What Makes Banks Special?
A study on banking, finance, and economic development

Biagio Bossone
The World Bank

Summary

Over the last two decades, the theory of finance has offered significant contributions to the understanding of banks, identifying specific aspects that qualify them as special intermediaries. Contributions have differently focused on functions (such as liquidity and payment services, credit supply, and information provision) for which banks have historically developed a comparative advantage vis-à-vis other types of intermediaries. In fact, the evolution of finance through greater market competition, progress in technology, and institutional development, has increasingly eroded the comparative advantage of banks. Not to mention that the spread of e-money might deal a fatal blow to conventional banking, generating entirely new ways of doing finance. By looking at money, production, and investment in an integrated way, this study argues that banks remain special as they lend claims on their own debt and have it accepted as money by the public. The study shows that banks and nonbank financial intermediaries perform complementary functions essential to the economy. It shows how risk-reduction policies in payment systems, banking asset allocation, and the deposit market affect the economy’s risk/efficiency tradeoff and the cost to generate resources to finance production. The study discusses the continuing relevance of banks in a world where nonbanks subtract business to them, financing production becomes a relatively less important source of banking business, and the use of e-money tends to spread fast. The study also looks at the way banking and nonbanking finance interact during different stages of economic development, producing various efficiency/stability configurations. It finally discusses the possible impact of the electronic age revolution on the role of banking and financial intermediation.

I wish to thank T. Beck, J. Caprio, J. Hanson, P. Honohan, A. Giustiniani, T. Glaessner, and R. Levine for their helpful comments. I am also grateful to L. Giammò for his comments on payment systems related issues. E. Mekhova has provided me with useful editing assistance. I am the only responsible for any errors in the text, and for the opinions therein expressed which do not necessarily reflect those of the World Bank. Last, but by no means least, I wish to thank my wife Ornella for all her enduring support.
Table of Contents

1. WHAT IS A BANK? .............................................................................................................................................3

2. WHAT MAKES BANKS SPECIAL: A REVIEW OF THE LITERATURE.........................................................5
   WHAT DO BANKS DO?........................................................................................................................................6
   Banks supply credit ............................................................................................................................................6
   ...and provide liquidity ......................................................................................................................................8
   In fact, they can’t do one thing without the other .........................................................................................9

3. WHY BANKS ARE SPECIAL: INTEGRATING MONEY, CREDIT, AND PAYMENT SERVICES..............................12
   SCHUMPETER AND THE MONETARY CIRCUITISTS .........................................................................................12
   BANKS AND FINANCIAL INTERMEDIARIES IN A PRODUCTION ECONOMY .............................................15

4. WHAT MAKES BANKS SPECIAL ................................................................................................................23
   BANKS HAVE SPECIAL POWERS ................................................................................................................25
   Banks extract seigniorage from the economy ..............................................................................................25
   Bank size and seigniorage ...............................................................................................................................26
   Banks tend to assume industrial corporate ownership and control ........................................................28

5. POLICY IMPLICATIONS: REDUCING RISKS SHOULD NOT KILL THE BANKS ..............................................30
   PAYMENT SETTLEMENT RULES: SOME MAY BE TOO COSTLY .................................................................31
   IS NARROW BANKING A GOOD IDEA? ........................................................................................................33
   SHOULD DEPOSITS BE INSURED? ................................................................................................................35

6. WILL BANKS REMAIN SPECIAL? .................................................................................................................37
   ARE BANKS STILL RELEVANT? ...................................................................................................................37
   IS PRODUCTION FINANCE STILL A MAJOR SOURCE OF BANKING BUSINESS? ........................................41
   WILL ELECTRONICS PUT BANKS OUT OF BUSINESS? ..................................................................................42

7. BANKS, FINANCE, AND ECONOMIC DEVELOPMENT ..................................................................................44
   BANKS, FINANCE, AND ECONOMIC DEVELOPMENT: PAST AND PRESENT ...........................................44
   Early industrialism: new industries, few banks ............................................................................................44
   Banks step in ..................................................................................................................................................45
   ...until the state takes it all ..............................................................................................................................47
   The “triumph of the small investor” ............................................................................................................48

8. CONCLUSIONS .................................................................................................................................................54

REFERENCES .......................................................................................................................................................56
“The banker...is not so much primarily a middleman in the commodity ‘purchasing power’ as a producer of this commodity... He stands between those who wish to form new combinations and the possessors of productive means. He is essentially a phenomenon of development, though only when no central authority directs the social process. He makes possible the carrying out of new combinations, authorizes people, in the name of society as it were, to form them. He is the ephor of the exchange economy.”

(Schumpeter, 1934, p. 74)

1. What is a bank?

Recently, in one of its School Brief series on finance, the Economist magazine has repeatedly described banks as intermediaries between savers and users of capital.¹ If the educational programs of renowned specialized media are reflective, as one would assume, of the knowledge that prevails among experts and practitioners of a human discipline at a certain point in time, one should take the Economist’s definition for banks, however simplistic it may sound, as a piece of widely accepted, basic economic knowledge. After all, Fama (1980) introduced his theory of commercial banking by defining banks as “…financial intermediaries that issue deposits and use the proceeds to purchase securities” (p. 39), and some two decades later the same definition is adopted by a leading text on banking theory (Freixas and Rochet, 1997).

Banks are nothing more nor less that financial intermediaries, the view holds. And to find evidence of this apparently unambiguous statement, one would only have to look at the history of banks since their archetypal forms came into being more than two millennia ago. However sophisticated their business has since become, banks in their very essence have always lent out the funds that were first placed with them by some patient money holders.

Nothing more, nothing less than intermediaries.

¹ See The Economist, various issues, October-December 1999.
But through the years, and still today, banks have been the object of particular concern from many quarters in most of the world, including governments, their publics, and the economic profession. For some reasons, they have been regarded as a *special* type of intermediary, one that needs differentiated treatment by the regulatory authorities and even special protective measures from competition and risk of failure. Indeed, there is no question that banks everywhere have received special treatment, to the point that one might legitimately believe that banks are special simply because they are specially regulated, or because regulators have made them so (Tobin, 1963; Kareken, 1985).

Is this concern with banks justified by a truly special nature of theirs? Do banks perform unique functions that can be complemented and maybe partially supplemented, but not quite entirely replaced by other types of intermediaries? And why in most cases have banks taken off earlier than nonbank financial intermediaries throughout economic development? Might this be the reason why they have attracted so much attention from regulators – which, to be sure, would equally turn to other financial intermediaries if these were to develop to a comparable extent?

In other words: are banks special because of some functional peculiarities, or because of historical developmental patterns? And, if history is the answer, did it just happen this way, or was there a necessary (natural) cause for it to go the way it did?

Over the last two decades, the literature in the theory of finance has inquired into the nature of banks trying to address what’s unique (if anything) about their function, and whether their uniqueness (if admitted) has survived the revolution in technology and the large and deep wave of financial liberalization that opened the gates to a flood of new intermediary types and financial instruments competing with banks and with their products for the public’s money. Progress in other branches of economic theory, such as the economics of information and contracts, and game theory, has been key in the search for insights. In the process, several answers have been given to these questions, and much has been learnt as to how intermediaries of various kinds operate in the economy; how different, similar, or even complementary their functions can be; and how different intermediation instruments coexist and may eventually replace each other in the market.

As technological progress transforms at exponential speed the ways and means of doing business in all sectors of the economy, it may be worth to re-explore the issue of the
special nature of banks in the light of the established (and still growing) body of knowledge on the subject. The question looms particularly relevant in the context of the debate on the relationship between finance and economic development: as the interest intensifies on how best to design financial systems conducive to high and sustainable economic growth in developing economies, gaining a better understanding of any distinctive role of banks and nonbank financial intermediaries becomes a key undertaking.

The study is structured as follows. Section 2 looks at the reasons given in the literature as to why banks are considered special intermediaries; major contributions and some recent works are discussed in the section. Section 3 points to the need to take an integrated view of money, production, and investment to gain a full understanding of the special function of banking and its complementary role to nonbank finance, and describes a functional approach to do so. Section 4 uses the approach to analyze the interactions between banks and nonbank intermediaries and their impact on the real economy. Section 5 evaluates how risk-reduction policies for payment systems, bank asset allocation, and the deposit market may affect the economy’s risk/efficiency tradeoff and the cost of generating production finance. Section 6 discusses the relevance of banks in a world where non-banks subtract an increasing share of their business, financing production becomes a relatively less important source of banking business, and the use of e-money tends to spread fast. Section 7 looks at the way banking and finance interact during different stages of economic development, generating various efficiency/stability configurations. The section discusses the possible impact of the electronic age revolution on the role of banking and financial intermediation.

2. What makes banks special: a review of the literature

Traditionally, the specialness of banks has been traced to the monetary nature of their (demand deposit) liabilities and to their running the economy’s payment system. Since the early experience of the deposit-taking institutions of the nineteenth century, banks have issued debt instruments that are accepted as means of exchange and payment on the basis of a fiduciary relationship among the agents using them and between the agents and the issuing banks.
Supplying transaction (and portfolio management) services is what characterizes banking according to Fama (1980), while Kareken (1985) emphasizes the central role of banks in managing the payment system.\(^2\) Corrigan (1982) adds to these functions the banks’ twofold role of backup sources of liquidity for all enterprises in the economy and of transmission belt for monetary policy.\(^3\)

Others have objected that with the evolution of financial markets and institutions none of the above functions is compellingly, and exclusively, pertinent to banks as such (see Golembe, 1983; Golembe and Mingo, 1985, and Goodhart, 1987). In advanced economies, transaction account facilities are supplied by non-depository - and even non-financial - institutions with access to payment clearing and settlement systems. Likewise, various other financial and non-financial entities can provide credit to business, while the backup-source-of-liquidity function in times of economic distress is in principle inconsistent with bank regulations aimed to prevent or forestall bank failures. Finally, where monetary policy is mainly conducted via open-market operations, government securities dealers (even more than banks) may act as transmission belt of monetary policy signals to the economy.

Research has thus looked for other features that may more specifically characterize banks as special financial intermediaries.\(^4\)

**What do banks do?**

**Banks supply credit…**

Diamond (1984) finds a special feature in banks acting as delegated monitors of borrowers, on behalf of the ultimate lenders (depositors), in the presence of costly monitoring. Essentially, banks produce a net social benefit by exploiting scale economies in processing the information involved in monitoring and enforcing contracts with borrowers. They reduce the delegation costs through a sufficient diversification of their

\(^2\) Kareken himself, however, and Fama (1980) before him, have envisioned regimes where payment transactions are divorced from banks and performed by nonbank entities or even by individual agents, which is what is increasingly happening these days.

\(^3\) In revisiting the issue almost two decades later, Corrigan (2000) confirms his original view.
loan portfolio. But if Diamond’s result shows that banks’ specialization in monitoring credits improves social welfare, neither it proves to hold for banks exclusively - since any kind of intermediary may equally benefits from portfolio diversification - nor does it explain why loan contracts are not replaced by more efficient risk-sharing, state-contingent contracts that reduce asymmetric information (such as equities).

What is in fact characteristic of bank loans is that their value is fixed in nominal terms and that they include collateral requirement clauses as well as costly bankruptcy provisions. By factoring ex-post information asymmetries and agency costs in the credit-making process, Gale and Hellwig (1985) show that such contract types – which they call standard debt contracts (SDCs) - are optimal financial arrangements that, on the one hand, save on the creditor’s costs of monitoring states of nature throughout the life of the loan and, on the other, give borrowers an incentive to minimize the risk of default and discourage them from hiding their true business performance.

The optimality of SDCs suggests a powerful argument to explain why banks have historically emerged as the first form of financial intermediation virtually everywhere when capitalistic production took place. However, SDC optimality is not robust against changes in the universal risk-neutrality assumption used by Gale and Hellwig in their model and does not hold in the case of ex-ante information asymmetries, where SDCs become exposed to adverse selection and moral hazard risks. Besides, as information and contract performance are crucial to the SDC optimality result, one would expect bank specialness to fade with the development of financial infrastructure since this provides agents with better information and more efficient contract enforcement technologies leading investors to prefer non-SDC contract types. Bank specialness is therefore a product of history, much like its own disappearance at some point.

Terlizzese (1988) uses ex-ante asymmetric information as a rationale for depositors’ preference to lend indirectly (by writing an SDC with a bank) over direct financing of individual entrepreneurs. As depositors are faced with a “lemon” problem, they generate a demand for delegated screening which banks have a comparative advantage

---

4 Bhattacharya and Thakor (1993) offer an extensive, critical review of the major issues and research areas relating to banking intermediation. For the public of Italian readers, Marotta and Pittaluga (1993) is an excellent reference.
to perform. In a repeated-game situation, the related agency problem is solved through reputation incentives.\textsuperscript{6} Wang and Williamson (1998), too, show typical debt contracts and intermediaries to develop in credit markets where screening borrowers is costly. This result is particularly useful since ex ante screening costs appear to be much more important than ex post state verification (auditing) costs for the functioning of credit markets.

...and provide liquidity

The credit view of bank specialness underscores the relevance of banks’ informational advantage vis-à-vis individual investors. Banks specialize in extracting and processing information concerning borrowers through their close relationship with them and in a way that is not replicable by individual investors. Bank information may be very exclusive and made unavailable to others. This is tantamount to saying that loans are illiquid to depositors and non negotiable in the market. Yet, as noted above, transferring funds from savers to fund users is not strictly peculiar of banks as such, and does not necessarily imply use of SDCs which the banks typically adopt. However, banks finance illiquid loans with liquid deposits which bear shorter maturity than loans, have nominally fixed value, and are available to holders on demand. The specialness of banks must thus rest with their capacity to provide liquidity services.

This is the core message of Diamond and Dybvig (1983). Instead of placing their endowments in an illiquid production technology, agents deposit them with banks in exchange for (interest-bearing) claims which entitle them to withdraw the deposits to finance future, unanticipated consumption needs (whose risk of materializing is assumed to be uninsurable). The agents thus gain greater consumption flexibility. Withdrawal requests, however, are served sequentially on a first-come-first-serve basis until the bank runs out of assets; thus, depositors face a positive risk of not being able to exercise their claims if the bank’s assets run out before their request for withdrawal is submitted. The return on deposits, on the other hand, enables the agents to achieve higher future consumption than if

\textsuperscript{5} Terlizzese (1989) shows state-contingent contracts (requiring no collateral) to be superior to SDCs in terms of efficient diffusion of information and risk sharing, under ex-ante informational asymmetries.

\textsuperscript{6} Interestingly, due to the ex-ante information asymmetry, banks should not find it possible to have depositors agree on deposit contracts contingent on states of nature, which provides an enlightening explanation for banks’ common use of SDCs to finance their assets.
they realized the illiquid assets. Banks therefore provide agents with a liquidity insurance service - or a higher degree of consumption flexibility - by exploiting the balance that holds continuously, under normal circumstances, between deposit withdrawals and renewals across depositors. This allows money to be shifted from patient holders to impatient consumers.

Yet, any non-convertible type of money (say, cash) would provide at least the same kind of consumption flexibility that Diamond-Dybvig bank deposits offer (although at a different risk-return tradeoff). The specialness of banks must therefore originate from their being able to transform liquidity into optimal illiquidity in the agent portfolios: they induce agents to replace safe cash for slightly riskier and still sufficiently liquid instruments. But this is possible if the banks use depositor’s cash to finance longer-term, illiquid productive assets that would not be financed by depositors directly, and if they extract from these assets enough rents to reward deposit illiquidity efficiently. Banks are thus intermediaries and their specialness must be traced to their ability to integrate optimal illiquidity creation with rent-extraction power.

In fact, they can’t do one thing without the other

Fama (1985) points to the specialness of banks as deriving from integrating credit and liquidity provision functions. By having borrowers hold deposits with them, banks can observe deposit movements and gain private information on borrowers, which they then feed into the processing of new loans. This gives banks a special role as information providers to capital market participants, who can incorporate the information embedded in banks’ lending decisions into their own evaluations. Although Fama’s theory may explain why transaction- and credit-related services have historically been integrated within the same type of institutions, it is not hard to imagine the informational advantage of banks being eroded by information agencies that specialize in providing information concerning the payment and credit status of individual agents.

---

7 The social welfare improvement due to private (bank) money backed by productive assets is discussed by Williamson (1999).
8 James (1987) and Lummer and McConnell (1989) have found confirming evidence to this hypothesis by looking at stock price variation of a number of firms after new loan negotiations or revisions and renewals of previous loans, were made public.
Goodhart (1986, 1987) looks at the banks’ peculiar asset and liability structure. Bank assets determine the nature of bank liabilities: holding assets mainly in the form of nominally fixed loans induces banks to issue liabilities largely in the form of deposits with guaranteed nominal capital value. This makes banks particularly fragile to perceptions of asset deterioration, to an extent that even requires the setting up of special safety nets. Goodhart emphasizes the exclusive nature of the information banks have on borrowers as the cause for loans’ uncertain value and non-marketability. Were it not for such special nature of banking and if bank assets could be evaluated and negotiated in the market, a new form of bank deposits could develop - Goodhart insists - very much similar to mutual fund shares, whose value would vary with the value of the assets and which, under some provisions, could even be used as means of payments.\(^9\) It is not clear, though, why banks should not be able to use more efficient risk-sharing types of contract and stick, instead, to SDCs for loans.\(^10\)

Calomiris and Khan (1991) investigate the specific liability issued by banks in the form of demandable debt to finance illiquid assets and show that, under costly and asymmetric information, demandable debt provides an incentive-compatible solution to the potential conflict of interest between (informed) bankers and (uninformed) bank depositors: the depositors’ right to withdraw their claims from a bank, if they become dissatisfied with their deposit returns, give them an incentive to monitor the bank. If enough of them agree on a negative assessment of the bank’s performance, they can call for liquidation. As a result, demandable debt induces bankers to pre-commit to a set of agreeable payoffs to depositors.

Integrating information-intensive lending and payment services distinguishes banks from other intermediaries, according to Goodfriend (1991).\(^11\) Systems to evaluate, monitor, and enforce loan agreements are productive both in originating loans to non-financial borrowers and in managing the lending associated with an efficient provision of payment services. In other words, efficiency gains can be extracted from jointly producing payments

---

\(^9\) See Goodhart (1987), but also Fama (1980) as a precursor of this idea.

\(^10\) If the information inferiority and risk-aversion of small depositors justify using SDCs for liabilities, this is not true for the asset side when the intermediary possesses exclusive access to borrowers’ information. Moreover, if banks diversify their assets sufficiently, there doesn’t seem to be compelling reasons why they should not issue SDCs as liabilities and finance assets with non-SDCs.

\(^11\) I am grateful to Thorsten Beck for bringing this work to my attention.
services and non-marketable loans. Similarly, since the same information can be used to establish the reliability of lending and deposit taking, payments services can be provided at lower cost by entities who also offer credit services.

Berlin and Mester (1996, 1998) emphasize access to interest inelastic (core) deposits as a special feature of banks. Core deposits insulate the banks’ cost of funds from exogenous shocks and allow them, in turn, to protect borrowers from exogenous credit shocks. Berlin and Mester find that banks funded more heavily with core deposits provide more smoothing of loan rates in response to exogenous shocks to aggregate credit. This type of protection gives to both banks and borrowers an incentive to undertake multi-period contractual agreements in which loans need not break even period by period, and allows banks to charge higher loan rates.

Finally, Diamond and Rajan (1998, 1999) view banks as superior devices to tight human capital to real (illiquid) assets. They emphasize the liquidity insurance services provided by banks to both depositors and borrowers as being made possible by their fragile capital structure: banks possess specific talents to collect the maximal value of loans to entrepreneurs and attract deposits from individual investors as they can commit to pass on them the amount they expect to collect from borrowers; they can do so because their fragile capital structure exposes depositors to illiquidity and credit risks and makes them prone to run on the banks if they think their claims are in danger; the run drives bank rents to zero and provides banks with an incentive to be credible borrowers. In this respect, bank runs are no longer an undesirable feature of deposit contracts but an essential inducement for banks to be credible debtors.

---

12 This incentive effect is the same as in Calomiris and Kahn (1991).
13 Banks in Diamond and Rajan are intermediaries, but their fragile capital allows them uniquely to improve social welfare by undertaking maturity and liquidity transformation with only a minimal amount of pre-committed financial capital. Note, however, that no “free-lunch” factor is at play here. Although Diamond and Rajan do not explore this issue, banks can perform their function only if their fragile financial capital is complemented by a reputational capital sufficient to generate enough trust from depositors that their claims will be honored by the banks at all times. Thus, banks may be seen as agents specializing in bridging the trust gap between savers and capital users (by appropriately investing in reputational capital buildup), where the cost for individuals to do so is too high.
3. Why banks are special: integrating money, credit, and payment services

Integrating the credit and liquidity functions of banks, as proposed by the theories reviewed in the last section, is necessary to gain an understanding the many sides from which banks interact with the real sector of the economy. Yet, all these theories share one limitation: all implicitly assume the pre-existence in the economy of some form of money (or, more generally, of some type of liquid claims on existing wealth) that can be deposited with banks and which banks can use to make loans. None of these theories analyzes banking as being characterized specifically by deposit creation.

Sir J. Hicks (1989) did so. He pointed that the banks’ ability to lend withdrawable deposits allows banks to hand over money to borrowers without giving up any cash (or third-party liabilities) simply by increasing their liabilities: banks create money.

Indeed, banks become special when their power is recognized to create money in the form of debt claims on themselves, which they lend out to borrowers and which the economy uses for payment. Banks create such new debts and promise to honor them from the moment borrowers draw on their loan accounts to issue payments.

The functional approach to banking and financial intermediation discussed in this section explores the role of banks as credit money producers, and discusses the implications that derive from taking such view of banks.\(^{14}\) The approach looks at how banks necessarily integrate liquidity, credit, and payment services as they issue new money to finance production. The approach shows how banking and nonbank intermediation serve complementary purposes in a production economy. But before describing the approach, let’s take note of earlier attempts to analyze bank money creation in a production economy.

**Schumpeter and the monetary circuitists**

The role of banks in nurturing economic development through money creation was analyzed by J. A. Schumpeter.

Schumpeter (1934) saw economic development as resulting from spontaneous and discontinuous changes in factor combinations, brought about by the entrepreneurs, that

---

\(^{14}\) For a formal presentation of this approach see Bossone (2000).
would alter an otherwise steady and identical circular flow of production and exchange, whereby the same products would be produced every year in the same way and each supply would be matched by an equal demand (and vice versa).

But Schumpeter also realized that, whereas in an ideally steady and identical circular flow all exchanges would be effected through means of payment supplied once and for all in given quantities and circulating at a given velocity, alterations to the circular flow could not possibly happen without creation of new purchasing power. This Schumpeter saw as the main function of banks.

Schumpeter thought of money creation via credit as the fundamental function through which banks enabled the entrepreneurs to adopt new factor combinations. Bank lending allowed new resources to be forced into new channels by giving entrepreneurs the power to exercise a demand for them.

For Schumpeter bank money (deposits) came into being in the process and for the purpose of granting credit to firms. Banks could add to the existing means of payment by lending promises to pay, and entrepreneurs could have access to bank money by mortgaging goods which they would acquire with the borrowed purchasing power. Thus, total credit could be greater than if it there were only fully covered credit.

Schumpeter saw that the bank credit structure projected beyond the existing commodity base. Unlike in the steady and identical circular flow where money claims were certificates for existing goods and past services, banks allowed for purchasing power to be created in anticipation of, and for the production of, new output: bank money was made up of claims on services and goods yet to be produced by the economy. Schumpeter saw that banks did not confine themselves to transferring existing purchasing power from depositors to borrowers: would they just do so, economic development would not be possible. However, his model was not articulated enough to identify different roles for banking and nonbanking financial intermediation, and to distinguish between production and investment finance.

15 In fact, Schumpeter noted that the creation of money by the banks establishing claims against themselves – as had been described by A. Smith and even earlier authors – had become a commonplace at his day.
The theorists of the *monetary circuit* have followed on Schumpeter’s steps.\textsuperscript{16} For over the last three decades, they have been studying the links between bank money and production, and have analyzed the properties of a monetary production economy modeled in the form of a circuit-type sequence process. Since production, incomes, and savings take time to generate, while firms need money to finance production until receipts are earned, new money need be created through bank credit. The circuit sequence thus starts when deposits are credited by the banks on the firms’ loan accounts. The process moves on as inputs are purchased, production starts, salaries are paid, and firms sell output and financial liabilities to income earners. The circuit closes when the firms use the proceeds from output and securities sales to pay back their bank debts.

Production in a circuit process can take place when money is lent by the banks to the enterprises, and equilibrium holds if all the money used in production re-flows back to the firms enabling them to pay off their debt. Banks create money as new loans generate deposits that are used to finance transactions until they re-flow to the banks as loans are repaid.

But while circuit provides a proper framework to identify the role of banks as producers of money, rather than mere intermediaries between savers and capital users, it fails to give a correct representation of financial intermediation by ignoring transaction costs and informational inefficiencies.\textsuperscript{17}

\textsuperscript{16} Monetary circuit theory has its major proponents in Europe - most notably France, Belgium, Italy, and Switzerland - where different groups of scholars are active under B. Schmitt, A. Parguez, and F. Poulon, J. Cartelier, A. Cencini, De Vroey, A. Graziani, and M. Messori, to cite just some of the leading names. On this side of the Atlantic, Marc Lavoie is a circuitist from Canada. In the United States, a circuit-based interpretation of the saving-investment process – which I adopt in this work - was propounded long ago and later reiterated by P. Davidson (1965, 1991), although this author may not be regarded as a circuitist. The major contributions to circuit theory do not constitute a fully unified body of economic knowledge, and in fact considerable diversity exists across them. Yet, some basic and important features are clearly distinguishable, providing common foundations for an integrated theory of money and production. Although the literature on the monetary circuit has grown rich since the pioneering work of Schmitt (1959), many publications are in languages other than English (for references, see Messori 1988 and references in Rochon 1999). Interested English-language readers should refer to Graziani (1990, 1996), Lavoie (1985), Rochon (1999), and Cencini (1995).

\textsuperscript{17} One notable exception on both counts is Messori (1988b). The role and the microeconomics of finance in a sequential monetary production economy with limited information and risk are further studied in Messori and Tamborini (1993, 1994).
Banks and financial intermediaries in a production economy

Revisiting the circuit approach to include the behavior of individual agents and institutions under risk and limited information enhances its power to analyze the functions of banks and financial intermediaries. I summarize below the main features of the revisited circuit approach.18

The economy includes four sectors: firms, households, the banking sector (including the central bank and commercial banks), and nonbank financial intermediaries. As will be discussed later on, banking functions and financial intermediation functions can be carried out by the same institutions; here, however, they are conceptually differentiated and are assumed to be performed by separate entities. For analytical purposes, such entities could indifferently be departments within the same institutions or subsidiaries of the same holding company.

Economic activity is represented as a one-period circuit process with a three (intra-period) phase sequence: a beginning phase (circuit-start), an interim interval, and an ending phase (circuit-end).19 Two commodities are produced in the economy for household consumption and capital investment from enterprises. The sequential steps of the circuit process are sketched below and the associated flow of funds is represented in Chart I:

1) At circuit-start (I), banks select borrowers among firms screened on the basis of their risk and efficiency, and negotiate with the selected firms the terms and conditions for one-period loans. The banks credit the negotiated loan amounts on the firms’ deposit accounts. The firms execute goods production using capital and labor, and use loans to pay wages to workers. Deposits are transferred from the firms’ bank accounts to the accounts of wage earners.

18 The exposition below reflects also a further revision of the circuit model that includes the central bank and the use of central banks for reserve requirements, free reserves accumulation by banks, and payment settlement purposes (Bossone, 2000b. Available on request to the author).

19 In the reality, the economy consists of multiple and overlapping circuit rounds as several bank loans of different maturity may be extended at different times, new production is carried out continuously, and debts on old production are retired at any time. Yet, the logic of each opened round having to come to a closure, as well as the conditions for it to do so, remain the same and can be analyzed in a simpler, one-round process description.
2) In the interim interval (II), household incomes are spent on consumption goods and/or saved. Unspent incomes go into demand deposits with banks and/or into long-term assets with (or through) financial intermediaries in the capital markets. Firms wishing to add to their stock of productive capital (*investing enterprises*) issue long-term securities in the financial market. Financial intermediaries evaluate the demand for funds from the investing enterprises and transfer savings to those most viable, in exchange for securities. The financed enterprises buy the capital goods needed. All money transfers and payments for goods and securities take place through book-entries on accounts held with the banks.

3) At circuit-end (III), the firms use their revenues from output sales to pay off their bank debt plus interest and the money originally created is destroyed.

**Chart I. Flow of funds in the circuit model**
Individual agents optimize their objective functions, involving satisfaction from consumption for households and profits for non-household agents, given prices and the information they possess. Households supply labor services to firms and receive wages. Their consumption varies with their expected income from real and financial assets. Households invest a share of their savings in long-term financial assets through financial intermediaries and hold bank deposits for transaction and precautionary purposes. Portfolio decisions depend on assets’ expected returns and on income variability.

Firms employ labor and capital to satisfy the demand for consumption and capital goods from households and firms, respectively. Firms finance production with short-term bank credit at circuit-start and repay banks at circuit-end. Their net profits are given by revenues from output sales less capital depreciation and credit repayment.20

The central bank advances reserve money to the banks on demand, which the banks hold as regulatory reserves in fixed proportion to deposits and as free reserves, including precautionary reserves and reserves for settling interbank payments.

Banks allow the circuit process to start; they provide short-term loans to firms to finance production. They possess technology and information to select and monitor the risk of borrowing firms and use incentives to ensure loan repayments. Since bank technologies and information are specific to the bank-borrower relationship, bank loans are not liquid.

The supply of loans to firms is positive in the loan interest rate and negative in the firms’ risk of default anticipated by the banks. The loan interest rate is negotiated between the bank and the firm. Banks increase lending to the point where the marginal return from lending equals the marginal default risk. The volume of credit to firms determines the quantity of inputs that firms can purchase in the factors market.

Money is created by the banks in the form of demand deposit claims from the moment firms’ loan accounts are credited with the agreed loan amounts. Deposits are acknowledgments of debt by the banks to all those who receive them as a result of payments or fund transfers. Banks promise to honor such debt from the moment funds are

---

20 In the formal model, relative prices are assumed constant. Thus, the interest rate determines the economy’s allocation of resources to current and future consumption and indirectly determines the demand for capital good by investing enterprises, which is taken to increase with the difference between capital marginal efficiency and the cost of funds.
drawn on the loan account by the borrowers. No cash circulates in the economy and all payments are made through deposit transfers across bank accounts (book-entries).

Banks extend credits to each other (subject to risk-based limits); thus, payments originating from the loans do not entail transfers of central bank money between banks unless and until receiving banks require paying banks to settle their debts in central bank money, or unless and until the economy’s settlement rules require interbank credit/debit balances to be settled in central bank money.

At circuit-end, banks receive debt payments from the borrowing firms. If firms are unable to pay off their debt, banks have to decide whether to renew (or roll over) the original loans, restructure them, or write off their losses. Deposits are destroyed (and net liquidity drawn from the system) when loans are repaid to the banks.

The model does not imply the existence of a money multiplier, although it includes central bank money. Since banks can lend new deposits and use interbank credit to finance deposit transfers, central bank money is not necessary for banks to generate loans, except that banks need to keep a balance between accumulated reserves and their deposit base (due to regulatory requirements and/or prudential practice). To the extent that banks can access reserve money from the central bank, banks may lend new deposits until it is convenient for them to do so after taking into account the cost of reserves. Thus, holding reserves affects bank money creation through an incentive effect. (In principle, the central bank may deny access to reserve money and ration bank money creation).

Financial intermediaries operate in the capital market, aggregate savings from several sources, and allocate them to capital users. Their liabilities are not accepted as money by the public and are not used in payment transactions. They are specialized in assessing the profitability and the risk of alternative investment options and the creditworthiness of the investing enterprises. Financial intermediaries finance investment with long-term funds generated by savings. They increase their supply of funds until the net-of-risk interest charged to the investing enterprises reaches its maximum, and ration supply thereafter.

---

21 As noted, banks may be asked by other banks to reduce or settle their debit positions. Also, unlike in the model above, banks are subject to demand for convertibility of their deposit in cash by the public.
22 Financial intermediaries can also act simply as brokers. This implies that the allocative and risk-management functions attributed to intermediaries would be performed by the households directly. However,
Box 1. The circuit model: objections and counter-objections

The circuit model stands two main types of criticism. Here they are briefly discussed.

**Objection:** The model is a single-period model and does not account for intertemporal behavior; as a result, while it captures the economy’s flow of funds, it does not reflect accumulation of stocks and stock-flow relationships.

**Counter-objection:** The revisited circuit approach discussed in this study represents the flows of funds that take place as a result of decisions made by different types of real-sector and financial agents acting in a sequence. It shows where and how money originates, where it goes, and where it ends along a logical time sequence. The approach may well incorporate agents’ stock allocation decisions based on long-term objectives and expectations and “trace” where and how the money initially created flows and accumulates along the circuit as a consequence of those decisions. The approach captures the effects of stock allocation decisions by determining their impact on the flow of funds, while stock allocation decisions can be modeled in ways that take into account the signals that agents extract from changes in the observed flow of funds. Also, the approach is flexible enough to incorporate economic and financial structural changes that may alter the direction of the money flows over time (see section 6).

**Objection:** The circuit model neglects the results of modern financial theory which explains the existence of banking and financial institutions as a response to frictions (e.g., costs of information, contracting, and transaction) impeding perfect market competition.

**Counter-objection:** The revisited circuit approach is consistent with the use of frictions as primitives to rationalize the role of financial institutions and instruments. One can invoke such primitives and yet use the circuit approach as an integrated setting to understand the different and complementary functions of bank and nonbank intermediaries, and to analyze the interactions of different sets of agents in a sequential setup. The approach can flexibly incorporate hypotheses on how frictions affect the behavior of individual agents and financial institutions, and can be used to assess their impact on the economy’s equilibrium.

Finally, through their financial departments banks, too, can act as capital market intermediaries taking long-term deposits from the public and issuing long-term loans to firms.23

---

23 One could think of the bank as having two separate departments - a monetary department and a financial department – with the financial department holding a deposit account with the monetary department, call it a financial deposit account (FDA). When the bank raises term-funds, it receives bank deposits from term-depositors’ accounts and credits them on the FDA. As the bank extend term-loans and borrowers start making payments, deposits move from the FDA to the accounts of the new claimers with their banks (the ultimate payees). The distinction between the monetary and the financial function of the bank emphasizes that the latter simply intermediates bank liabilities issued under the former.
The supply of long-term funds from intermediaries cannot exceed their collection of savings from the economy. In fact, intermediaries may choose not to raise more funds than what they can (or wish to) invest. Similarly, if savers prefer deposits to less liquid assets, intermediaries would be rationed on their demand for funds by a limited supply from savers. The investment actually funded is thus the least between the supply and demand for long-term funds.

In each period, the economy’s equilibrium requires that firms sell all their output and the investing enterprises raise enough funds to settle their contract obligations with the capital goods producers. If output remains unsold, firms need to finance its costs. From the financial standpoint, at circuit-end all deposits must be destroyed, all money incomes must be spent in consumption and/or investment, and firms must pay off their debt. The circuit closes as producers use the proceeds from output sales to clear their debt with the banks. If there deposits are left outstanding, banks need to roll over the credit positions left outstanding by those borrowers who have not settled their obligations.

At circuit-end, producing firms - as a group - can appropriate no more than the money originally injected in the system. This is equal to the initial bank loans but does not include the money needed to pay interest. Thus, firms as a group may at most be in a position to pay off the principal debt to the banks, while they need additional money to be able to settle the interest debt share (unless banks accept payments in kind). This payment problem has nothing to do with the capacity of the firms either to produce and sell additional output, or to extract a larger surplus from production by lowering the relative price of labor; it exclusively rests on the liquidity constraint inherent in a circuit process where a given stock of money, which does not cover for interest payments, is advanced by the banks to the firms at circuit-start.\textsuperscript{24}

Resorting to the financial market to raise the needed extra cash through debt is not an option for the firms, since in equilibrium all existing savings are absorbed by investments and since extra funds can be made available to the firms only if the banks place their interest income in the capital market. But this is impossible as it requires loan

\textsuperscript{24} In fact, increasing productivity or decreasing the real value of wages can only increase firms’ real profits (or material surplus) but can in no way help them raise additional cash, if the overall stock of cash is given.
interest payments to have been made already. Interest payments (in money) can therefore be effected only if banks inject new money in the circuit (through credit).^25

Chart II exemplifies the changes that would take place in the balance-sheet of each agent acting in the circuit sequence described above, as a result of their interactions. To highlight how banks create and inject money in the system, the example does not include the central bank.

---

^25 The sequence would be: banks lend short-term to firms → firms pay interests → banks receive interest income and place it in the capital market → firms borrow the (newly generated) long-term funds from the capital market and clear off the bank debt → the newly created money re-flows to the issuing banks and is thus destroyed. Thus, as banks inject fresh liquidity, firms can convert their short-term interest debt obligations into long-term liabilities through the financial market. This mechanism bears implications that will be discussed later on.
The balance sheets report individual transactions originated in the circuit process described in the text. Each transaction generates fund transfers that changes the asset and liability composition of the balance sheets. The numbers in parentheses indicates the place of each transaction in the sequence. Items in shaded areas indicate transactions initiated by their respective agent. All agents start with zero asset and liability stocks. At circuit-end, stocks accumulates as a result of net flows. Firm F holds deposits at bank A; household H holds deposits at bank B and securities deposits with intermediary I; intermediary I holds deposits at bank B. Firm F produces both consumption and capital goods and repays its debt to A using the proceeds from sales to H and from (long-term) lending from I. At circuit-end, the position of all agents is balanced, the economy has accumulated new savings and investment that add to the existing stock of financial wealth and physical capital, respectively. Note that if H were to hold bank deposits at circuit-end, F would not be able to pay off its debt to A, and A would have to either rollover its credit to F or write off the corresponding loss.

### Chart II. Transaction effects on balance sheets in the circuit process

<table>
<thead>
<tr>
<th>Bank A</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Loan Account F</strong></td>
<td>0</td>
</tr>
<tr>
<td>(1) + Loan claim</td>
<td>+100</td>
</tr>
<tr>
<td>(2) - Debt pay out</td>
<td>-100</td>
</tr>
<tr>
<td><strong>Deposit Account F</strong></td>
<td>0</td>
</tr>
<tr>
<td>(1) + Loan deposit from A</td>
<td>+100</td>
</tr>
<tr>
<td>(2) - Payment to H</td>
<td>-100</td>
</tr>
<tr>
<td>(3) + Payment from H</td>
<td>+80</td>
</tr>
<tr>
<td>(4) + Loan deposit from I</td>
<td>+20</td>
</tr>
<tr>
<td>(5) - Debt pay out</td>
<td>-100</td>
</tr>
<tr>
<td><strong>Debit with B</strong></td>
<td>0</td>
</tr>
<tr>
<td>(2) + Transfer to H</td>
<td>+100</td>
</tr>
<tr>
<td>(3) - Transfer from H</td>
<td>-80</td>
</tr>
<tr>
<td>(4) + Transfer from I</td>
<td>-20</td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Firm F</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposits with A</td>
<td>0</td>
</tr>
<tr>
<td>(1) + Loan deposit from A</td>
<td>+100</td>
</tr>
<tr>
<td>(2) - Wage payment to H</td>
<td>-100</td>
</tr>
<tr>
<td>(3) + Sales to H</td>
<td>+80</td>
</tr>
<tr>
<td>(4) + Loan deposit from I</td>
<td>+20</td>
</tr>
<tr>
<td>(5) - Debt pay out</td>
<td>-100</td>
</tr>
<tr>
<td><strong>Physical capital</strong></td>
<td>20</td>
</tr>
<tr>
<td>(4) + Real investment</td>
<td>+20</td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bank B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Credit Account A</strong></td>
<td>0</td>
</tr>
<tr>
<td>(2) + Transfer to H</td>
<td>+100</td>
</tr>
<tr>
<td>(3) - Transfer from H</td>
<td>-80</td>
</tr>
<tr>
<td>(4) - Transfer from I</td>
<td>-20</td>
</tr>
<tr>
<td><strong>Deposit Account H</strong></td>
<td>0</td>
</tr>
<tr>
<td>(2) + Payment from F</td>
<td>+100</td>
</tr>
<tr>
<td>(3) - Payment from F</td>
<td>-80</td>
</tr>
<tr>
<td>(3) - Payment to I</td>
<td>-20</td>
</tr>
<tr>
<td>(3) + Payment from H</td>
<td>+20</td>
</tr>
<tr>
<td>(4) - Transfer to F</td>
<td>-20</td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Household H</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deposit with B</strong></td>
<td>0</td>
</tr>
<tr>
<td>(2) - Wage payment from F</td>
<td>-100</td>
</tr>
<tr>
<td>(3) - Payment to H</td>
<td>-80</td>
</tr>
<tr>
<td>(3) - Payment to I</td>
<td>-20</td>
</tr>
<tr>
<td><strong>Financial Wealth</strong></td>
<td>20</td>
</tr>
<tr>
<td>(1) + Securities investment with I</td>
<td>+20</td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intermediary I</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deposits with B</strong></td>
<td>0</td>
</tr>
<tr>
<td>(3) + Payment from H</td>
<td>20</td>
</tr>
<tr>
<td>(4) - Transfer to A</td>
<td>-20</td>
</tr>
<tr>
<td><strong>Obligations</strong></td>
<td>20</td>
</tr>
<tr>
<td>(2) + Securities deposit H</td>
<td>+20</td>
</tr>
<tr>
<td><strong>Loans</strong></td>
<td>20</td>
</tr>
<tr>
<td>(4) + Loan claim on F</td>
<td>+20</td>
</tr>
<tr>
<td><strong>Net worth</strong></td>
<td>0</td>
</tr>
</tbody>
</table>
Bank A creates deposits by lending to firm F, while intermediary I only transfers bank deposits from household H to F by mobilizing its account at bank B. The two shaded areas in A’s balance sheet denote that money is an asset-liability created as a loan is issued to F, entitling F to an equal deposit claim on A to be used to finance its purchases. To the extent that the use of such a deposit claim by F generates transfers of monetary value from bank A to bank B, credits from B to A are necessary to support the deposit claims transferred.

In a model where deposit transfers would have to be backed by central bank money, bank A would need to start its operations with a capital endowment to buy/borrow central bank money. And even if A were to borrow from B, the borrowing should consist of central bank money.

4. What makes banks special

When the economy is considered in its various sequential stages, as in the simplified model above, very distinctive roles appear to be played by the credit market on one side, where liquidity is created to finance production, and the financial market on the other, where existing liquidity accumulated by savers is allocated to investments. This, in turn, implies a distinctive role for banks and financial intermediaries, whereby:

1) Banks allow the circuit to start by providing new money to production. Such money is in the form of banks’ own liabilities, or debt claims on the banks themselves, made available to borrowers under credit contract terms. Banks do not intermediate existing money, but add to it every time they extend new credit to firms in the form of new deposit claims, while simultaneously committing to honor the deposit claims already outstanding. Banks specialize in selecting borrowers to whom they allocate the new deposit claims.

2) Financial intermediaries collect existing liquidity (bank deposits) from savers with long liquidity positions (who, in this case, part with their liquidity) and allocate it to
investors with short liquidity positions. Unlike banks, the money intermediated by financial intermediaries does not represent claims on the intermediaries themselves; such money always consists of claims on banks (i.e., deposits) and, as such, it can only move across bank accounts. Thus, while intermediaries transfer money across agents with different liquidity preferences, in no case do they create money. By funding investment, financial intermediaries enable capital goods producing firms at circuit-end to appropriate the money spent in production and to service their short-term debt obligations to banks.

What then characterizes banks from other financial intermediaries is that (i) banks issue debt claims on themselves that are accepted as money by the public, and (ii) inject money into the economy by lending claims on their debt.

Banks and financial intermediaries thus perform functions that are different and complementary in originating money and in making it circulate and re-flow to its point of origin in the circuit process. These functions may be carried out either by separate entities or jointly, under more universal institutions that bring banking and nonbanking activities together. In many cases, banks operate also as financial intermediaries engaging in fund transfers from savers to users while nonbanks can offer liquidity services on certain types of liabilities, but this should not shade the distinctive nature of bank-money creation and financial intermediation.

This way of looking at banks has two implications for bank money. First, from the moment it is created, bank money is simultaneously an asset and a liability both for the bank that issues it and for the borrower that receives it (Schmitt, 1982; Cencini, 1995): banks do not create value and the newly issued money takes on value only when and if it is associated with new production.

Second, whereas in non-circuit models banks intermediate pre-existing value in the form of claims on commodities or reserve assets, in circuit models – like in Schumpeter’s view - banks issue claims on real resources yet to be generated by a production process that

---

26 In principle, however, agents could conventionally accept nonbank liabilities as money for transaction purposes. In this (hypothetical) case, the intermediary would de facto become a bank.

27 This function was well recognized already in the thirties by Bresciani Turroni (1936).
those same claims allow firms to set in motion. Banks thus carry a potential for both production and inflation (Mathieu, 1985; Rachline, 1993), and whether either prevails after money has been issued is also a function of how well banks select their risks.  

In sum: Banks are special because they can finance new production by creating money; this makes production financing cheaper than if banks were only able to intermediate claims backed by pre-existing real resources.

**Banks have special powers**

The exclusive power to create money and to advance it at an interest differentiates banks from financial intermediaries in some unique ways, giving them a special influence on the economy’s resource distribution process and industrial corporate governance structure.

**Banks extract seigniorage from the economy**

Where banks create money through lending, they command net real resource transfers from firms, that is they extract seigniorage from the economy.  

As discussed above, at circuit-end firms as a group may, at a maximum, appropriate just enough money to pay off their principal debt. When firms repay the principal, no net real resource transfer is involved from them to the banks. Payments of interest, however, require firms to give in a share of their output in exchange for no additional real resources from banks; hence, the net transfer.

That there is a net real resource transfer from firms to banks is explained by the nature of the interest rate on loans. Since banks create liquidity through lending, and to the extent that funds advanced through commercial lending do not originate from transfers of

---

28 This understanding of money goes beyond the “inside money” concept introduced by Gurley and Shaw (1960), whereby money is a claim by some private agents on others, and a liability for its issuer.

29 This conclusion differ from neoclassical models where banks play a passive and neutral role (see, for instance Fama 1980) on price formation.

30 For a discussion of seigniorage in an economy with a central bank and a (commercial) banking sector, see Baltensperger and Jordan (1997).

31 Banks may temporarily shift the obligations to the future by refinancing their debt over longer-term maturity, but at some point they would have to relinquish real resources anyway as term lenders would not postpone their credit exposure indefinitely. The net real resource transfer involved in loan interest payments is recognized by Graziani (1988).
pre-existing resources, the interest charged on loans may not be conceptualized as a compensation to foregone consumption and intermediation services.

In the circuit process, the interest rate (net of the resource costs involved in lending, including the remuneration of deposits) is a pure *rent* that banks extract from borrowers by virtue of their *exclusive* power to create money.

Such exclusivity rests fundamentally on both government and society restricting the acceptance of money to liabilities issued by selected agents. No rent would in fact be involved if, hypothetically, all agents could issue their own liabilities and have them accepted as money by everybody else. This same argument explains why different market structures of the banking sector, resulting in different market entry rules and in stronger or weaker competition between banks, alter the seigniorage power of banks.

Seigniorage is not extracted by agents that only intermediate existing liquidity. Unlike commercial loan interests and like payments for inputs, interest payments on long-term obligations represent production cost items against the company’s revenues and, correspondingly, income compensations to both investors for temporarily parting with liquidity and to intermediaries for providing intermediation services. As such, these payments do not involve *net* real resource transfers from firms to savers and intermediaries.

**Bank size and seigniorage**

The way money enters and flows through the circuit entails also a direct relationship between the size of a bank and its seigniorage power. Size is here defined as the share of a bank’s deposit base over the economy’s total deposits, and includes both the volume of deposits and the number of depositors. When bank A extends a loan to a borrower it stands a chance that, once used for payments, the newly created deposit claims are transferred to other banks where the payees hold their deposit accounts. The smaller the size of A, the higher the chance of the claim being transferred to other banks. When claim transfers occur, bank A incurs open interbank debit positions which it becomes liable to offset when so requested by the creditor banks or by the extant payment settlement rules: offsetting a debit position through settlement funds neutralizes the seigniorage from deposit creation.
Banks extend credit to each other charging risk-sensitive interest rates and under limits (caps) that vary with the size and quality of the borrowing banks. Typically, the smaller the size of A, the tighter the maximum net debit position that other banks accept from A. Smaller banks, thus, have a stronger incentive than large ones to limit money creation and to hold relatively larger reserve assets against their debit positions. Less stringent interbank debit caps thus lower bank demand for reserve assets. In the limit, if banks extended unlimited credit to each other, they would have no incentive (and indeed no reason) to hold reserve assets. Unlimited (and costless) interbank credit would be like having only one bank in the system, holding deposit accounts for all agents and managing all payments across its own books: in principle, the bank could issue deposits at will, without regard to having to raise liquidity for interbank obligations settlement, or to ensuring convertibility of its deposit liabilities into reserve assets.

In the absence of unlimited interbank credit, it becomes relevant for a bank to expand its size. By expanding its deposit base and by integrating credit, deposit, and payment activities, the bank reduces the obligations arising from its lending and enhances its seigniorage from money creation. By broadening its deposit base, the bank increases the number of payments settled within its own books and, ceteris paribus, decreases the number and the volume of its interbank debit positions. It may increase its lending with a lower risk that this will generate deposit transfers to other banks, and with a higher probability that deposits will flow back in its own books from other banks as a result of payment activity.

On the incentive side, unlimited interbank credit would have the same effects as full deposit insurance or suspension of deposit convertibility in stemming the risk of run on banks from depositors, and would eliminate restrains on deposit overissue.

The above arguments show that credit creation is not upper-bounded by deposits as it would be if banks were pure intermediaries transferring deposits across agents. Besides the role played by borrower-specific factors in credit supply decisions, optimum money

---

32 Contrast this with the extreme case of a system with only one bank. Money would circulate as book-entries within the same bank’s balance sheet. In this case, irrespective of how fast existing deposits shift hands, and abstracting from inflationary and confidence-related concerns, the bank could create new money at will, unimpeded by debit offsetting requirements. Note that in this case the bank would be in a position to extend loans of any maturity with no concern about asset-liability maturity mismatch. Related to size is also the
creation through lending in a bank credit economy is determined by each bank taking into account its interbank obligations, its means available to discharge such obligations, and the acceptability of its own debt claims to other banks.\footnote{33}

Finally, the previous discussion on size and seigniorage suggests that banks have an incentive to capture greater shares of the deposit market. Since this allows them to clear a larger number of deposit transfers within their own books, banks may gain significantly from attaching payment services to their lending-deposit activity. This conclusion complements Goodfriend’s (1991) argument, discussed earlier, which explains that banks specialize in integrating credit, deposit, and payment services due to informational advantages.

**Banks tend to assume industrial corporate ownership and control**

Firms settle their bank debt only by transferring real resources to the banks. Real resource transfers may include direct sales of output, payments out of profits, and transfers of fixed asset (e.g., real estate) and equity claims. Settlements may also involve transfers of savings borrowed from the financial market, although this solution is not sustainable over time.

Of the nondebt settlement instruments, direct output sales make a small share of total transfers, while fixed asset claim transfers find a natural bound in the firms’ endowment of marketable assets. Thus, in the long run, firms settle their debt by using gross profits and/or by selling equity to the banks or the public. Equity sales allow firms to transfer to the banks claims on their future income as well as control over their future income-earning capacity.

For given (desired) levels of profitability, using profits to settle bank debt requires that firms extract larger surpluses from input-owners or that they use inputs more intensively, raising productivity. Using profits may thus be hampered by mutually

---

\footnote{33} A contrario, this point could be taken to imply that, to the extent that banks ultimately have to mind about preserving the convertibility of their own money into reserve assets, even a free banking regime would pose limits to money creation. These would derive from self-imposed restraints set by individual banks on their own lending volumes. I am grateful to J. Hanson for emphasizing this point in conversation.
inconsistent distributional plans of firms and input-owners, rigidities to productivity growth, and slow business activity.

When profits are constrained by any of these factors, firms’ demand for liquidity driven by interest debt payment requirements needs to result in equity transfers from the firms’ original owners to the banks and/or the public. A market for corporate control would evolve which might take various forms depending on the institutional characteristics of the economy involved: where banks dominate the financial system, enterprise ownership claims tend to accumulate in bank portfolios, with the claims likely being exchanged in exclusive (bilateral) settings under noncompetitive and nontransparent conditions. On the other hand, in economies with more market-based financial systems, enterprise equity would be absorbed by the public, directly or indirectly (through institutional investors), in much more competitive and transparent contexts.

The market for corporate control would also be influenced by the business cycle. By compressing corporate profits, economic slowdowns and recessions should favor noncompetitive transfers of ownership from firms to banks, since exclusive bank-firm relations make it relatively easier and cheaper for both to convert debt into equity within the same bank portfolios. With sustained economic activity, however, larger corporate profits and the public’s greater appetite for risk and need for risk diversification should encourage transfers of ownership to households, other firms, and nonbank financial institutions through competitive market mediations.

But whether corporate ownership is prevalently transferred to banks or to the public may have a different impact on money supply. As shown earlier, if firms transfer ownership to the banks, the interest debt is settled directly as debt is exchanged for stocks and no additional money is needed in the circuit to make interest payments. On the other hand, if corporate ownership shifts to the public and to nonbank (financial and non-financial) entities, interest payments need to continue to be made from the firms to the banks, requiring new money to be injected in the system.
5. Policy implications: reducing risks should not kill the banks

Banks create money and lend it to those who use it to mobilize resources and produce output, before new incomes (and savings) are generated. In fact, banks create money against future incomes deriving from the production financed with the new money. Bank money creation thus entails a forward-looking process that moves from credit to output to debt service. To the extent that the money advanced as credit is backed up by real output ex post, this mechanism provides the economy with a greater potential for mobilizing real resources than one where real resources are needed upfront to back up new credit. This argument is the same as Schumpeter’s and carries with it important implications for the role of banks in economic development.

On the risk side, a system where banks can create money is definitely riskier than if banks were to fully cover their credit. Because banks issue liquid liabilities to finance illiquid activities, they bear significant illiquidity risks, they themselves generate risks and are vulnerable to risks originating elsewhere in the economy. Also, due to their large intra-sector and inter-sector open positions, banks may transmit shocks across the economy. Governments, therefore, introduce rules to prevent such risks, especially with a view to averting systemic consequences.

Given the tradeoff between cost-efficiency and risk associated with generating production finance, it is important for policymakers to make sure that the measures to protect the economy from systemic risks strike a right enough balance between the two terms of the tradeoff. This is important most of all for developing economies where risk-prevention policies from regulators may result in high costs for generating production finance.

Three issues are taken up in this section to assess how public policy to reduce risks interact with the incentives of banks as money producers. The first deals with payment settlement rules, the second discusses the narrow banking proposal, and the third looks at deposit insurance.
Payment settlement rules: some may be too costly

By supporting banks’ mutual debit positions arising from deposit transfers (payments), interbank credit in a circuit process allows banks to create the money necessary to finance economic activity. ³⁴ Each bank must be sufficiently creditworthy to get credit from the other banks and, in turn, must be willing to extend credit to the others (if these are sufficiently creditworthy) as deposits move across their books.

In the absence of distortions, lending to other banks places each bank’s capital at risk and gives banks an incentive to monitor each other’s creditworthiness and risk of money overissue. In other words, interbank credit provides each bank with market incentives to preserve its own credit standing and to ensure that other banks preserve their own asset quality. Banks are thus induced to price interbank credit based on the risk of the borrowing banks and to limit their exposures (or deny credit altogether) to overissuing banks.

Different payment settlement rules may affect differently the cost to the banking system to finance economic activity and to induce discipline into banks.

Under correspondent banking arrangements, banks hold accounts with their correspondents, through which they clear (incoming/outcoming) payments, and settle net end-of-day balances on the books of the central bank. Correspondents extend credit to their client banks and other correspondents through credit lines or overdraft agreements under pre-set terms and limits. Up to these limits, they commit to make good all incoming payments from their clients until new funds arrive on the client accounts and credits are repaid. ³⁵ The cost of money creation is determined by the specific types of interbank credit agreements. The reciprocity of such agreements and the adoption of netting contracts between correspondents can make the cost of interbank credit much cheaper than if banks were to settle their (gross) mutual obligations in central bank money (see below). Correspondent banks have incentives to assess, monitor, and control the risk of their client banks. Due to the nature of their relationships with client banks, they possess appropriate information channels to perform risk monitoring.

---

³⁴ Smith and Weber (1998) show, analytically and historically, the welfare superiority of payment systems with interbank credit.

³⁵ Assume that banks A and B hold correspondent accounts with one another and agree on mutual credit lines. If A’s borrower pays B’s client out of his loan account, A runs a liability vis-à-vis B. Bank A may finance its
Under *multilateral net settlement* rules, banks are required periodically to settle in central bank money their multilateral net payment obligations outstanding at the time of settlement. This requirement does not prevent banks from using (free of cost) interbank credit during the interval between settlement cycles to create the money needed to finance transactions. In fact, net interbank debit (credit) positions might in principle be very small and yet the banks’ mutual gross obligations could be as large as necessary to satisfy deposit transfers of any size. Multilateral netting rules thus provide an economically convenient way for banks to generate production finance. On the risk side, they may not necessarily structure bank relationships such as to provide banks with the mutual information channels available under correspondent banking. Still, rules can be designed to strengthen the incentives for banks to peer-monitor against money overissue and asset quality deterioration, including interbank liquidity- and loss-sharing cooperative agreements that make banks liable to take up a share of the liquidity needs or of the losses deriving from any of them failing to settle.

Under *gross settlement* rules, banks are required to settle each (gross) interbank deposit transfer in central bank money. Gross settlement virtually eliminates settlement risk as well as the possibility of systemic transmission of the failure of individual banks to settle. But this comes at a cost. Recent studies have focused on the efficiency costs of gross settlement rules,\(^{36}\) including payment gridlock problems and suboptimal intertemporal consumption effects (Kahn and Roberds, 1999).

Circuit theory adds significantly to the cost list. The gross settlement requirement implies that no deposit claim can be issued to borrowers and used for payments if it is not fully backed by central bank money. In other words, banks can issue and mobilize deposits only if they possess *already* enough real resources to back up the new circulation. This tightens the resource constraint on production finance generation. To be sure, banks can borrow/lend central bank money from/to each other but, although this allows them to economize on scarce reserves, it does not alter the implication that - for the economy as a whole - the real resources needed to borrow/purchase central bank money must pre-exist to bank deposits.

\(^{36}\) See Kahn and Roberds (1998) for a review of the literature and a contribution to the issue.
Thus, gross settlement rules replace a bank credit economy where banks can create money by exchanging cheaper (and mutually cost-offsetting) IOUs with a more costly \textit{cash-in-advance} economy where banks face a real resource constraint on generating production finance, that is, where the money they create must be backed up by real resources \textit{ex ante}, rather than \textit{ex post} when new output has been produced.

The cost of gross settlement rules can be reduced, and settlement risk satisfactorily contained, by combining features of netting and correspondent banking in some kind of \textit{hybrid} arrangements where a number of larger banks would act as correspondents and clearinghouses for smaller client banks, establish reciprocal interbank credit relations, and hold settlement accounts with a central agent. Each correspondent/clearinghouse would clear the payments of its client banks within its own books and deal with the other correspondents/clearinghouses for the payments between their respective client banks. Only net balances between correspondents/clearinghouses could be settled in central bank money either at the end of the day or at intervals during the day. The net balances between each correspondent/clearinghouse and its client banks could be covered by interbank credit lines or overdraft facilities (subject to limits) that would reduce the cost to smaller banks for production financing.

The hybrid arrangements would make production finance generation cheaper than gross settlement rules, while they would make market discipline on risk management stronger than under netting or correspondent banking. In particular, the large correspondent/clearinghouse banks would exert a strong disciplinarian role on the smaller client banks and among themselves.

\textbf{Is Narrow Banking a good idea?}

Risk reduction lies at the heart of the various \textit{narrow banking} proposals periodically recurring in the literature and the policy debate. These proposals aim at removing illiquidity and insolvency risks from the banking business by requiring banks to back up their demand deposit liabilities with safe short-term assets.\footnote{The proposal was first introduced by Simons (1934, 1935). See also references in Wallace (1996) and Thakor (1996). For a discussion, see Goodhart (1987).} Noting that the magnitude of such assets outside the banking system exceeds the magnitude of banks’ demand deposit
liabilities, the advocates of this approach assume that there is an unsatisfied demand from the public for safe assets that banks do not match because of existing distortive incentives — such as deposit insurance, government regulation, or bailout commitments — which encourage banks to hold illiquid portfolios.\textsuperscript{38}

In fact, one should wonder whether such a radical approach would not end up by killing the essential function that banks perform in capitalist economies. Wallace (1996) shows that the narrow banking solution eliminates the role of banking altogether. As he puts it, “…using narrow banking to cope with the potential problems of banking illiquidity is analogous to reducing automobile accidents by limiting automobile speeds to zero” (p. 9).

Such conclusion is only reinforced by circuit analysis. The narrow-bank concept conflicts with the understanding of the bank as a money producer and circuit-starter. If banks can create money through lending to finance production, a high leverage is intrinsic to their incentive structure whereby liquid debt claims (deposits) stand necessarily against less liquid assets (loans to firms).

Under narrow banking, banks cannot finance production. But then: How would money be injected in the economy under narrow banking? How would production be financed? What would be the net social welfare cost from financing production differently than under traditional banking?

One way to finance production under narrow banking would be through specialized nonbank financial intermediaries. The central bank creates money through money market operations. Nonbanks could borrow money directly from the central bank or from the money market, and transfer it to the producing firms. This solution raises the same limitations of gross settlement rules.\textsuperscript{39}

The social welfare implications of financing production though nonbank intermediaries, thus, depend on the way money is made available to the economy. Under central bank lending or money market operations, the intermediaries must have enough

\textsuperscript{38} See, for instance, Kareken (1985).

\textsuperscript{39} In this respect, gross settlement rules may be seen as a different way to achieve a kind of narrow banking: whereas the latter requires a 100-percent reserve on deposit stocks, gross settlement rules impose a 100-percent reserve requirement on the flows deriving from deposit transfers (payments).
collateral or securities to raise the reserves needed to back up their lending. As discussed, this constraints production, raises the cost of generating production finance.

Under narrow banking, production can be financed also in an alternative way. Firms may sell liabilities in the financial market bearing maturity terms consistent with the production cycle they finance. Even assuming that all firms can do so, two main problems arise: first, new production must be financed with the stock of money already existing in the economy; to the extent that such money is somebody’s credit (either central bank’s credit or credit from whoever has originally purchased reserves from the central bank), its reutilization in new activities diverts it from re-flowing to its original creditor, the interest rate and the velocity of money would both rise as new production needs financing. The central bank accommodates must accommodate the new money demand with the same consequences as above.

Second, in the absence of liquidity transformation by banks, production finance under narrow banking requires savers to hold in their portfolio larger shares of illiquid claims on the firms. Unless this happens, either some production plans go unfulfilled or new money has to come from the central bank, with again the same effects as above.

The conclusion is that whereas narrow banks may emerge as a market response to an unsatisfied demand for safe assets, thereby filling an incomplete segment of the financial market, it would be wrong to transform by decree all traditional banks into narrow banks as this would create a severe market incompleteness: banks as highly leveraged institutions are necessary to reduce the cost to generate production finance.

**Should deposits be insured?**

Governments concerned with systemic risks may resort to deposit guarantee mechanisms to protect depositors (especially smaller ones) from the risk of losing their money if their banks go bust. If banks are seen as providers of liquidity insurance services to depositors under a sequential service constraint (Diamond and Dybvig, 1983), it apparently pays to prevent banks from the risk of being run by depositors (for right or wrong reasons) by introducing an insurance that guarantees the promised returns to all who withdraw their money. This insurance eliminates the incentives for depositors to participate in a bank run and thus preserves confidence in the banks.
The merits of deposit insurance have been questioned on grounds of moral hazard. Diamond and Rajan (1998, 1999) and Calomiris and Kahn (1991) argue that bank runs have good incentive effects on banks as they induce bankers to commit to be credible debtors (see section 2). Deposit insurance neutralizes such incentives and weakens the discipline that banks derive from their fragile capital structure.

In the circuit framework discussed above, deposit insurance can be shown to have even more serious perverse effects on the banks’ incentive structure. When bank deposits are protected with insurance, the incentive for banks to manage the credit risk on their interbank lending is weakened. This is not just because the commitment of each bank to be a credible debtor to its own depositors is lessened by the insurance coverage but also because, as a lender to other banks, each bank knows that the borrowing banks are not subject to runs from their depositors if their quality deteriorates. In fact, banks may even attract new deposits if they perceive their own risks to be threatening their liquidity/solvency position. Such incentive alteration may worsen the quality of interbank credits and bank peer-monitoring, leading to wrong pricing of interbank credit facilities, too generous interbank credit limits, and inadequate risk control. This would likely result in over-extension of credit to risky banks and, ultimately, in an inaccurate market selection of banks. And since risk-sensitive interbank lending is crucial to support the banks’ ability to safely create money via loans to firms and to select best banks in the long run, poor management of interbank credit risk may result in individual banks overissuing money and financing lesser quality firms.

In the circuit framework, therefore, deposit insurance not only tends to raise the moral hazard of individual banks, it debilitates the capacity of the whole banking system to enforce market discipline upon and within itself.

One could experiment with the idea of having depositors insuring their deposits against the risk of bank failures, by buying policies from private insurers. As the insurers would be forced to correlate policy premiums with bank risks (certainly under stronger pressure and with better information and expertise than government bureaucrats in state-run deposit insurance funds), they would have an incentive to monitor banks closely, thereby exerting discipline on the banks’ risk behavior. On the other hand, market-determined premiums for insurance would lead depositors to select and monitor their deposit banks
more carefully. Of course, the idea requires a much thorough investigation but it essentially emphasizes the need for depositors and banks to evaluate and manage risks free of moral hazard, whereas the current deposit insurance schemes tend precisely to do the reverse.

6. Will banks remain special?

This question arises spontaneously after observing that almost everywhere in the industrialized world, over the last two decades, traditional banking (deposit/lending) activities have lost relevance in quantitative terms. Such loss is the result of the impressive growth of commercial paper markets and of the money market mutual funds thrift industry, the increasing ownership of banks by securities firms and commercial enterprises, and the proliferation of quasi-monies or money-substitutes with associated transaction account facilities comparable to bank-issued transaction balances. Correspondingly, banks have had to diversify their activity from traditional deposit/lending to nonbank intermediation and financial services provision.\(^\text{40}\)

For many, the quantitative loss reflects a qualitative change: traditional banking has lost importance vis-à-vis other forms of financial intermediation, and banks - as we have come to know them so far - will either disappear or become practically indistinguishable from other financial institutions. Can banks remain special in the future? Will they be replaced by nonbank intermediaries? Or will they have to become nonbank intermediaries to survive in the new world? What will happen to them in the new age of electronic money?

Are banks still relevant?

In the industrial world, following financial liberalization, nonbank financial activities have subtracted considerable market shares to commercial banks. Increases have been observed in the market share of institutions holding securities instead of loans. At the same time, domestic banking sectors have undergone profound reorganization with banks consolidating into fewer and larger units.

An interesting interpretation of the disintermediation problem has been provided by the core-deposit theory of Berlin and Mester (1996) recalled earlier: as banks have increasingly been forced to pay market rates on an increasing share of their funds, core deposits have shrunk and multi-period relationship lending has become less and less feasible, causing banks to lose some of their comparative advantage vis-à-vis nonbank intermediaries.41

Table 1. Banking and financial indicators for G7 countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial claims (%GDP)</td>
<td>4.4</td>
<td>5.2</td>
<td>5.8</td>
<td>5.6</td>
</tr>
<tr>
<td>Financial instruments (%Fin. claims)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deposits</td>
<td>25.0</td>
<td>22.1</td>
<td>20.9</td>
<td>19.4</td>
</tr>
<tr>
<td>Loans</td>
<td>27.5</td>
<td>24.1</td>
<td>23.2</td>
<td>22.0</td>
</tr>
<tr>
<td>Equities</td>
<td>12.3</td>
<td>14.3</td>
<td>15.2</td>
<td>17.8</td>
</tr>
<tr>
<td>Bonds</td>
<td>10.7</td>
<td>12.4</td>
<td>11.2</td>
<td>13.8</td>
</tr>
<tr>
<td>Household deposits (%Fin. claims)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans to firms (% Tot financing to firms)</td>
<td>43.9</td>
<td>41.7</td>
<td>39.9</td>
<td>36.6</td>
</tr>
</tbody>
</table>

Source: Calculations on BIS data.

The circuit approach suggests an alternative interpretation that accounts simultaneously for bank disintermediation and consolidation. Since liberalization has caused the demand for core deposits to decline, individual banks - especially smaller ones - have faced a higher risk of running increasing interbank open debit positions as a result of

41 The concept of core deposits must be used cautiously. In a circuit process, increases in the interest elasticity of the demand for deposits do not imply destruction of existing deposits but simply their increasing velocity of circulation. In principle, to the extent that banks extend sufficient credit to each other and/or that their size allows them to minimize interbank exposures, the decline of core deposits should not upset lending.
lending. Banks have thus started to run higher reserve holding costs against the risk of defaulting on their interbank obligations. This has put pressure on many of them to reduce lending (causing disintermediation), to increase their deposit market share (through mergers and acquisitions) so as to capture larger re-flows of deposits from lending, and to expand the range of financial services provided to their clients alternative to core deposits. Unsuccessful banks, on the other hand, have been forced out of the market, again leading to industry consolidation.

Moreover, the advent of competitive nonbank financial intermediaries has facilitated the life of investing enterprises seeking to raise investment funding from the capital market. This has lowered savers’ demand for core deposits vis-à-vis securities and equity instruments, while improving the capacity of producing firms to appropriate the money needed for debt service at circuit-end by accessing the capital market. As a result, banks may have had lesser need to refinance the original short-term commercial loans and may have reduced their commercial lending.

But, can we imagine a world where nonbank intermediaries replace banks entirely? This question brings us back to a world dominated by narrow banks, discussed earlier. Suppose that such intermediaries are securities dealers with collateralized access to central bank funds, who can redistribute liquidity among themselves through the money market. Suppose also that all firms have access to securities dealers and that bank deposits are issued by, and held at, narrow banks. Deposits are used both as liquid assets and for transaction purposes, and all payments are managed by the narrow banks. Nonbank intermediaries finance new production by accessing the central bank or the money market, and funds are transferred through book entries at narrow banks.

The problem with this setup is that since all finance to production requires central bank money, and all borrowings from the central bank must be collateralized, securities dealers need to fully back up their loans to production with capital: a much less efficient solution than lending to the economy through highly leveraged institutions like banks, which can finance loans by issuing uncollateralized debt and only use a small capital base to support lending.

Alternatively, assuming that enough liquidity is already in the system, the nonbank intermediaries could finance new production by issuing securities to the public with a
maturity structure similar to that of the assets financed by them. But this would deprive the markets of the liquidity services that banks provide by transforming maturities. On the other hand, if investors wish to stay liquid and keep their claims with the narrow banks in lieu of holding nonbank liabilities, production finance would not be available in sufficient quantity since narrow banks are (by definition) prohibited from lending to firms. It thus turns out that banks allow for an efficient completion of the economy’s financial markets; they complement intermediaries in an essential way and cannot be replaced by them.

To be sure, financial and regulatory innovations enable bank money to circulate more rapidly and nonbanks to issue quasi-monies and to create bank credit substitutes.
Box 2. Are mutual fund shares real money?

Mutual funds pool savings from many individuals in exchange for (market-valued) shares in the funds, and invest them in various activities. Money market mutual funds invest in high-grade, short-term securities that offer market returns on share accounts that permit check-writing and wire fund transfer privileges.

A close inspection at money market mutual funds, however, shows that they do not make banks and their money-creation function any less special, although they allow bank deposits to circulate faster in the economy: First of all, when mutual funds issue shares they receive in exchange bank deposits that they initially hold with their bank or banking department: deposits are not replaced but move from individual to mutual fund portfolios. Second, when mutual funds invest in securities, they actually transfer the deposits originally purchased to the accounts of the securities issuers who, in turn, can access the borrowed deposits to finance their expenses: deposits are not replaced in lending and spending operations. Third, when shareholders instruct their mutual fund to make payments out of their share accounts, the mutual fund draws down its bank deposits and transfer them to the payees’ bank accounts: deposits are not replaced as transaction devices and payments take place as deposit transfers on bank books. Of course, the mutual fund can finance its payment obligation through a credit line from its correspondent bank, or it can simply realize its securities investment; yet both options still involve bank deposit transfers. Mutual funds do not create money; they only transfer it from savers to users.

Mutual funds with many shareholders can exploit the law of large numbers and economize on the deposits that individuals would hold in their absence. But banks stand ready to purchase securities from mutual funds (or to lend deposits to them) if the funds do not have enough liquidity to cover for their payment obligations: mutual funds need banks to stay liquid.

However, to the extent that quasi-monies do not replace deposits as means of payment (see Box 2), nonbank credit substitutes can only serve to channel existing bank money from savers to capital users more efficiently. Replacing banks with nonbank intermediaries would constrain the productive potential of the economy and make production finance more costly to generate.

Is production finance still a major source of banking business?

Even some circuitists today observe that the significance of bank lending to industrial and commercial undertakings has all but declined in some advanced industrial countries relative to household credit to consumption and second-hand asset purchases, including for speculative trading. But this does not diminish the role (and the relevance) of
banks as producers of money, although it modifies the standard circuit flow of funds discussed earlier.

In the case of consumer credit, high industrial productivity growth requires nowadays ways to expand and accelerate consumption. As households are willing to commit part of their future incomes to fund higher current consumption beyond their current incomes, bank consumer credit allows firms not only to appropriate enough money to pay off their debt at circuit-end (as in the standard circuit process), but also to make extra profits corresponding to household current dis-savings. As corporate profits are invested in the capital market, banks can borrow them at maturities that match those on their consumer loans and balance their position. In this new flow of funds, firms are the net savers (instead of households) while banks are both suppliers of fresh money (to firms and households) and borrowers of funds in the capital market (from the firms).

In the case of credit to second-hand asset purchases by households (including durables, houses, and financial assets), banks issue new deposits and finance second-hand asset acquisitions that in turn generate incomes and new funds to the capital market. Banks can then refinance their positions in the capital market as loan accounts are drawn down on borrowers’ payment orders. In this case, the households are the fund users while the sellers of the assets and the intermediaries involved provide the savings by investing their profits from asset sales and intermediation services, respectively, in the capital market. As money suppliers, banks allow deposits to accommodate the higher demand for funds generated by the increasing volume of financial transactions driven by trade of existing asset stocks.

In all the above cases, the efficiency of the economy’s financial infrastructure (in particular, in allowing rapid fund transfers across the economy) is key to ensure than banks refinance their positions in the market promptly and smoothly.

Will electronics put banks out of business?

What is the future of banks with the advent of electronic money? For reason of clarity, the concept of e-money here used rules out innovations introduced with the only purpose to move around bank money more quickly - such as home banking, electronic bill paying, ETF, and ATMs – and refers exclusively to computer money-forms that represent
liabilities of nonbank entities, or special segregated bank accounts, and that actually substitute for conventional monies.\textsuperscript{42}

E-money is initially issued to clients by network operators on network accounts in exchange for conventional bank deposits. The e-issuers hold the original bank deposits, which are perfectly matched by what the e-issuers owe to their clients. Assuming full acceptability of the new money instrument and full trade security, the e-money is used by cyberspace shoppers to finance purchases in the e-malls. Merchants are paid in e-money and convert their e-receipts into conventional bank deposits. Accordingly, their bank accounts are credited, while shoppers’ bank accounts are drawn down equivalently.

At some point, the e-issuers may start offering e-deposit facilities, so that merchants can either hold their e-value with the issuers or use it for further payments in the net. This could extend to other payers and payees along the trade chain. Thus, once issued, the e-money stock could in principle never return to the conventional banking system; it would become an indefinite leakage for it and grow independently of it.

Are banks to disappear in this scenario?

Quite the opposite. Even though individual banks might be unable to make it in the new world, banking as such would see its own rebirth - though under new disguise. As the new e-deposits become accepted in the system, some e-issuers would face an incentive to move away from fully backing e-deposits with conventional deposits and start using the float to make e-loans. Further down the road, as the size and reputation of the e-issuers grow and consolidate throughout the net community, they would have an incentive to lend e-deposits by acknowledging new debts (in addition to existing e-deposit claims) and promising to honor them as the e-loan accounts are drawn down: history would repeat itself.

Conventional banking would be replicated in the cyberspace and any nonbank entity aiming to issue e-deposits in the cyber-circuit would have to possess at least those same skills than bankers have developed for centuries when dealing with more material money stuff. E-banks would still be \textit{special}: they would create e-money by lending their own e-debt and will remain a very distinct specie for a while longer.

\textsuperscript{42} For an extensive presentations of the many issues relating to e-money, see Solomon (1997).
7. Banks, finance, and economic development

Banks provide the raw material for a monetary production economy to work and grow (regardless of the technological form of money).

The positive impact of finance on economic growth has received significant empirical support during the nineties, starting with the contribution by King and Levine (1993). Research showed also that counties with more developed banking systems and liquid capital markets have experienced the most rapid growth (Demirgüç-Kunt and Levine, 1996), confirming the importance of complementarity between banking and nonbank financial intermediation.43

In the following, based on the elements derived from the circuit approach and using a highly-stylized description, I will digress on how banking and financial development may determine various efficiency/stability configurations across different stages of economic development. I will also consider the possible role of banking and finance as technology evolves, in the light of what seem to be some of the most pressing conditions that need to be in place for progress to deliver its benefits to largest possible communities of people.

Banks, finance, and economic development: past and present

Early industrialism: new industries, few banks

Early in the process of industrialization - as in 18-century Europe - economies are dominated by manufacturing firms. Family-owned firms specialize in profit-oriented production and commercialization of commodities. Production and trade are financed through issues of bank-guaranteed commercial paper.

While the acquisition of the capital goods used in production is funded with the owners’ own funds and with business internally generated surpluses, capital goods producers finance their inventories through commercial credit from input suppliers and contract sales. Banks have a very limited role.

---

43 For a comprehensive analysis of the role of financial development for economic growth, see Levine (1997). For recent empirical works on financial structure and economic development see www.worldbank.org/research/interest/intrstweb.htm
Firms determine the total level of resource employment for the given production technology. Since they finance the acquisition of capital goods internally, they – as a group – determine the share of the economy’s resources allocated to consumption and investment. To the extent that labor market negotiations set nominal wages and trading involves commodities already produced (that is, no forward contracts are allowed), firms (as a group) also determine the volume of output that they appropriate at circuit-end.

Workers’ saving decisions that turn out to be inconsistent with firms’ production plans prompt commodity price adjustments to the point where all money spent by the firms is appropriated by the firms themselves at circuit-end.

The circuit process of the early industrial economies is rudimentary and relatively stable overall (to the extent that, and for as long as, its distributional effects are accepted by the society). The limited role of banks as money suppliers constrains production and saving generation, while capital accumulation is limited by the lack of organized finance for long-term investment and depends almost exclusively on the capitalists’ sources of personal wealth. Bank seigniorage is low or nonexistent, there is no transfer of ownership rights from firms to banks or the public, and equity capital is the domain of a restricted class of owners.

**Banks step in…**

The integrated role of banking and finance evolves when the demand for capital equipment and labor inputs intensifies and financial needs exceed personal wealth - as in Europe and America of late 1800s-early 1900s.

Banks support increasing production by lending new deposit claims. Moreover, bank loan and deposit contracts are cost-efficient in environments with poor financial infrastructure, scarce information, and a high demand for safe and stable returns on money from savers: banks become the main financial players.

But since the demand for credit to produce capital goods increases and the supply and demand for capital goods and financial capital specialize across different groups of agents, the risk rises that at circuit-end firms fail to appropriate the money borrowed from the banks. Also, as the demand for financial capital outgrows internally-generated funds and the agent investment horizons become more extended, investing enterprises need
increasingly to rely on a stable supply of long-term funds and to minimize uncertainty about their ability to attract and retain funding.

To assist the investing enterprises, banks develop investment banking and financial intermediation functions. With capital markets in their infancy and relatively few (as well as unsophisticated and not well-informed) large investors, banks are well positioned to draw long-term funds from savers and transfer them to investing enterprises on acceptable terms. As more bank credit is injected into production, financial intermediation provides investing enterprises with patient money. Once used to fund investments, such money enables producing firms to stay current on their bank debt service obligations at each circuit round close: financial intermediaries complements banks in an essential way.

However, with increasing specialization and decentralization of both commodity production and financial services provision, firms as a group lose control of output allocation to consumption and investment, the profitability of individual firms becomes more volatile, and the net real resource transfer from firms to banks (due to seigniorage) increases.

Banks and firms thus grow a strong common interest to protect firms’ cash-flow generation, and develop exclusive relationships whereby firms gain from banks stable access to long-term capital - even at times of adverse contingencies - in exchange for equity positions. Exclusive relationships give banks considerable rent-extraction power over the firms being financed. Universal banks emerge with increasing control of firms and manage to clump firms operating at different vertical levels of the same commodity sectors under single parent entities, with the purpose to stabilize the flow of funds from savers to end-users and ensure regular bank credit recovery at circuit-end.

The system may become more stable overall as firms face a lower risk of not appropriating the full money at circuit-end, and circuit closure is more safeguarded. In the long run, however, exclusive relationships lead to market information opaqueness, illiquid investment positions, restrictions on competition in both the real and financial sectors, and lesser reliability of the price system as an allocation device. Eventually, exclusive relationships weaken the incentives to efficiency, technological innovation, and new industry creation (Rajan and Zingales, 1999).
...until the state takes it all

To a large extent, the expanding role of the state in finance - as experienced in the industrial countries after the 1930s and almost everywhere in the world in the aftermath of the second world war (with the exception of the United States which remained a market-oriented economy) - replaces investment banking at the core of the financial side of the circuit, both by directly absorbing savings and by centrally managing funds allocation through financial repression, directed and subsidized lending, public money transfers, and control of financial institutions.

With the state acting as a financial intermediary, fund-savers become less concerned with the reputation of private-sector financial institutions, as they place savings with state-owned or state-controlled banks under the implicit guarantee of (perceived) unlimited solvency of the public sector. Also, the large share of public spending on aggregate output stabilizes the flow of funds that firms need to appropriate at circuit-end.

Under state finance the (actual) cost of funds bears little or no relation to risk- and information-related factors. Fund rationing is decided centrally by the government, based on macroeconomic or political policy objectives, and irrespective of risk factors.

In extreme forms of state finance, no seigniorage is extracted by the banks from the enterprises. In fact, as the state advances money to firms at below-equilibrium (typically negative) real interest rates to push production and resource employment above their market equilibrium, the net real resource transfer from the firms to the banks is low (negative). No market for corporate control emerges since the state guarantees the industrial firms (especially the large ones) against circuit breakdowns and insolvency.

By directing the mobilization and allocation of funds, state-controlled finance ensures greater stability of fund supply to end-users and producing firms but, for the same reason, it causes less efficient selection of investing enterprises and higher moral hazard from borrowers. The state role as a financial intermediary severs producers’ and investors’ decisions from the profit motive. This results in a highly leveraged corporate sector and in large and inefficient investments with poor output performances.

In more extreme forms, this situation is typical of corporatist and centrally planned-economies, where the circuit operates under state control. Liquidity for production is created at will by a centralized banking system, funds are funneled back to the state
through forced savings and taxation, and from there they are allocated to investing enterprises to finance capital goods purchases and ensure circuit closure.

Historically, both these economic models have proven effective in delivering the required output for limited periods of time, while they have irreversibly faded in the long-run. Their success is possible when the meta-economic objectives of the society (such as preparing for war, restoring order after periods of economic upheaval, or affirming the supremacy of central planning) are strong enough to mobilize the necessary individual efforts beyond commensurate individual economic rewards, and the output required from the collective effort addresses specific needs (e.g., building up heavy industry) that are accepted as absolute priorities by the collectivity.

As the two conditions weaken, both systems turn unable to preserve the necessary steam to production: the circuit is deprived of the capacity to allocate funds to best uses, and circuit closure can only be achieved at the cost of larger fiscal deficits and higher inflation.

**The “triumph of the small investor”**

The economic efficiency and welfare losses resulting from the failure of centralized models to reconcile individual efforts and rewards call for a correction of the economy’s incentive structure. This is accomplished by repositioning the circuit on a decentralized-decisions setting. Such route has been followed by the industrialized countries and a growing number of emerging economies over the last two decades, in particular in response to the need for restoring correct financial incentives.

Banks are allowed to compete in the markets for deposits and credit, to step into new markets, and to expand their range of services to clients. Nonbanks issue money substitutes and provide payment services. Capital markets develop owing to financial infrastructure development and entry of institutional investors. Competition leads the latter to calibrate their investment strategy on savers’ risk-return preferences, while efficient financial infrastructure allows small savers to directly enter the financial markets as investors (as is today the case in the United States). Small investors are increasingly well placed to influence, with their investment choices, the institutional investor strategies due

---

44 I take this title from Chernow (1997), which inspires much of this subsection.
to strong competition and investor-oriented laws and regulatory systems. They may rely on fast improving and increasingly safer technologies and infrastructure to move their own funds in the capital market at decreasing transaction costs and based on better quality information and best profit opportunities and underlying risks.

Financial decentralization raises both the risk sensitivity of small investors and their influence on the financial resource allocation process. Competitive financial investment institutions seek to achieve and retain all the flexibility needed to rapidly adjust their portfolios and investment strategies to changes in market conditions and small investors’ preferences. Information disclosure from market participants and attentive regulation and supervision of their behavior take hold as individual supply functions of investable funds become more sensitive to profit opportunities and risk factors. This complicates the survival of exclusive bank-firm relationships and, as financial market efficiency grows, informational advantages by individual agents are competed away and no longer assure them permanent extra-rents. Competition between banks and nonbanks increases. Also, synergies from joining banking and nonbanking financial intermediation functions increasingly lead to conglomeration or mergers of banks and nonbanks.

In a decentralized financial environment, funds can be more easily withdrawn from fund-users that are perceived to be riskier, bringing them under stronger market discipline. On the other hand, markets become subject to large price and funding volatility due to higher sensitivity of financing decisions to shifts in risk perceptions, new information, and changes in market opinion and sentiment. Thus, while financial resource allocation under “small investor” finance is more efficient than in alternative regimes, the circuit process is more vulnerable to a less stable supply of funds that may hinder its normal closure. Moreover, although the quality of both capital and output is higher, real investments may turn out to be more volatile and the spread of disturbances across sequential circuit rounds more likely to occur. Greater uncertainty may tighten credit to production and reduce bank flexibility in assisting illiquid firms at circuit-end.

Finally, in a “small investor” financial system, liberalized banking markets allow banks (especially large clearing banks) to regain greater seigniorage power than under state finance, although competition among banks and between banks and nonbanks in traditional banking activities may limit the seigniorage power of individual banks. On the other hand,
with developed capital markets, corporate ownership and control tend to shift to the wider public.

**Banks, finance, and economic development: what’s next?**

The “age of the small investor” is already a reality in the most advanced economies. It will be more and more pervasive as progress in knowledge, technology, and financial infrastructure enables larger masses of individual agents to take position in the financial markets. Of course, the age of the small investor lies still far ahead in the future of many developing countries that today lag much behind the post-industrial world. Yet late comers usually jump to their future much more rapidly than it took for the early pioneers to discover theirs. Also, late comers often skip intermediate stages, benefiting from the experience of their forerunners and cutting many corners.

Thus, trying to see how banking and financial intermediation might tend to evolve in the age of the small investor may well be in the spirit of searching for guidance to the future of many. With no pretence of articulating future comprehensive scenarios, one might anticipate one or two interesting developments just by looking at how finance will need to cope with the issue of trust in the new age.

The age of the small investor is the age of rapid progress in communications and information technologies. Indeed, small investors would not be relevant at all if it were not for the irresistible power that technology grants them at an increase pace. As technological development proceeds, thus, the options and opportunities available to investors continue to grow while transaction costs decline further.

What is the prospective role of banks and intermediaries in all this? Will they lose ground to direct finance, as savers and fund users are less and less inhibited by time, space, and information constraints to meet one another directly in the e-market arena?

A few observations. First, the openness and speed that the e-age allows individuals to enjoy in managing their choices generate larger networks where possible contacts and feasible options become countless. This is what creates opportunities; however, it is also where the impossibility for each network node to know the others comes into play as a critical limitation of direct transactions. Big numbers run counter the individuals’ need to overcome the natural mutual distrust that derives from not knowing each other well,
especially in the face of complex transactions and risks. And no matter how much information is available to establish the trustworthiness of potential counterparties to transactions and to assess risks, this information is hard to collect, assemble, and process; it requires time, resources, and skills. And, by the way: how can one be sure that the information available is reliable and its source trustworthy, to start with?

Second, transactions in the network require that the agents trust the network, that is, its infrastructure. The more immaterial the mode of exchange, the greater the agents’ need to make sure that immaterial transactions do not “dematerialize” their value for good!

Who’s going to bridge across these legitimate trust gaps? How can societies (especially in developing countries) benefit from the new technologies in the presence of large trust gaps?

There are basically two categories of institutions that will play a crucial role in making networks grow and become consolidated realities. They are, again, the banks and the financial intermediaries, with very important and specialized tasks to perform each.

**Banks**

In a world where trust matters more than ever, even less than in the conventional world can money creation be left to anybody to perform, although it was earlier argued that the incentive to enter this business will be stronger than under conventional banking. Also, contrary to the current trend toward domestic gross settlement payment systems around the globe, the high cost of such systems might push back the reintroduction of correspondent banking and netting arrangements (though strengthened as necessary to deal with systemic risks) or it might lead to the introduction of hybrid arrangements (see section 4 and also McAndrews, 1997).

Size will matter even more to banks. First, a large deposit base will be necessary for them to extract greater seigniorage in an environment where the increasing velocity of deposits and the larger circulation of quasi-mones make core deposits less relevant. Second, only large and highly visible banks with strong reputational capital will be able to attract depositors’ trust: the more money will be immaterial, the more people will want to know who produces it. Large size and large capital base will be crucial in both respects.

Crucial will also be the visibility that banks will be able to earn in the market as trustworthy institutions. Since the use of bank capital as a reputational signaling device
involves significant economies of scale (Hughes and Mester, 1998), banks will have one more incentive to reach for larger sizes in a world where reputation matters more.

Financial intermediaries

In a network-based world, where trust gaps may be formidable for the reasons discussed above, nonbank intermediaries play an essential role. They will be integral to the infrastructure serving individual agents and will provide the seal of trust necessary for individuals to use the networks. Their trust provision will be enhanced by the competition or contestability of the markets where they operate. More and more financial intermediaries will combine communications, information, and financial knowledge to supply agents with secure exchange mechanisms and reliable, high-quality information. They will produce and process the information needed to enable millions of anonymous individuals to interact and trade in the web, while their reputational capital and expertise will be necessary to validate the quality of the information exchanged. Nonbank intermediaries will be in a middleman position to bridge knowledge between savers and fund users and to best direct customers in counterparty selection and deal makings.

As networks bring in more participants and business opportunities, such knowledge will be useful for the intermediaries themselves to provide risk aggregation and diversification services that cannot be performed by individual agents, or that may be too costly for individuals to perform. It will also allow intermediaries to benefit simultaneously from the externalities associated with large numbers of customers and from the opportunity to package specialized and customized services based on the knowledge of customers’ individual preferences and needs.

In short, although in the future the circuit process will look increasingly sophisticated, its basic banking and financial functions will remain fundamental and fundamentally the same. Banks and nonbanks will become even more crucial as individuals take a more active interest in saving and investing.
Less banking, more intermediation

Far from making banking redundant, technological progress will nonetheless further the consolidation process of banks with nonbanks due to the impact of information on the financial sector structure.

In the early developmental stages, countries are characterized, inter alia, by poor financial infrastructure, including weak contract enforcement and deficient legal rules and systems, inadequate accounting/auditing practices and professions, non-existent investor protection guarantees, lacking mechanisms for asset resolution, costly information on business counterparties and trade opportunities, and backward telecommunication and information technologies.

Under these circumstances, transaction costs are prohibitive for individuals and nonbank institutions to engage in investment finance and to manage diversified financial portfolios. In such a context, banks can fill the gap between savers and capital users more efficiently than any other intermediary types, owing to their peculiar relationships with borrowers. Through such relationships, banks can supplement the lack of financial infrastructural services by accessing the information directly from the borrowers. This grants them with sufficient rent-extraction power which they use to access liquidity from depositors in exchange for a credible promise to transfer back to them a share of the rents extracted (see once more Diamond and Rajan 1999, 199b).

As financial infrastructure develops, the banks’ quasi-monopoly of information is broken and information is no longer concentrated within exclusive bank-borrower relationships. As a result, the value of the firms become known to the market and the banks lose their comparative advantage on extracting rents from capital users: banks lose markets to nonbanks and seek to recover such losses by entering the nonbanking business. Indeed, this seems to be actually taking place in the industrial world (Allen and Santomero, 1999).

At the same time, as competition in the nonbanking sector reduces net margins, nonbanks may have more than an incentive to merge and conglomerate their activities with banks in an attempt to gain some shares of their seigniorage from e-money creation.
8. Conclusions

This long journey through the theory of banking and finance has sought to show that banks are - and remain - special institutions with a fundamental role to play in market economies.

Over the last two decades, the mainstream literature on the theory of finance has offered significant contributions to the understanding of banks, identifying specific aspects that qualify them as special intermediaries. Contributions have differently focused on functions such as liquidity and payment services, credit supply, and information provision - all areas where banks have historically developed a comparative advantage vis-à-vis other types of intermediaries.

But the evolution of finance - through greater market competition, progress in technology, and institutional development - has increasingly eroded this comparative advantage. Not to mention that - as many foresee - the future success of electronic monies might deal a fatal blow to conventional banking, generating entirely new ways of doing finance.

This study has shown that, once money, production, and investment are looked at in an integrated way, banks and nonbank financial intermediaries can be seen as performing complementary functions essential to the economy.

The study has also shown that not recognizing the specialness of banks may lead to policies that place heavy - and perhaps unnecessary - burdens on real economic activity and diminish the market’s capacity to conjugate efficiency and risk optimally.

The study argues that banks remain central to the economy, irrespective of whether production financing loses ground as a source of core banking business.

The study has looked at the way in which banking and finance interrelate differently in the course of economic development, leading to different efficiency/stability configurations. Finally, noting that the expansion of communication possibilities will increasingly require societies to deal with the problem of overcoming mutual distrust among anonymous individuals, the study has argued that banks and financial intermediaries are uniquely placed to bridge trust gaps in the areas of money creation and fund intermediation.
The study finally predicts that banking will not only survive the advent of the e-money, but that it will play a main-actor role in producing it. It suggests that, as one recognizes the formidable incentive implicit in issuing money via lending, banking appears bound to remain an important reality, notwithstanding the technological transformation that money may undergo over time as a vehicle of purchasing power. In so far as technological progress speeds up the velocity at which virtual money circulates, the incentive to issue money through lending may increase manifold, carrying with it problems of poor credit quality and money overissue and, most of all, raising the prospect that any nonbank entities may want to act like banks and reap rents from seigniorage.

The study leads to conclude that, after all, the e-money will not change what is special about banks. Irrespective of its technological and material nature, bank money has always been “virtual” since it has always consisted of promises issued by entities specialized in attracting public trust. Any steps further into the era of e-finance will only reaffirm the inherent and attractive virtuality of money, and will always call on banks to give money a real content and to preserve it.
References


_____, 2000b, Do Banks Have a Future? A study on banking and finance as we move on into the Third Millennium. Mimeo.


______. 1990, The Theory of the Monetary Circuit, Économies et Sociétés, 24, 6, June, 7-36.


Schmitt B., 1959, L’Équilibre de la Monnaie, Revue d’Économie Politique.


