1. Introduction

Should one ask the layman whether banks and finance are important to understand the working of the economy, the answer would be immediate and straightforward: Yes! In fact, the importance of financial markets and banks for the working of market economies is at the centre stage in popular discussion, as well as in current debates on economic policy. Yet, as it is generally recognized (see, e.g., Gertler, 1988; Goodhart, 2005–2006), for a large part of the 20th century, spanning from the late 1930s to the 1980s, mainstream macroeconomics put banks and the financial system to backstage, or even expelled them completely from its theoretical representations of the economy.¹

In the meantime, the financial system has expanded enormously in its complex interaction with the real side of the economy. Banks too big to fail trade on global markets, marketing innovative financial products; a whole shadow banking industry has emerged, with complex and non-transparent links to the banking industry; financial institutions of various sizes and definitions trade in derivatives or other non-standard contracts involving massive financial flows; firms’ and families’ budgets take advantage of getting credit from the global financial system, whose poor transparency in terms of capital requirements and indebtedness suddenly became so evident since 2007.

1.1 THE DISTURBING PUZZLE OF BANKS AND FINANCE

An explanation of such an evolution of macroeconomic theory could be that, during the period when banks and finance were essentially ignored, the financial side of market economies worked in a relatively smooth way and, thus, economists tended to be concerned with different issues and topics, regarded as more urgent and challenging. Such an explanation, though containing an element of truth, cannot be regarded as fully satisfactory. The theoretical interest, or its vanishing, in the working of the financial side of the

¹ We define mainstream economics as the evolving set of theories, prominent in academic communities at the research and teaching levels, during a certain historical period. For the idea of mainstream economics as a flexible, evolving core of theories, see Colander et al. (2004).
economy and its interactions with the real side cannot be simply related to the occurrence, or absence, of serious disturbances like the Great Depression of the 1930s and the crisis of the late 2000s. Finance and credit always play such a vital role in modern capitalism that they cannot be the concern of economic theory only at critical times. There must be, in our view, deeper and more general reasons why macroeconomics evolved in the way it did. Our book aims to explain such a puzzling evolution.

The dominant families of macroeconomic models in mainstream macroeconomics substantially avoided incorporating banks and finance in their basic analytical structures. As a related and intertwined question, the debt structure within the private sector was also cancelled, either by aggregating the private debts of heterogeneous agents, or by simply assuming that the macroeconomic behaviour of the economy could be effectively represented by models with a single representative agent. It was only in the last 30 years or so that the interest in the study of credit and financial markets and their interrelation with the ‘real’ economy grew significantly. We propose an explanation of all this which is essentially based on the analysis of the way in which banks and financial markets have been conceptualized in mainstream economics, not only in Monetarism and New Classical macroeconomics, but also in Keynesian and Neo-Keynesian macroeconomics (the post-war Keynesianism of the Neoclassical Synthesis).2

Our historical narration reconstructs the state of affairs in contemporary macroeconomics by considering the historical context in which macroeconomists elaborated their theories, the evolution of the research technologies which they explored as effective analytical tools, and finally the vision of the market economy that different scholars had in mind when looking for operational models to be assumed as reference standards. We deal with these aspects by focusing our narration on the evolution of core theories. Given the long-term historical perspective of the book, we cannot deal in any detail with economic history, changing economic policies, or the sociology of research in academic communities. We write a history of ideas, and we focus on selected authors and selected works without any pretence of exhaustiveness. Our aim is to critically explore the threads and issues in the evolution of ideas that we regard as most relevant.

In doing so, we try to provide elements that can help answer some critical questions: do banks and finance really make a fundamental difference for our understanding of macroeconomic events, so that ‘forgetting’ them means missing crucial aspects of phenomena like fluctuations and growth,

---

2 For brevity, here we adopt the conventional label ‘Neoclassical Synthesis’; in Chapters 6 and 7 below, we address in some detail the differences among the various scholars, notably Patinkin, Modigliani and Tobin.
Introduction

3

Macroeconomic stability or macroeconomic policies? Can macroeconomists ignore banks and finance in their modelling and interpretative strategies with only a minor cost to bear? Should macroeconomists stick to models and theories deprived of any explicit reference to banking and finance, or even to money?

By calling attention to these issues, the book aims at restoring a view of financial markets and banks as primary actors in the functioning, or malfunctioning, of market economies, to capture a more realistic vision of the ‘visible hands’ at work, rather than the impersonal operation of ‘invisible’ and impersonal market forces. In monetary market economies, the entrepreneurial activities in banking and finance are among the visible hands, which operate within a complex system of institutional settings to promote the inter-temporal coordination among millions of heterogeneous, independent agents. Financial networks may be exposed to major shocks and be severely disrupted, with the consequence of blocking growth or amplifying business fluctuations. Notwithstanding the recent developments mentioned above, in our view, macroeconomic theory has still a long way to go to reach a satisfactory account of the complexities of financial markets in models and theories which the profession uses in education, in conceptual analysis, and in policy advice.

In the following sections of this chapter, we briefly outline the main issues, topics and economists that will be considered in detail in the successive chapters; but before proceeding it is worth reminding that our historical reconstruction is concerned only with mainstream macroeconomics. Although well aware of the fact that some important contributions to the topics with which the book deals come from economists outside the mainstream, a systematic and detailed consideration of their work is beyond the scope of the book. We limit ourselves to look only at some contributions by Minsky, one of the few non-mainstream economists whose analysis of recurrent episodes of fragility of economic systems with a well developed financial sector has attracted the attention of some mainstream economists especially after the recent crisis.

1.2 FROM THE 19TH CENTURY TO THE 1930s

In the 19th century, scholars addressing questions in the realm of economic theory from the systemic perspective of growth and fluctuations had again and again directed their attention to price adjustments and price instability, the regulation of the money supply, banking policy and the stability or the fragility of the financial system, the sequences of waves of credit expansion or contraction, the fluctuations of expectations and confidence in financial markets.
Thornton’s early analysis of liquidity crises, Tooke’s, J. S. Mill’s or Lord Overstone’s narrations of speculative booms and busts in the so called overtrading theories of recurrent financial crises, the banker John Mills’s psychological explanation of speculative waves, Jevons’s theory of recurrent business cycles, the two Marshall’s interpretation of booms and depressions, are a few prominent interpretations of economic instability. All these scholars evoked speculative bubbles due to imitative expectations, followed by disappointed expectations of capital gains, with collective failures in the rational planning of budget constraints, the domino effects due to bankruptcies and the consequent out-of-equilibrium processes of adjustment. From various perspectives, these scholars analysed how in monetary economies, after some real or monetary shock, including shocks to the quantity of money, out-of-equilibrium adjustment processes take place. Their theories made recourse to observable phenomena of monetary illusion, due to either the outright mistaken expectations about future prices by many private agents or to the misalignment of monetary variables in the dynamic process that leads, eventually, to an adjustment to long-term equilibrium values.

In these various accounts of financial crises and business cycles, banks and their reckless financing of speculative investment had a prominent explicatory role. The overtrading interpretation of financial crises was intertwined with the ‘cycle of credit’ that fuelled immoderate speculation, and later accelerated the financial collapse, because of the credit crunch, the crises of confidence and the panic rush to liquidity that followed the bubble burst. Bankruptcies and their domino effects were an essential aspect in the picture of economic instability, both for their effect in influencing expectations and for their impact on spending and production.

During the 19th century, however, economic thought suffered a kind of schizophrenia: if monetary instability had a prominent role to account for historical events, price theory gave pride of place to equilibrium values in their various descriptions, be they classical or marginalist. The hardcore of economic theory focused on barter economies, where relative prices were transparently set in a non-monetary environment. The emphasis placed on dynamic disequilibrium phenomena in credit and financial markets ambiguously coexisted with the emphasis placed on the stability of ‘natural’ values in classical political economy, and it quite openly clashed with the focus on equilibrium exchange values in the so-called marginalist revolution. In classical thought, out of equilibrium, price theory pointed to the smooth adjustment to long-term equilibrium values, established independently of the

---

3 We use the controversial expression ‘marginalist revolution’ for brevity, with no pretence to discuss here the problem of its continuity or break with respect to classical political economy, or the remarkable differences of approach among ‘marginalist’ scholars.
supply of money, finance and the banking system. Classical growth theory gave exclusive role to real variables. The schizophrenia became even more visible with the emergence of mathematical models of rational, maximizing behaviour within the new paradigm of maximizing equilibrium that various scholars explored in campaigning for a science of political economy grounded in rigorous mathematical language.

At the turn of the 19th century, the gulf between monetary theory and relative price theory appeared wide open to the innovative scholars actively involved in building mathematical models of competitive market equilibrium. The equilibrium analysis of consumers’ and producers’ rational, optimal choices was disconnected from the study of business fluctuations and the monetary economy. Both Jevons and Walras addressed the question, suggesting that it was the inevitable result of the two-stage construction of political economy as a science. Statics, the equilibrium theory of relative prices in a transparent barter economy, was the scientific foundation on which to build business cycle dynamics in the future (Ingrao, 2013, pp. 575–ff.). Since the late 19th century, there was some uneasiness in such promises of future solutions. Wicksell addressed the thorny question as a major topic in his research; in 1898, in Interest and Prices he underlined the dichotomy between equilibrium exchange values and the absolute level of prices. Marshall tried to bridge the gulf between the foundations of economics, dealt with in his Principles of Economics, and monetary theory, by devoting a late separate volume to money and credit (Money, Credit and Commerce, 1923).

Along the years, in the various editions of his major work Eléments d’Économie Politique Pure ou Théorie de la Richesse Sociale, Walras made an effort to coherently include a technology of monetary payments within the equilibrium construction of pure political economy, justifying (and indeed imposing) the demand for cash. He dealt, in particular, with the mathematical expression of the quantity theory of money. Whether he succeeded or whether the equilibrium frame of pure political economy is structurally an ideal barter economy, disconnected from any monetary description of transactions, has been the object of controversy to the present day. Walras devoted considerable attention to monetary questions in his writings on applied political economy.

---

4 On the distinction between statics and dynamics, see Walras (1900, pp. 259–260, 298 and 301–302).
5 ‘The exchange of commodities in itself, and the conditions of production and consumption on which it depends, affect only exchange values or relative prices: they can exert no direct influence whatever on the absolute level of money prices’ (Wicksell, 1898[1936], p. 23).
6 Bridel (1997) carefully scrutinizes Walras’s attempts to put money into his general equilibrium construction; he concludes that they failed. Baranzini (2005) addresses the controversial relation of Walras’s monetary theory in pure economics and in applied economics.
but the coherence of these with his pure theoretical construction is also the object of ongoing controversy.

The uneasiness about the open gap between static equilibrium theory, and monetary theory, the theory of the trade cycle or economic dynamics in general, fully came to light in the early 20th century, when it took front stage at the frontier of research. Since the early 20th century, many scholars signalled it, with reference to Walras’s general equilibrium theory or to other neoclassical theories of relative prices in the context of a transparent barter economy. The question was addressed either from the perspective of how the quantity theory of money could be integrated into equilibrium theory or from the complementary perspective of how to reconcile static price theory with the explanation of dynamic phenomena in monetary economies.

Both perspectives involved new controversies about the quantity theory of money. There was a shared preoccupation with credit expansion or contraction via banks’ loans. Since the early 20th century, and until the 1930s, a number of scholars (among whom were Wicksell, Fisher, Robertson, and Schumpeter) insisted on the role of private banks in creating money if not without limits, certainly within the movable limits set by highly elastic constraints. In the first quarter of the 20th century, macroeconomics as a discipline distinct from microeconomics had not yet emerged. Banks, as institutions that create money through credit, were studied within fluid disciplinary borders, dealing with monetary instability, investment and growth, or the explanation of business cycles. Their ability to create money was the object of concern, the transfer of bank deposits becoming the customary means of payment in the business community. In Europe, banks had been prominent in the financing of investment for development in France, Germany, Italy, and other countries; a highly decentralized banking system characterized the U.S. until the institution of the Federal Reserve system in 1913.

By the mid-1920s, both for dramatic historical reasons (the post-war instability) and for the inner logic of development of the equilibrium paradigm, which was taking the centre-stage in economic theory, the reconciliation of competitive market equilibrium and business cycles in a monetary economy became a major theoretical issue, which the most creative economists of the time addressed as an urgent and unsolved question. Hayek had dealt with it since the late 1920s; in 1933 he argued that the aim to unify coherently equilibrium price theory and dynamics was the task of his generation, since till then little advance had been done to bridge the gulf. From a different

---

7 ‘The most characteristic feature of the work of our generation of economists is probably the general endeavour to apply the methods and results of the pure theory of equilibrium to the elucidation of more complicated “dynamic” phenomena. Perhaps one might have expected all generations of economists to have striven to approach nearer to reality by gradually relaxing the
perspective, Keynes complained that neoclassical theory was the theory of a
counter-crypto economy, and that this prejudiced the understanding of unemployment,
a theme to be found in his writings since the 1920s (Keynes, 1933[1973]). Myrdal wrote *Monetary equilibrium* (1939), published in Swedish in 1931, having in view the task to go further than Wicksell in reducing the gap between monetary theory and equilibrium theory. Hicks had in mind the same purpose when writing *Value and capital*, published in 1939. In the midst of the turbulence following the first World War, in the early 1920s, or later during the contagion of the international financial crisis and the spreading of the depression, a variety of analyses emerged that focused on banks and financial markets in the macroeconomic scene, notably by Hawtrey (1919), Robertson (1926[1949]), Fisher (1922, 1932, 1933), Hayek (1933, 1939), Keynes (1931[1972]d, 1931[1972]a.)

On the distinction between monetary versus real theories of the business
cycles a word of caution is in order. Whether the disturbances to full
employment and market equilibrium arise from real forces in the economy
(such as new investment, productivity shocks in agriculture, innovation and
technological change, or ‘sudden changes in the channels of trade’, as Ricardo
named them) or from monetary disturbances is a question debated since the
19th century. Different scholars from Thornton to Ricardo, from J.S. Mill to
Marx, from Jevons to Wicksell and Marshall, underlined either the primary
role that real forces play or the primary role of monetary disturbances, or
some combination of both. Jevons, who introduced the study of business
cycles based on the statistical analysis of time series in the mid-19th century,
articulated the theory of business cycles as caused by shocks to agricultural
crops due to the periodicity of solar spots. He was a real business cycle theorist
ante litteram, arguing that the ultimate cause of fluctuations are productivity
shocks, with no connection to money or finance; but he never dreamt of
severing the analysis of real business cycle from monetary phenomena, and
he explored how the shocks to agricultural crops finally affected British
manufacturing markets, affecting expectations in the business world and credit
markets. Also the other above-mentioned authors who pointed their
finger towards real factors examined the transmission mechanisms via the links with
the financial sector to explain fluctuations in prices, income and employment.
Often they made recourse to speculative price bubbles to account for recurring
commercial crises.

---

degree of abstraction of pure theory. Yet advance in this direction was not great during the fifty
years preceding say 1920’ (Hayek, 1933[1939], p. 135).

8 Myrdal underlined the difficult coexistence of the theory of general economic equilibrium
with dynamic analysis. He rejected the simplified dichotomy between equilibrium values in real
terms and monetary variables.
According to the various theories, the links between the real and financial sectors were established by considering the expansion or contraction of credit by financial intermediaries, fluctuations in nominal interest rates, divergence between the market rate of interest and the real rate of return on investment, the delay of monetary wages to adjust to inflation or deflation, liquidity crises in the banking industry, the systemic bankruptcies of improvident investors, and so on. In the 20th century the Austrian school underlined how the divergence between the banking rate of interest and the real rate of return on investment put into motion a train of events that involved both real investment and monetary phenomena. Schumpeter, a proponent of real business cycles due to innovative change, did not fail to connect investment to bank credit. Whether the ultimate cause is real or monetary, or a mixture of both, the working of the system of payments, the credit structure, the degree of risk in the balance sheets of households or firms, the solidity of banks and financial institutions, affect the way in which market economies react to shocks. Markets could nurture inside shocks, progressively building up to disequilibria and financial fragility.

1.3 THE KEYNESIAN PARADIGM, THE MONETARIST COUNTER-REVOLUTION AND BEYOND

In the lively and controversial climate of the debates recalled above, since the late 1930s and during the 1940s, macroeconomic theory finally emerged as a separate field of research with the coming to dominance of the Keynesian paradigm. Unfortunately, under the dominance of Keynesianism from the late 1940s to the early 1970s, banks, financial intermediaries and financial markets faded away from macroeconomic models. Commercial banks came to play a merely passive role. Attention was focused on central banks, which were assumed to be able to implement effective policies to control the money supply and stabilize the economy.

The lack of attention to the financial system and its effects on the working of the economy may appear quite paradoxical, given the emphasis that Keynes, in the 1920s and 1930s, laid on the banking system and finance. In reality, however, Keynes himself in *The General Theory* had expunged banks from his analysis of the functioning of the economy; a choice that can be explained by his wish to stress the importance, in an uncertain world, of liquidity preference expressed as demand for (idle) money – something that cannot be easily done when the existence of bank loans make the supply of money endogenous – and his adoption of an equilibrium method, as opposed to the dynamic method of his earlier major work, *A Treatise on Money*.

As a consequence of Keynes’s choice, in the post-war years, also the authoritative Keynesian scholars, who built the theoretical scaffolding of the
so-called ‘Neoclassical Synthesis’, designed macroeconomic models, where
the financial side of the economy is collapsed into the equilibrium equation
between the demand and supply of money, and the money stock is an
exogenous policy variable under the assumption of a stable money multiplier.
The well-known Modigliani-Miller theorem (Modigliani and Miller, 1958)
contributed to such an evolution of macroeconomic analysis as it stated that,
under the hypothesis of perfect markets, the value of a firm is unaffected by
how that firm is financed. In principle, the inside debt structure of firms in the
private sector of the economy appeared to be irrelevant for macroeconomic
stability.

The Keynesian dominance came to its end after the radical critique carried
out by Milton Friedman and other scholars who promoted the so-called
monetarist counter-revolution. Friedman criticized American Keynesians for
having ‘forgotten’ money; but the new monetarist focus on the money stock
did not bring with it a renewed attention to the financial system as a whole.
Friedman pointed to the primary role of erratic monetary shocks affecting
the stability of nominal income, under the assumption that the central bank
could easily control the appropriate money aggregates, and, hence, promote
long-term price stability. The use of simplified operational models, which
Friedman and other monetarist scholars favoured, induced them to look at
macroeconomic models as ‘black boxes’, into which the financial sector could
be put with no long-term impact on the crucial relationship between the two
operational macroeconomic variables, the money stock and nominal income.

Monetarism was followed by New Classical Macroeconomics (NCM), and
inside money and finance were given an even less significant role to play.
During the 1970s, in New Classical Macroeconomics the monetary aspects
were essentially confined to considering the effects of ‘monetary surprises’
caused by unanticipated changes in the money stock controlled by the central
bank (Lucas, 1972, 1981). The explicit choice to conceive of the macro-
economy as a system of markets in full equilibrium, the equilibrium path
being disturbed only by the transitory misconceptions of relative prices due
to monetary surprises, implied the a priori exclusion of financial markets as a
source of disturbances to equilibrium, failed adjustments, systemic collapse or
low-growth traps.

Eventually, money disappeared altogether in the new generation of real
business cycle models (RBC). Monetarism finally turned into ‘Monetarism
without money’, as the point of arrival of such a paradoxical evolution has
been named (Laidler, 2015, p. 19). Since the early 1980s, real business cycle
models shifted their focus from monetary to technological shocks. Money, let
alone financial markets, disappeared from the theory of business cycles, which
were reduced to optimal responses to real shocks by the isolated representative
agent. In one-agent economies, financial intermediation is redundant.
In due course, the ability of RBC models to explain the real world was challenged by the emergence of New Keynesian Economics (NKE), whose salient feature is the importance that market imperfections have for the explanation of the working of the economy. The conflict among NCM, RBC and NKE in the late 1980s gave rise, some years later, to a sort of convergence among macroeconomists on a number of topics and issues, which has been called the ‘New Neoclassical Synthesis’ (Goodfriend and King, 1997). The key features of the new synthesis may be summarized as follows: i) macroeconomics has to be based on rigorous inter-temporal general-equilibrium foundations; ii) the agents’ expectations are rational; iii) there exist imperfections and frictions that are relevant for the working of the economy and they make policies (especially monetary policy) effective; and iv) the most advanced analytical tools are dynamic stochastic general equilibrium (DSGE) models.9

Thus, the most significant New Keynesian contribution to the New Synthesis essentially is the emphasis on imperfections. Initially, NKE mostly concentrated on imperfections in the goods and labour markets; banks and financial markets remained outside the mainstream analytical picture. They were missing in most DSGE models, based on the fictional hypothesis of a single representative agent. Over the years, and especially after the 2007–2008 financial crisis, there has been a growing number of writings concerned with credit and financial markets.

1.4 BANKS, FINANCE AND GENERAL EQUILIBRIUM

In reconstructing the crucial passages briefly outlined above, our focus will be on how different authors and different theoretical strands dealt with two separate, but highly intertwined problems: the nature of banks and other financial firms as active agents in monetary market economies; the visible coordination activities, which make monetary market economies work.

In the first perspective, we shall underline the distinction between banks seen as mere intermediaries between savers and borrowers, and banks seen as institutions that create money through credit or, more in general, on the role of banks and other financial firms as strategic players in contested, non-perfectly competitive markets, where they fight for market spaces through product innovation, innovative approaches to risk management, oligopolistic strategies,

---

9 Another important aspect of the New Neoclassical Synthesis, on which we do not dwell in the book, is that monetary policies are no longer based on the control of the money supply but on ‘inflation targeting’: central banks fix the policy interest rate and, consequently, the supply of money is endogenous (on inflation targeting, see e.g. Bernanke et al., 1999b).
or even sheer fraudulent behaviour. The money supply is endogenous via credit creation by banks, which is not perfectly controlled by the central bank.

Keynesians of the Neoclassical Synthesis still saw banks as able to create money through credit, but in a passive way. The quantity of money in circulation was conceived as essentially determined by the central bank. In fact, it was assumed very often that the quantity of money created by commercial banks was a fixed multiple of the hard money issued by the central bank (the money multiplier). In the 1960s, the prevailing view of banks underwent a further change: they came to be regarded as institutions that are not essentially different from any other financial intermediary (Tobin, 1963).

The monetarist counter-revolution did not bring about any significant change in the conceptualization of the banking system, if not for signalling the systemic risk of banks’ runs intrinsic to a less than 100 per cent reserve banking system. The crucial role of central banks and the passive role of commercial banks were theorized by Friedman, who campaigned to impose 100 per cent reserve requirements in deposit banking. If banks are passive intermediaries, or if specific legislation constrains them to be passive intermediaries, it is reasonable to abstract from their role, when analysing the macro-economy. Essentially this same line was followed by most of the earlier New Keynesian literature. Taking banks into account appears to be only a complication to avoid without any relevant theoretical consequence, especially at the textbook level.

In the macroeconomic literature, it is certainly well known that in a decentralized market society financial markets promote the inter-temporal coordination of spending and saving decisions among different agents. And yet it is not so often recognized that banks are strategic actors in financial markets, and, as a consequence, in market economies at large. Far from being passive, their role as intermediaries between lenders and borrowers is linked with their role as competing innovators. They competitively create the money contracts which they sell, tailoring them to the clients, anticipating the emerging needs on the demand and the supply side. They design, and market, new financial products to capture aggressive investors, to convince prudent lenders, to encourage potential borrowers. The mechanical reading of the money multiplier forgets that the banks generate financial innovation that is a major engine promoting growth, or it may be a major factor of financial fragility and macroeconomic instability, if not accompanied by effective regulations, or sound new practice, to deal with transparency and solvency.

In the second perspective, we shall advance a straightforward critique of the auctioneer metaphor that dominated mainstream macroeconomics because optimal inter-temporal equilibrium was the assumption adopted in modelling strategies, or because these were conceived under the dominant, theoretical influence of neo-Walrasian general equilibrium models. The ambiguous, monetary technology that Walras had tried to build was left aside in the evolution of
the general-equilibrium theory. The image of competitive markets developed into a dichotomous frame, which left a durable imprint in post-war macroeconomic theory. The conception of an equilibrium system of competitive markets turned out to be logically split into the set of equations setting relative prices in goods and services markets, where neither money nor financial intermediation play any role, and the added quantity theory equation needed to set the general price level, and eventually anchor the system to monetary values. If the quantity theory equation appears to solve the problem of integrating monetary values into the pure price theory, the two complementary sets of equations are deeply disconnected with regard to their analytical foundations.

The efforts to integrate monetary theory and general equilibrium theory (the ‘peak of neoclassical thought’ according to Samuelson) clashed again and again with this structural, theoretical difficulty. Notwithstanding the efforts that the best minds in economics had devoted to the task in the years of high theory from the 1920s to the late 1930s, the thorny question was still open in the 1960s. In 1967, Arrow noted that the failed relation between macroeconomics and microeconomics was a scandal in price theory (Arrow, 1967, pp. 734–735). Arrow’s scandal was so much of a challenge that it was worthwhile pointing to it almost fifty years later. The ‘Arrow’s scandal’, as Thomas Sargent called it in 2015 quoting Arrow’s passage, was still relevant for macroeconomics at the opening of the 21st century. It was exceedingly difficult to explain why money should be there, or how it could be put there following a stringent line of research.

Why had the task to unify value theory and monetary theory, which was to be the primary aim of theoretical economists in the 1920s, so dramatically failed in the 1960–1970s? Lucas’s disheartening 2013 summing up acknowledges that in contemporary mainstream macroeconomics, money is not there (Lucas, 2013, pp. xxvi–xxvii). Not only it is not there; it is a puzzle, and a difficult one to solve.10 The difficulty, or more precisely the theoretical impossibility, to include money in neo-Walrasian general equilibrium models is acknowledged by a large specialized literature.11 As Sargent noted in his review of Lucas’s monetary essays, in the neo-Walrasian world money as cash makes no sense, since there are no bilateral exchanges, and thus no need of a generally accepted

---

10 ‘Beliefs aside, a successful policy to deal with a monetary or liquidity crisis will need to be based on some understanding of how real effects of monetary shocks come about, some kind of theory. This has been a central unresolved issue for economists at least since David Hume addressed it in the eighteenth century. This is the theme of my Nobel Lecture, Chapter 16 here, but no resolution is offered in that essay’ (Lucas, 2013, p. xxiv).

11 As for general equilibrium theorists, Debreu explicitly recognized that money is absent from his Theory of Value (Debreu, 1959). Hahn repeatedly discussed the issue with a negative answer; he convincingly argued that money has no place in the Arrow-Debreu model (see, e.g., Hahn, 1965, 1982, 1987).
medium of exchange. In the multi-lateral system of complete spot and forward markets, where transparent transactions are perfectly coordinated ex ante for the whole horizon of exchanges, there is no need for money as a means of payment. The system of payments is mimicked as if it were a perfectly centralized, coordinated system of mutual credits, none of which is at risk of insolvency or default. The fictional auctioneer, calling equilibrium spot and forward prices at infinite speed of adjustment, guarantees that no agent defaults, or fails to accomplish what is stated in contracts. The rationality assumption guarantees that each planned budget is balanced over the relevant horizon. In principle, there are no financial or credit markets: each and every trader is simply adjusting his or her own optimal inter-temporal plan balancing present with future sales and purchases. In perfectly competitive markets, moreover, each and every bond or share, if any are conceivable in these fictional markets, would be perfectly transparent in terms of expected risk and returns to all traders.

There is no need for money even as a unit of account, since every single good or basket of goods may be chosen as numeraire, provided its price is by assumption chosen to be equal to one. In market economies the need of money as a unit of account is related to the convenience in the standardization of computations, to facilitate communication among communities of traders, the collection of taxes or other payments, the setting of price tags in various locations. It helps because of the limited cognitive abilities of traders in processing and comparing lists of prices and values in different units. A fortiori, in the fictional Arrow-Debreu markets there is no need for money to transfer general purchasing power from the present to the future, as the liquidity component of portfolios. Every good is perfectly liquid at the prevailing equilibrium price, once general equilibrium is instantaneously achieved. No risk is run that your neighbour will cheat you by deferring payment, or failing to pay according to commitment; no sanctions are required to impose compliance with signed contracts. If risks arise, they are fully covered by contingent forward exchanges, according to the states of the world.

12 ‘A major source of Arrow’s “scandal” was that in the Arrow-Debreu model of general equilibrium, all trades are multilateral; they are accomplished through a credit system that comprehensively nets out claims. There is no role for cash because there are no bilateral transactions’ (Sargent, 2015, p. 47).

13 In a recent review essay, Lagos et al. (2017, p. 372) motivate the emergence of New Monetarism on the plain evidence that in the Arrow-Debreu model agents do not trade with each other, but ‘they merely slide along budget lines’. They add: ‘money is not essential in standard theories’ (p. 375).

14 The horse that Marshall compared to money to discuss the relative utility of a real good versus an inventory of general purchasing power, such a horse is as liquid as an ounce of gold or a skyscraper in Manhattan.
The absence of money in modelling the economy implies *a fortiori* the absence of finance and credit, apart from the spurious multilateral credit system of neo-Walrasian markets mentioned above. Theoretically, it is conceivable that some intermediaries might supply credit to their debtors in terms of stocks of goods, that is to say, loans in kind, but the essence of a modern credit system is to supply purchasing power that is not constrained in terms of the goods to be acquired with it. There are restrictions on the transactions approved for each credit line or instrument, or open to each specific borrower, such as investment goods, housing or durable consumer goods, if the purchased goods stay as a guarantee to the creditor; but the nature of credit is to supply general purchasing power that borrowers will be free to use according to their best knowledge of their specific circumstances and purposes. Credit and finance are structurally linked to monetary values, and embedded into a monetary economy using a medium which is representative of general purchasing power.

Moreover, since the 1970s, and notably after the radical turn that expelled money from business cycle theory since the 1980s, the research technology most widely adopted in mainstream macroeconomics refrained from going through explicit aggregation procedures. These were bypassed by making the assumption of a strict equivalence between the macroeconomic behaviour of the economy and the behaviour of an ideal, representative household. In real business cycles models, there is not even a barter economy, as we reminded above, and no possible asymmetric information (Mishkin, 2011, pp. 13–4). The research technology based on the representative agent postulates the irrelevance of heterogeneous agents for the dynamic macroeconomic outlook. The cultural roots of this approach date back to Pigou’s welfare economics, and the development that followed through Ramsey’s aggregate model of the national household which optimally allocates saving through time.

In the last quarter of the 20th century a large theoretical literature on business cycles was based on Ramsey’s optimal saving model and the Solow-Swan aggregate growth model, more than on neo-Walrasian general equilibrium models with heterogeneous agents and goods. Those who introduced this shortcut argued that a general equilibrium system with a multitude of consumers is perfectly equivalent to an economy with a single representative household, pretending that such equivalence has been demonstrated. They bypassed all the radical questions that the technical literature on general

---

15 Wicksell contemplated credit in kind in describing an ideal exchange economy. Historically, in the share-cropping system in agriculture, landlords advanced seeds or nourishment to their poor farmers. In the cottage system in pre-industrial revolution times, merchants advanced cotton or wool to cottage weavers. Credit in kind is obviously not the main character of the contemporary banking or financial industries.
equilibrium had raised since the late 1970s on the poor results to be reached within the original Arrow-Debreu model. Their pretence has been criticized in highly technical general-equilibrium literature, but criticisms were ignored or regarded as irrelevant.\(^{16}\)

We cannot deal here with these technical debates, the importance of which, however, must not be underestimated. Whatever the theorems, it is easy to see that the single household assumption erases a priori heterogeneous borrowers and lenders from the macro economy, including the constellation of diverse banks and financial bodies. The rational, single household optimally adjusting inter-temporal choices cannot enter into borrowing and lending with itself; it cannot be exposed to any risk of illiquidity or insolvency; it runs no risk of not complying with its own optimal choices, or not knowing its own patrimonial solidity. The single household moves along its optimal path whatever the exogenous shocks that might hit it. The choice of this research technology was justified by the aim to demonstrate theorems within the sophisticated but still tractable macroeconomic models that it permitted to build. The choice was replete with implications concerning what such simplified lenses allow us to see of the real world around.

1.5 WHY SHOULD BANKS AND FINANCE COME BACK TO MACROECONOMICS?

Since the 18th century, market economies worldwide work within the setting of a fully developed monetary system; financial institutions form an articulated system of markets, regulated both by private contracts and public regulations. In the network of markets which form the skeleton of our economies, financial institutions play a crucial role.

In the historical development of market economies, the monetary system and the network of financial markets have undergone decisive changes in terms of their extension and pervasiveness, and in terms of variety of organization, radical innovation in contracts, technologies of transactions, working practices and public regulations. From the first industrial revolution to the present, the smooth functioning or the maladjustment and crises of market economies have gone hand in hand with radical changes in the financial system. Innovations in financial regulations and/or in the functioning of financial institutions and their accepted business practices have marked epochal changes in the history of market economies.

\(^{16}\) Research on aggregation demonstrates that no one-to-one mapping may be established between a properly general equilibrium economy and a single agent’s economy, unless very strong restrictions are introduced (Kirman, 1992, 2006; Hendry and Muelbauer, 2018).
Banks and specialized firms in financial markets emerged and acted as Schumpeterian entrepreneurs, introducing radical innovations. They opened the way to the diffusion of new means of payments, new loan contracts, new insurances, new ways of raising funds for investment and innovation, new ways of hedging, a new world of transactions and business practices creating opportunities for enhanced coordination, but also for enhanced risks of systemic collapse, due to their becoming deeply ingrained into the current management of the consumers’ and non-financial firms’ balance-sheets. Notwithstanding the recurrent cases of malignant speculation in financial markets, nobody could conceive the development of contemporary market economies all over the world separately from the intertwined development of payment systems, credit contracts, mortgages, insurance policies, hedging funds, stock exchanges, and so on. It is so much so that today it is even problematic to draw a neat dividing line between finance and the real economy when dealing with large corporate businesses.

In the light of the historical experience of capitalist economies, the need for embodying banks and finance into theoretical reasoning and modelling is obvious. The issue, however, requires more elaborated theoretical consideration. The starting point is clear: no auctioneer exists to make the complex system of markets on which our society depends work smoothly. There is no a priori, inter-temporal coordination of people’s decisions in market transactions; no central authority which might collect the appropriate information and impose whatever optimal allocation in real time. The working of markets depends on a set of norms, social practices, and institutions to regulate the production and allocation of resources in decentralized, bilateral transactions taking place in societies undergoing continuous innovation and change. Markets include systems of partial coordination, which require trained staff operating within the constraints of law, norms and shared conventions. Trained staff work daily to smooth arising disequilibria, which generate endogenous innovation and change. Markets work, more or less effectively, thanks to visible persons, visible logistics and communication systems, visible monetary arrangements and financial contracts.

Systems of law, social norms, or shared conventions are the foundations of the trust that permits the encounter in bilateral exchanges of private people or organizations, having in principle no reason to trust each other with respect to their reciprocal commitment to sign contracts and to not violate them. People trading in markets trust and rely on visible conventions which are shared by the other agents too. Money is a social, shared, convention regarding the common standard of value, in which bilateral partners denominate their contracts. Money is no veil; it is the instrument by which millions of transactions are expressed in a common language of value that permits mutual understanding and trust in exchange. The objects traded in financial markets are essentially
Introduction

17

present and future flows of purchasing power denominated in money. By their nature, financial markets deal in contracts defining trades in abstract flows of purchasing power denominated in some currency or package of currencies. These contracts might be backed up by some real goods, or anchored to some real goods, but not necessarily so. In any case, collaterals have to be priced in money values.

In the decentralized markets of contemporary societies no transparent information is fully available a priori, and no inter-temporal equilibrium prevails in complete, forward markets. No central authority covers the risks of each and every contingency over the whole planning horizon. Markets work thanks to the dynamic strategies of busy agents, who take care of partial coordinating activities within changing institutional settings, and under changing conditions. Financial institutions, be they banks or other private agencies (or even government agencies), are among the busy bodies whose strategic investment in human capital, effectiveness in communication and accumulated wealth help the coordination of bilateral monetary transactions in a global world of radical uncertainty and asymmetric information. They exist because information about creditworthiness and risk of default is asymmetric and costly, in the same way as acquiring the proper expertise to evaluate expected returns and risks of investments is time-consuming and costly. They exist because financial intermediation operates within regulations by law and custom that impose high transactions costs, due to the necessity to impose sanctions to protect traders in an environment of asymmetric information, and radical uncertainty governing the future money value of the assets in people’s portfolios.17

Banks and other financial intermediaries work as visible hands to allocate money flows to potential buyers, or to sell investment products to potential savers, coordinating choices not only in financial markets, but at the junction of these with the markets for goods, services, or properties. They may succeed or fail, like other visible bodies and brains at work in market economies. They bet on making choices today that are compatible with uncertain future flows of income, pricing assets whose future prices are volatile, assessing risks, smoothing possibly emerging imbalances, providing for buffers, and so on. They produce innovative change in conventional practices and rules, opening improved opportunities for welfare and growth, or increasing the risks of disequilibria, imbalances, or conflicts.

17 In some societies people could directly trade in bilateral transactions their present wealth against future flows of consumption, or acquire their present consumption against flows of future labour services. The debtor could settle the debt for present consumption by enailing himself/herself for future forced work due to the lender. Against the present payment of a dowry the family could enslave the daughter into a monastery, which promised her lifelong consumption. In our societies these contracts are luckily forbidden.
Banks manage the system of payments, that is, the accounting technology to support legal, bilateral transactions in the commonly accepted means of payment. As accountants, banks monitor the solvency of traders, who pay through the payment instruments having their mark, such as credit cards or deposit transfers. The collapse or the malfunctioning of the system of payments has severe consequences in terms of higher transaction costs and loss of confidence; it implies the more or less extensive paralysis of transactions and, hence, of the market economy. Bilateral transactions require trust in the currency that is exchanged against goods or services, in terms of its purchasing power and liquidity. As the historical experience has proven, the collapse of the system of payments drastically reduces the opportunities for otherwise useful and welfare improving transactions; it drastically shrinks the real economy.

Credit creation by banks accompanies macroeconomic fluctuations through the financing of real investment or consumers’ spending, or through the swelling of price bubbles. Banks inject flows of purchasing power into the balance sheets of borrowers, who channel it into the property markets, the stock exchange, the markets for consumer durables or investment goods, with differential effects on prices or production in the markets where they are channelled. In extending loans, banks perform an allocative function: they evaluate creditworthiness and risks of default. In market environments dominated by imperfect, asymmetric information and volatile asset prices, banks provide the human capital to assess for each liquidity constrained buyer the capability for inter-temporal substitution. Their role in asserting creditworthiness is crucial as regards potential borrowers who, because of the modest size of their business and wealth, cannot provide guarantees to potential lenders; banks assess their ability to incur debt today to be repaid with future income.

Under prudential practices, information gathering and public regulations, banks eventually manage their auctioneer’s job, and help rationed borrowers get their sustainable financing. If the perception of the systemic risk of default

---

18 "The bulk of money is in the form of commercial bank liabilities, and banks can behave very differently over time. The form of their liabilities, their capital base, their confidence and their risk appetite can and does alter over time, both cyclically and more permanently. The whole question of whether certain segments of the economy can access funds beyond their current income depends crucially on the behaviour of the banks. If there is a supply shock to money, with certain groups now getting more, or less, access to funding, for example when banks provide mortgages to a wider group of households on easier terms, will this not feed back into the IS curve? Of course it will" (Goodhart, 2007, p. 58).

19 Notably, they finance the middle or small size firms, which have no access to the stock exchange or alternative channels of finance. By financing their current activities, they help them to balance through time, flows of sales against flows of costs; they help finance their investment expenditures, or their buying of properties. Similarly, they provide purchasing power to households to buy consumer durables or enter mortgage markets.
increases, inducing banks to be selective in lending, their prudential behaviour cuts out from access to expenditure potential borrowers, with effects of credit rationing, or credit crunches. Banks’ role at the junction with real markets is partially blocked; they no more provide buffers to smooth disequilibria. The allocative function requires competent staff and organization, and it cannot be easily substituted for.

If the fictional hand of the auctioneer is removed from the macroeconomic picture, coordination failures evidently show up. Some firms or consumers may become insolvent, or go bankrupt. The relevance of single bankruptcies, and their spillovers into other balance sheets, with real effects on spending, are related to the complex structure of the financial system. The systemic spreading of insolvency among the heterogeneous firms acting in the economy depends on the existing buffer stocks in both the real and the financial sectors of the economy, which are linked to the financial structure with respect to the distribution of debts and financial wealth. The wealth’s basis of balance sheets are money values, which depend on volatile prices of properties, shares and other real or financial assets. Bankruptcies in the banking and financial industry, if and when they happen, create systemic effects in the macro-economy via expectations and confidence, and the reduction in the velocity of circulation of money.

The above are only a few aspects and features of modern sophisticated market economies, but they are sufficient to show how taking account of the financial sector and its interaction with the real sector should necessarily be a fundamental constituent part of an economic theory aiming to understand, and possibly improve, the world in which we live. Other questions, of course, are open and in need of more satisfactory answers. Should the economists’ attention concentrate on banks or other financial institutions? Which is the difference, and which is the relation between them? Which financial flows are provided by bank credit and which by issues at the stock exchange or other funds channelled in financial markets? How is shadow banking related to banking properly?

To accomplish such a task, we argue, requires to go beyond the analytical and methodological strictures that still characterize the current mainstream. More specifically, we need to reject the pretension that an exclusive recourse to models, as sophisticated as they may be, can provide a fully satisfactory understanding of banks and finance in their complex interrelation with the real economy. It is necessary to mobilize a wide range of available cognitive instruments, including historical knowledge, the critical exploration of economic ideas and their evolution over time, the consideration of the complexity of human behaviour as well as the complexity of paths and trajectories that depend on crucial events, institutions and the social and political context. It is the lesson we draw from the great scholars whose ideas we examine and
discuss in the book. ‘Giants’ of the past like Wicksell, Fisher, Schumpeter, Robertson, Keynes, Hicks and Tobin developed their theories, carried out their analyses and formulated their policy prescriptions in the context of wide cultural horizons, where historical explanations, the knowledge and understanding of the evolution of economic concepts played a crucial role together with the elaboration of new views and interpretations.

1.6 PLAN OF THE BOOK

The book is divided into two parts. Part I (Chapters 2 to 6) is concerned with the period spanning from the early years of the 20th century to the years immediately following World War II, when banks and their functioning enticed the attention of many scholars even though, at the same time, there emerged views and positions that can explain the subsequent theoretical developments that led to the fading into ‘oblivion’ of banks and other financial institutions.

Chapter 2 looks at how Wicksell and Fisher, acknowledging the growing importance of banks, took account of them in their attempts to elaborate what they regarded as a more satisfactory quantity theory of money. Chapter 3 is devoted to considering Schumpeter’s and Robertson’s contributions. They were not particularly interested in the analysis of the general price level per se and focused on the role of credit and banks as fundamental factors in capitalist processes of change. The Great Depression of the 1930s was, of course, a major concern of many economists; in Chapter 4 we examine Fisher’s and Keynes’s viewpoints on the effects of the bank and financial crises and, in particular, their deflationary effects, a topic that returned to interest several in the aftermath of the late 2000s crisis.

Finally, Chapters 5 and 6 are devoted to Keynes’s major theoretical contributions in the 1930s and the debates following the publication of The General Theory. More specifically, Chapter 5 looks at the evolving of Keynes’s views of banks from A Treatise to The General Theory. We argue that the well-known fact that, in passing from one book to the other, Keynes let banks virtually disappear, can be explained by two crucial factors: his focus on liquidity preference expressed as a demand for money and his abandonment of a dynamic/sequential approach in favour of an equilibrium method. Part of Chapter 5 and the whole of Chapter 6 are concerned with criticisms of Keynes’s General Theory. We concentrate on criticisms of the liquidity preference theory of the interest rate, to which a modern loanable funds theory was opposed, and the debate on the so-called wealth (or Pigou) effect.

Part II (Chapters 7 to 9) is concerned with macroeconomic theory from the post-war years to the recent developments of the late 20th and early 21st centuries. Since the early debates on The General Theory, one of the claims
Introduction

of the critics of Keynes was that his major results did not stand, or they had to be significantly qualified, when the analysis is carried out by taking account of all the relevant interrelations among markets. In due time, this criticism became a claim for the adoption of a proper Walrasian general-equilibrium framework to deal with macroeconomic phenomena. In Chapter 7, we look at how this problem was tackled by the economists of the Neoclassical Synthesis, though from different perspectives and points of view. We look in particular at the contributions of Patinkin, who was one of the earliest and most important representatives of the general-equilibrium approach to macroeconomics, but who also perceived the difficulties of introducing money into such a framework. Attention is also paid to the innovative approach to money and finance followed by Gurley and Shaw, who significantly influenced Tobin and his attempt to introduce financial markets into a general-equilibrium framework. Tobin’s work was interesting and original, but his research project encountered significant difficulties and remained at the margins of the mainstream.

Chapter 8 deals with the main exponents of the ‘anti-Keynesian counter-revolution’, starting with Friedman’s new monetarism and ending with the Real Business Cycle (RBC) approach. With the partial exception of Friedman, who essentially remained a Marshallian, all the major anti-Keynesian economists of the time emphasized their commitment to develop macroeconomics within an alleged modern Walrasian, or Arrow-Debreu, framework. But they encountered great difficulties in dealing with money, let alone financial markets, within such an environment. These difficulties, however, were underrated or, more prosaically, put under the carpet and the solution to them essentially was to abandoning any pretension to give money and financial markets any significant role to play in standard macroeconomic models.

Both Lucas’s monetary surprises and RBC technological shocks proved largely unable to give a reasonable account of economic fluctuations experienced by actual economies. Chapter 9 explores the developments following these failures, which ended up with the confluence of mainstream macroeconomics to the so-called New Neoclassical Synthesis, with the concurrence of New Keynesian Economics (NKE), characterized by the central role that it gives to market imperfections. Thanks to the abandonment of the hypothesis of perfectly competitive markets, and notably thanks to the emphasis on informational imperfections, there emerged a new generation of macroeconomic models that make some steps forward in including banks or finance in their interaction with the real economy.

However, these models encounter a number of difficulties, which essentially derive from their still close and strong connection with the previously dominant paradigms of the New Classical Macroeconomics. We look at some such difficulties and the issues and questions that still remain open and in need of more satisfactory answers. More in particular, we concentrate on the
difficulties inherent in the very notion of imperfection and those that the analysis of financial and credit markets encounters when carried out by using models that are structured on the coexistence of steady states and exogenous random shocks, relegating economic change and innovation to the sphere of unexplained phenomena and underrating, if not ignoring altogether, the crucial role of the endogenous dynamics of financial markets in the evolution of crises and phases of severe depression.

Finally, the chapter examines some contributions from behavioural economics which, by relying also on psychological research, offer analyses of the functioning of financial markets far from those of the efficient markets hypothesis and closer to Keynes’s approach. Human behaviour, at the individual as well as collective level, cannot be fully explained by the assumption of well-informed agents optimizing over infinite horizons.

The concluding Chapter 10 summarizes the main results of our research and outlines a number of critical issues and topics that require further developments by approaching them in a novel and richer way than has been done so far.