well explained by Godley himself, in his tribute to Augusto Graziani.

In order to finance production, the entrepreneur must obtain the funds necessary to pay his workforce in advance of the sales taking place. Starting from scratch, he must borrow from banks, at the beginning of each production cycle, the sum which is needed in order to pay wages, creating a debt for the entrepreneur and, thereby, an equivalent amount of credit money, which sits initially in the hands of the labour force. Production now takes place and the produced

good is sold at a price which enables the debt to be repaid inclusive of interest, while hopefully generating a surplus – that is, a profit – for the entrepreneur. When the debt is repaid, the money originally created is extinguished. An entire monetary circuit is complete. (Godley, 2004, p. 127)

Post-Keynesians, and in particular circuit theorists, as alluded to earlier, thus make a distinction between ‘initial’ and ‘final’ finance (Graziani, 1984; 2003). This distinction has recently been reasserted by Cesaratto and Di Bucchianico (2020). Initial finance is usually short-term finance, through bank loans and corporate paper. Final finance refers to more long-term market instruments. Davidson (1992) refers to ‘construction finance’ and ‘investment funding’ to distinguish the two steps of the financing process. Firms that purchase capital goods must be able to finance them. Besides their retained earnings, firms must capture the savings of households, either directly or indirectly, through the banking system or other financial institutions such as mutual funds and insurance companies. This final finance phase of the production process closes the monetary circuit of production.

Table 4.35 helps us to make the distinction. The production sector of the transactions–flow matrix is subdivided into a capital account and a current account. Both accounts must sum to zero. The capital account column describes the components of final finance. In this simple model, the accumulation of fixed capital or the increase in inventories of firms at the end of the period, say a quarter or even a year, is financed by only three possible sources: new share issues, new borrowing from banks, or undistributed profits (retained earnings).

By contrast, the grey cells in the matrix represent initial finance. At the beginning of the production process, as the monetary circuit begins, firms must borrow the funds needed to pay wages to employees, and begin the production of new goods, ΔIN (recall that we assume a closed vertically integrated production economy, where the only costs are wage costs). The amount borrowed is exactly equal to total wages in the current period. It is the first stage of the monetary circuit of production. It does not matter whether loans are taken for the production of consumption or investment goods: both in fact require initial finance, as explained by Keynes (1973, xiv, p. 220) when in 1937 he added the ‘finance motive’ to his scheme.

At the very beginning of the circuit, therefore, firms have a debt towards banks but also a claim in the form of a bank deposit. This first stage is in fact very short: firms will draw on their line of credit only when they need to pay wages and begin production. This is done either through cheques or through electronic transfers, as is usually the case nowadays. But as soon as wages are paid, they become an income for households and workers. At the very moment they are paid, and hence before households begin consuming the newly received income, these funds simultaneously become household saving, kept in the form of bank deposits, ΔD . The grey cells depict this second stage. The third stage – final finance – as already pointed out, is represented by the capital account of the production sector, when firms collect back funds from households.

Making use of the grey cells of matrix 4.31 allows us to fall back on some useful accounting principles, especially when considering the operations of firms. As goods are produced but remain unsold, they become an additional component of inventory stocks, ΔIN . Consistent with the best accounting practices, it is important that inventory goods be valued at their current production costs (at their replacement cost, and not at their

expected sale price). In our vertically integrated economy, the cost of production of these stocks is exactly equal to the wages paid to households within this period. In this sense, the value of the increase in inventories, ΔIN , is exactly equal to wL , which is what emerges from the grey cells in the current account column of the firms.

This illustrates the ‘quadruple accounting’ principle that was put forward by Morris Copeland (1949) and is highlighted in the System of National Accounts (SNA) of the United Nations (2009). This principle is also mentioned by Minsky (1996, p. 77). Since each column and each row must sum to zero at all times, any transaction requires at least four recorded changes for the matrix to balance out. For instance, if a bank decides to grant a loan ΔB to the production sector, it must also create, as a counterpart, deposits ΔD of an equal amount. The capital account column of the banking sector then sums to zero. Moreover, the new loan must also be recorded as an additional liability in the capital account of the producing sector. This ensures that the row of loans sums to zero. But a fourth entry must be changed to take into account the fact that the producing sector now holds an additional amount of bank deposits, which allows the row of deposits to sum to zero as well. As a result, we have a minimum of four accounting records for each transaction. In Table 4.35, the operation requires even more entries, because the firms immediately dispose of their deposits by paying wages to their employees, who then deposit the funds in their own bank accounts.

A balance-sheet matrix with shadow banking

The role of the shadow banking sector relative to *bona fide* banks was assessed in the previous section. Quite a bit of effort has been devoted over the last few years to incorporate the shadow banking sector and its various components to stock-flow consistent models. Indeed, initially, Taylor (2008) complained that even the more complex models in the book of Godley and Lavoie (2007a) only handled a plain banking sector. Lavoie (2009d) provided an initial answer, showing how aspects of shadow banking could be introduced within a balance-sheet matrix. Since then, several such balance-sheet matrices have been proposed, and in some cases modelled, such as those found in Nikolaidi (2015), Botta et al. (2015), Sawyer and Passarella (2017), Caverzasi et al. (2019) and Botta et al. (2020).

Table 4.36, taken from Lavoie (2022), presents such an institutionally-rich balance sheet. While there are three tangible assets – inventories, fixed capital and houses h priced at p_h , there is a multitude of financial assets to which households, firms and also financial institutions have access. A key feature is the way in which securitization has been handled, taken from Botta et al. (2015). It is assumed that a proportion z of the mortgages and of other loans granted by banks is being passed on to special purpose vehicles (SPV), who then transform them into asset-based securities (ABS), while the amount of mortgages and loans left on the balance sheets of commercial banks is reduced by the percentage $1-z$. Structured investment vehicles (SIV), who are pure intermediaries, purchase part of these long-term asset-based securities and finance their acquisition by issuing short-term asset-based commercial paper (ABCP), which are themselves purchased by firms and money market funds in search of yields comparable to those of treasury bills.

As discussed earlier, investment banks purchase treasury bills and the securitized assets, essentially through their own funds and repos. These repos are mostly issued by money market funds, who get their funds from households and firms that have large cash

Table 4.36 A balance-sheet matrix incorporating the shadow banking system

	Households	Firms	Commercial banks	Investment banks	SPV	SIV conduits	Money market funds	Central bank	Govt	Σ
Tangible Capital		+ K								+ K
Inventories		+ IN								+ IN
Homes	+ $p_{ir}h_h$		+ HPM_b	+ D_{ib}				- HPM_{cb}	+ HPM_g	+ $p_{ir}h_h$
Cash	+ HPM_h		- D				+ D_{mmf}			0
Deposits	+ D_h						- MMF			0
MM funds	+ MMF_h	+ MMF_f								0
Mortgages	- MO		+ $(1-z_1)MO$		+ z_1MO					0
Loans	- B_h	- B_f	+ $(1-z_2)B$		+ z_2B					0
ABS/CDO				+ ABS_{ib}	- ABS	+ ABS_{siv}				0
ABCP		+ $ABCP_f$				- $ABCP$	+ $ABCP_{mmf}$			0
Repos		- CP	+ R_b	- R_{ib}			+ R_{mmf}			0
Corporate paper			+ CP_b				+ CP_{mmf}			0
Shares	+ $p_s S_h$	- $p_s S_f$	- $p_s S_b$	- $p_s S_{ib}$						0
T-bills		+ TB_f	+ TB_b	+ TB_{ib}			+ TB_{mmf}	+ TB_{cb}	- TB	0
Advances			- A					+ A		0
Net worth	- NW_h	- NW_f	- NW_b	- NW_{ib}	- NW_{cpv}	- NW_{siv}	- NW_{mmf}	- NW_{cb}	- NW_g	- $(K+IN+p_{ir}h_h)$
Σ	0	0	0	0	0	0	0	0	0	0

balances and wish to obtain better rates of return on them. As one can see by looking at their column, money market funds in counterpart hold several different liquid financial assets, such as bank deposits, treasury bills, asset-backed corporate paper, corporate paper, and repos. Finally, there is a central bank, that provides cash to households, reserves and advances to banks, and deposit facilities to the government. Despite the complexity of the matrix relative to that of Table 4.34, there are still components missing, such as hedge funds and long pools of cash such as mutual funds. Still, Table 4.36 helps to see how interconnected are the balance sheets of the various financial institutions.

4.6.5 The Limits of SFC Accounting

Alongside sector-based balance-sheet matrices, sectoral transactions–flow matrices provide the core relationships of monetary production economies. To these, along with the dynamic equations that link stocks and flows, we need to add behavioural equations pertaining to each sector of the economy. Any model offers specific behavioural equations that provide the closure of the model and account for its particular results and derived conclusions. Nevertheless, post-Keynesians like Godley believe that the core accounting equations and the dynamic stock–flow equations actually constitute a framework that constrains the range of possible results. In other words, certain configurations or dynamics are just plain impossible, because they contradict the accounting core. As pointed out by Taylor (2004, p. 2), stock–flow consistent macro modelling and the core accounting equations ‘remove many degrees of freedom from possible configurations of patterns of payments at the macro level’.

According to Godley and Cripps (1983), any fully coherent model that adequately represents a monetary economy of production will necessarily give some medium- and long-term results that are essentially identical, irrespective of the values given to the various parameters. Such a model must be able to provide a linkage between all stocks and flows, while taking all transaction flows into account, including the budget constraints. In addition, stock constraints must be accounted for, such as the portfolio constraints and asset adding-up constraints that we have already discussed. Economic agents target certain ratios that will guarantee the confluence of both stocks and flows, for instance desired sales to inventory stocks ratio in the case of firms. In fact, this may arise without the expressed knowledge of agents. For example, when households consume each year a given ratio of their income and of their wealth, they implicitly define a long-term steady relationship between wealth (a stock) and disposable income (a flow).

Unfortunately, things are not so simple. While it may be relatively easy to agree on the main structural features of a simplified economy, different economists see the behaviour of firms or banks or even households in many different ways. Despite a possible common structural framework that should constrain the range of possible results, as Godley and Cripps (1983) initially hoped when they proposed a first version of the stock–flow consistent approach, it turns out that different economists will still disagree on behavioural equations and the appropriate closure of the model. Thus the results obtained with these different models will differ, as has been confirmed when new SFC models, with assumptions slightly different from those of the earlier ones, produced different trajectories. Moreover, Kappes and Milan (2020) show that, even within the same model, in this case the GROWTH model of Godley and Lavoie (2007a), different weights used in the

adaptative expectations of a single equation generate different transitional paths as well as different steady-state ratios of household wealth to corporate capital stock. This shows that the tested SFC model is subjected to path-dependency, thus underlying once more that time is important in post-Keynesian models.

Notwithstanding the above, models or analyses based on proper accounting and flow-of-funds are more likely to provide useful information than models that ignore the interdependence of the real and the financial sides of the economy. Indeed, Bezemer (2010) argues that those making use of such tools hold a much better record in predicting financial crises and their possible causes and consequences.

NOTE

- * Besides relying on and extending the 1992 *Foundations*, full sentences have been taken from the following publications: 'Money, credit and central banks in post-Keynesian economics', in Eckhard Hein and Engelbert Stockhammer (eds), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2011, pp. 34–60; 'A primer on endogenous credit-money', in L.P. Rochon and S. Rossi (eds), *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, 2003, pp. 506–43.